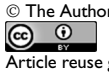
Original Manuscript

Social Psychological and Personality Science
1–12



DOI: 10.1177/19485506221099146
journals.sagepub.com/home/spp

Reducing Explicit Blatant Dehumanization by Correcting Exaggerated Meta-Perceptions

Alexander P. Landry[1], Jonathan W. Schooler[2], Robb Willer[1], and Paul Seli[3]

Abstract

If *explicitly, blatantly dehumanizing* a group of people—overtly characterizing them as less than human—facilitates harming them, then reversing this process is paramount. Addressing dehumanization among American political partisans appears especially crucial, given that it has been linked to their anti-democratic hostility. Perhaps because of its overt nature, partisans recognize—and greatly exaggerate—the extent to which out-partisans explicitly, blatantly dehumanize them. Past research has found that when people perceive they are dehumanized by an outgroup (i.e., *meta-dehumanization*), they respond with reciprocal dehumanization. Therefore, we reasoned that partisans' dehumanization could be reduced by correcting their exaggerated meta-dehumanization. Indeed, across three preregistered studies (N = 4,154), an intervention correcting American partisans' exaggerated meta-dehumanization reduced their own dehumanization of out-partisans. This decreased dehumanization persisted at a 1-week follow-up and predicted downstream reductions in partisans' anti-democratic hostility, suggesting that correcting exaggerated meta-dehumanization can durably mitigate the dark specter of dehumanization.

Keywords

dehumanization, political polarization, meta-perceptions, social distance, affective polarization

Explicitly, blatantly dehumanizing a group of people—overtly characterizing them as less than human–deprives them of the moral elevation distinguishing humans from "lower" animals and has been consistently linked to antisociality, discrimination, and aggression (Kteily & Landry, 2022). Explicit blatant dehumanization remains troublingly pervasive in the modern world, and not only in egregious instances of genocide (e.g., Smith, 2020). It also predicts anti-immigrant xenophobia (e.g., Landry, Ihm & Schooler, 2021; Markowitz & Slovic, 2020), animus toward racial outgroups (e.g., Jardina & Piston, 2021), and anti-democratic hostility between political partisans (Landry, Ihm, Kwit & Schooler, 2021; Moore-Berg et al., 2020). This has led prominent activists (Ames, 2019), human rights organizations (Heath, 2018), and scholars (Smith, 2020) to issue urgent calls for interventions addressing dehumanization's corrosive presence in contemporary intergroup relations.

Reducing Dehumanization by Targeting Meta-Dehumanization

The dark specter of dehumanization has inspired work attempting to mitigate it, most notably by fostering intergroup contact (see Capozza et al., 2014 for review), undermining humans' sense of superiority over nonhuman animals (Costello & Hodson, 2014), and individuating targets by highlighting their multiple social identities (e.g., Albarello et al., 2018). However, these interventions focused on rather subtle forms of dehumanization, such as the denial of "uniquely human" traits to outgroup members. Although subtle dehumanization likely plays an important role in intergroup relations, explicit blatant dehumanization is a consistently stronger predictor of hostility (Kteily & Landry, 2022). Therefore, more recent work has sought to reduce explicit blatant dehumanization (Gallardo et al., 2021; Moore-Berg et al., 2022).

One promising strategy draws on intergroup *meta-perceptions*, or how individuals believe members of an outgroup perceive their own group. This work begins from the assumption that humans are motivated to view the groups they identify with as superior to relevant outgroups (Tajfel & Turner, 1986). When people perceive an outgroup

[1]Stanford University, CA, USA
[2]University of California at Santa Barbara, USA
[3]Duke University, Durham, NC, USA

Corresponding Author:
Alexander P. Landry, Department of Organizational Behavior, Stanford University's Graduate School of Business, 655 Knight Way, Stanford, CA 94305, USA.
Email: alandry@stanford.edu

devalues their own group, they often respond with reciprocal negativity in an effort to restore their ingroup's relative status (e.g., Branscombe & Wann, 1994). Indeed, people who think an outgroup dehumanizes their own group–*meta-dehumanization*–experience social-identity threat and respond with reciprocal dehumanization (Kteily et al., 2016; Landry, Ihm & Schooler, 2022). Therefore, Kteily et al. (2016) reasoned that they could reduce Americans' dehumanization of Middle Eastern Muslims indirectly with an article suggesting that these Muslims recognized Americans' humanity. Indeed, this article not only reduced Americans' meta-dehumanization but also led them to dehumanize Muslims less in turn.

However, recent research did not find an effect of this intervention on reduced dehumanization–relative to a neutral control condition–in 4 of the 5 studies that examined this outcome (Pavetich & Stathi, 2021). This suggests important boundary conditions to the efficacy of meta-dehumanization interventions. For instance, meta-dehumanization interventions may be less efficacious (a) when participants' dehumanization of the outgroup is already relatively low–which was true of Pavetich and Stathi's samples (S. Stathi, personal communication, September 1, 2021)—or (b) when participants' meta-dehumanization is relatively accurate. These considerations inspired our attempt to reduce dehumanization in a context marked both by high levels of dehumanization and highly inaccurate meta-dehumanization.[1]

Exaggerated Meta-Dehumanization Among American Partisans

Substantial numbers of Democrats and Republicans explicitly, blatantly dehumanize one another (Cassese, 2021; Landry, Ihm, Kwit & Schooler, 2021; Moore-Berg et al., 2020). Perhaps because of its overt nature, partisans recognize that out-partisans explicitly, blatantly dehumanize them and even think they do so to a far greater extent than is actually the case. For instance, a nationally representative sample of Democrats and Republicans thought that out-partisans dehumanized them 186% more than they actually did (Moore-Berg et al., 2020). Such exaggerated meta-dehumanization is a particularly striking instance of a general tendency for individuals to attribute overly-negative intentions to outgroup members (Lees & Cikara, 2021).

Fortunately, overly-negative meta-perceptions are corrigible, and correcting them can improve intergroup relations (Lees & Cikara, 2020; Ruggeri et al., 2021). For instance, partisans who learned that out-partisans do not view their party's actions as negatively as they thought subsequently attributed less obstructionism to these out-partisans (Lees & Cikara, 2020). In this study, the researchers corrected partisans' meta-perceptions by presenting them with the actual (less negative) attitudes of the outgroup obtained from a prior study. This corrective approach may be particularly effective because it presents participants with an "objective" reality and nudges them to update their own meta-perceptions to match it, rather than asking them to directly change their attitudes about the outgroup (Lees & Cikara, 2021; Moore-Berg et al., 2020). Indeed, the efficacy of this intervention has been demonstrated in nine countries (Ruggeri et al., 2021).

Negative Affect and Meta-Prejudice. Although groups that are dehumanized are typically also disliked, dehumanization and dislike (i.e., *negative affect*) are empirically separable and have distinct neural underpinnings (Bruneau et al., 2018; see Kteily & Landry, 2022 for review). However, dehumanization and negative affect are tightly linked in partisan relations, so we sought to determine whether our intervention would reduce dehumanization while controlling for negative affect (Landry, Ihm, Kwit & Schooler, 2021; Moore-Berg et al., 2020). Moreover, just as dehumanization and dislike are distinct perceptions, feeling dehumanized and feeling disliked are distinct meta-perceptions that both contribute to intergroup conflict. And as with meta-dehumanization, partisans also overestimate how much the other side dislikes them (i.e., *meta-prejudice*; Landry, Ihm, Kwit & Schooler, 2021; Moore-Berg et al., 2020). Although we were primarily interested in correcting meta-dehumanization because previous work suggests that it inspires greater hostility than meta-prejudice (Landry, Ihm & Schooler, 2022), we also examined the effects of correcting meta-prejudice.

Anti-Democratic Hostility. Reducing partisans' dehumanization appears particularly crucial because it is linked to their *desire for social distance* from out-partisans—which impedes bipartisan unity (Iyengar et al., 2019)—and their *anti-democratic spite*—a willingness to subvert democratic norms to harm out-partisans (e.g., support for gerrymandering; Landry, Ihm, Kwit & Schooler, 2021; Moore-Berg et al., 2020). Therefore, if correcting partisans' exaggerated meta-dehumanization successfully reduces their own dehumanization of out-partisans, one may also expect to see downstream reductions in these indices of anti-democratic hostility. However, previous interventions that corrected partisans' meta-perceptions have had little influence on their anti-democratic attitudes (Voelkel et al., 2021; see also Broockman et al., 2020; Mernyk et al., 2021). Nonetheless, since dehumanization may be a particularly crucial driver of hostility (e.g., Landry, Ihm & Schooler, 2022), it is possible that a meta-perception correction reducing dehumanization would produce a downstream reduction in anti-democratic hostility.

The Present Research

We aimed to reduce American partisans' dehumanization by correcting their tendency to exaggerate how much out-

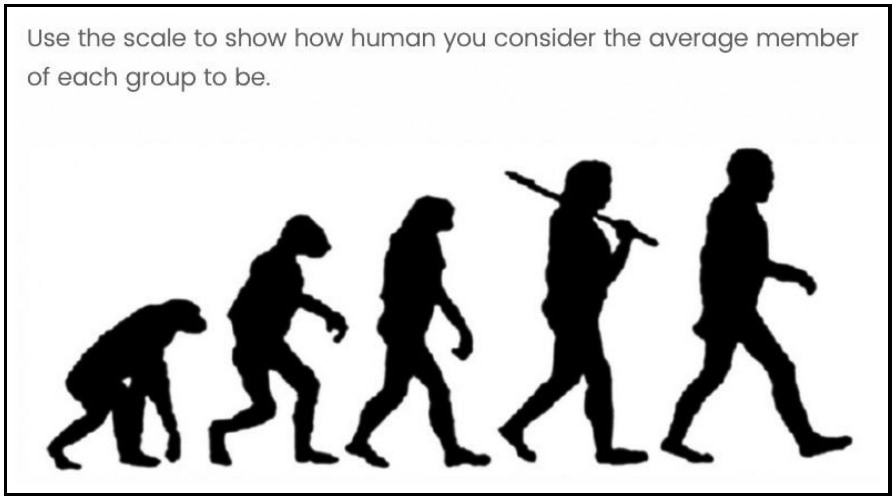


Figure 1. The "Ascent of Man" Measure of Dehumanization (Kteily et al., 2015).
Note. Participants rate how human they consider various groups to be on a slider scale ranging from 0 (leftmost primate ancestor) to 100 (fully modern human).

partisans dehumanize them. Our intervention first introduced partisans to the construct of dehumanization and a popular measure used to study it (the "Ascent of Man" scale; Kteily et al., 2015; Figure 1). They were then informed of the results from a recent study demonstrating exaggerated meta-dehumanization among American partisans (Moore-Berg et al., 2020, Study 1), namely that "Democrats and Republicans both rate the other side as more human than the other side thinks they do" (see Supplemental Material for the full intervention). We evaluated this intervention in three preregistered experiments (N = 4,154).

Our focal interest was whether the intervention reduced dehumanization. However, given that negative affect also explains unique variance in partisans' anti-democratic hostility (Landry, Ihm, Kwit & Schooler, 2021; Moore-Berg et al., 2020), we also investigated whether the intervention reduced negative affect in its own right. Moreover, we tested whether the intervention directly reduced two indices of anti-democratic hostility: partisans' desire for social distance and anti-democratic spite. Finally, we investigated whether reductions in anti-democratic hostility were mediated by parallel decreases in dehumanization and negative affect.

Study 1

Method

This study was preregistered at https://aspredicted.org/83N_23C.

Participants. Six-hundred and twenty participants were required to detect a small effect of our intervention (d = .2) with 80% power.[2] We initially recruited 1,450 participants on Amazon's Mechanical Turk.[3] After applying our preregistered exclusion criteria (see Procedure), we were left with a final sample of 1,107 participants (52.8% Republican; 47.2% Democrat[4]; M_{age} = 42.02, SD_{age} = 12.71; 81.8% White, 8.3% Black; 6.6% Asian, 3.3% other; 56.3% female, 43.5% male, 0.2% other). We were sensitive to detect an effect of d = .15 with 80% power.

Measures

Demographics. Participants reported their ethnicity, age, gender, income, education, and political affiliation (strong Democrat, strong Republican, moderate Democrat, moderate Republican, lean Democrat, lean Republican, Independent, or other).

Dehumanization. We followed previous research by creating a dehumanization composite from two measures (Kteily et al., 2015). One measure was the Ascent of Man scale (Figure 1). Participants rated the humanity of Democrats and Republicans (counterbalanced) using a slider scale ranging from 0 to 100. Scores were reversed such that higher scores indicated greater dehumanization. The other measure asked partisans to rate how well eight animalistic traits (e.g., "savage," "like animals") described out-partisans on a Likert-type scale ranging from 1 (*not at all*) to 7 (*extremely*; Democrats: a = .95, Republicans: a = .96). We created the dehumanization composite by standardizing these two measures and averaging them together (r = .68, $p < .001$).[5]

Negative Affect. We used a feeling thermometer (Haddock et al., 1993) to measure partisans' negative affect toward out-partisans. Participants rated how they felt toward Democrats and Republicans (counterbalanced) using a slider scale ranging from 0 (*extremely cold*) to 100 (*extremely warm*). Feelings toward out-partisans were

reversed such that higher scores indicated greater negative affect.

Desire for Social Distance. Participants responded to three items asking how comfortable they would feel if their doctor, child's teacher, and child's best friend was a member of the other party using a slider scale ranging from 0 (*not at all*) to 100 (*very*; Bogardus, 1947). We averaged these items together and reverse-scored them to obtain a measure of desired social distance from out-partisans (Democrats: a = .94; Republicans: a = .91).

Anti-Democratic Spite. We measured anti-democratic spite with four items[6] from Moore-Berg et al. (2020; Democrats: a = .77, Republicans: a = .79), using a Likert-type scale that ranged from 1 (*strongly disagree*) to 7 (*strongly agree*): "[Ingroup] should redraw districts to maximize their potential to win more seats in federal elections, even if it may be technically illegal"; "It's OK to sacrifice US economic prosperity in the short term in order to hurt [outgroup's] chances in future elections"; "[Ingroup] should do everything they can to hurt the [outgroup], even if it is at the short-term expense of the country"; and "[Ingroup] should do everything they can to limit the ability of [outgroup] to implement their preferred policies."

Procedure. Participants first provided demographic information and were directed to the end of the survey if they did not report affiliating with or leaning toward the Democratic or Republican Party (n = 78). Partisans were then randomly assigned to receive the meta-dehumanization intervention or to an empty control condition. Partisans receiving the intervention were asked two comprehension check questions and were directed to the end of the survey if they did not answer both of them correctly (n = 169).[7] Partisans then completed measures of dehumanization and negative affect (counterbalanced) and then anti-democratic spite and social distance (counterbalanced). We embedded an attention check item in the spite measure ("For quality purposes, please select 'Strongly disagree'") and excluded those who failed this check (n = 24). Finally, after completing the dependent measures, we asked participants whether they recommend we use their data and excluded those who said we "definitely" or "maybe" should not use their data (n = 22).[8]

Results

Descriptive statistics and reliabilities for the dependent measures in each study are presented in Supplemental Tables S2 to S3. Following our preregistered analysis plan, we first evaluated whether the intervention reduced partisans' dehumanization of out-partisans, controlling for their negative affect. We performed a one-way analysis of covariance (ANCOVA) with a binary variable denoting experimental condition (0 = control, 1 = intervention) as the between-subjects factor, scores on the dehumanization composite as the dependent variable, and negative affect as a covariate. Those receiving the intervention expressed significantly less dehumanization than those in the control condition, $M_{\text{Diff(Adj)}}$ = 0.32, F(1, 1,104) = 42.29, d = .40, SE = 0.05, 95% confidence interval (CI) = [0.22, 0.42], p < .001.[9]

We then performed three independent-samples t tests to determine whether the intervention reduced partisans' negative affect, desire for social distance, and anti-democratic spite. Indeed, the intervention significantly reduced negative affect, M_{Diff} = 6.68, t(1,105) = 4.28, d = .26, SE = 1.60, 95% CI = [3.66, 9.98], p < .001; social distance, M_{Diff} = 8.11, t(1,105) = 4.25, d = .26, SE = 1.82, 95% CI = [4.52, 11.85], p < .001; and spite, M_{Diff} = 0.17, t(1,105) = 2.13, d = .13, SE = 0.08, 95% CI = [0.18, 0.31], p = .024. Finally, we conducted two parallel mediations to test whether the intervention's effects on reduced (a) social distance and (b) anti-democratic spite were independently mediated by decreased dehumanization and negative affect.[10,11] The effect of the intervention on reduced desire for distance was indeed mediated by decreased dehumanization, β_{indirect} = −.18, SE = 0.03, 95% CI = [−0.22, −0.13], and negative affect, β_{indirect} = −.10, SE = 0.02, 95% CI = [−0.14, −0.05]. The effect of the intervention on reduced spite was also mediated by decreased dehumanization, β_{indirect} = −.21, SE = 0.03, 95% CI = [−0.26, −0.16], although the indirect effect through negative affect did not reach significance, β_{indirect} = −.01, SE = 0.01, 95% CI = [−0.03, 0.00].

Study 2

In Study 1, correcting partisans' exaggerated meta-dehumanization substantially reduced their own dehumanization of out-partisans. Moreover, the corrective intervention also tempered their negative affect and anti-democratic hostility. Although promising, these conclusions are limited by a potential confound: We selectively screened out participants in the intervention condition who failed the comprehension checks, which introduced differential attrition that could have accounted for these effects (ns = 634 vs. 473). Therefore, the primary goal of Study 2 was to replicate the effects of the meta-dehumanization intervention while also mitigating differential attrition.

In Study 2, along with the meta-dehumanization intervention, we also evaluated an intervention that corrected partisans' exaggerated meta-prejudice. This intervention informed partisans of another finding from Moore-Berg et al. (2020; Study 1), namely that "Democrats and Republicans both like the other side more than the other side thinks they do" (see Supplemental Materials for full

intervention). Comparing these interventions allowed us to determine their relative efficacy in reducing dehumanization and anti-democratic hostility.

Method

This study was preregistered at https://aspredicted.org/9ZN_9L5.

Participants. Seven-hundred and eighty-seven participants were required to detect a small effect of our intervention ($d = .2$) with 80% power. On the basis of exclusion rates from Study 1, we opened the study to 1,500 participants on Prolific Academic. After applying the same preregistered exclusion criteria as in Study 1 (see Study 1 Procedure), we were left with a final sample of 920 participants (71.6% Democrat; 28.4% Republican; $M_{age} = 45.84$, $SD_{age} = 16.55$; 80.5% White, 10.7% Black; 5.1% Asian, 3.7% other; 52.2% female, 47.1% male, 0.8% other). We were sensitive to detect an effect of $d = .18$ with 80% power.

Measures. All measures were identical to those in Study 1.

Procedure. The procedure was identical to Study 1 except where otherwise noted. Participants were randomly assigned to receive either the meta-dehumanization intervention, the meta-prejudice intervention, or to an empty control condition. To achieve parity in exclusion rates across conditions, participants in the control condition were randomly assigned to receive one of the interventions *after* completing the dependent measures and were then excluded on the basis of failed comprehension checks in the same manner as those in the other two conditions.

Results

We first tested for differential attrition. There was no relationship between binary variables contrasting the control condition with the intervention conditions and attrition status (control–meta-dehumanization contrast: $r = .04$, $p = .269$; control–meta-prejudice contrast, $r = -.06$, $p = .097$), suggesting differential attrition was indeed mitigated.

We proceeded with our preregistered analyses, first testing whether the interventions reduced dehumanization with a one-way ANCOVA. A multicategorical variable denoting experimental condition (0 = control, 1 = meta-prejudice intervention, 2 = meta-dehumanization intervention) was entered as the between-subjects factor, the dehumanization composite was the dependent variable, and negative affect was a covariate. A significant effect emerged, $F(2, 916) = 15.90$, $d = .36$, $p < .001$, with Tukey HSD planned contrasts confirming that the meta-dehumanization intervention led to less dehumanization than both the control condition, $M_{Diff(Adj)} = 0.36$, $SE = 0.05$, 95% CI = [0.23, 0.49], $p < .001$, and the meta-prejudice intervention, $M_{Diff(Adj)} = 0.24$, $SE = 0.06$, 95% CI = [0.12, 0.37], $p < .001$. Those receiving the meta-prejudice intervention also expressed marginally less dehumanization than the control condition, $M_{Diff(Adj)} = 0.12$, $SE = 0.07$, 95% CI = [−0.01, 0.25], $p = .065$.[12]

We performed another ANCOVA to determine whether the interventions decreased negative affect, entering this as the dependent variable and dehumanization as a covariate. We observed significant differences, $F(2, 916) = 16.33$, $d = .36$, $p < .001$, such that the meta-prejudice intervention led to less negative affect than both the control condition, $M_{Diff(Adj)} = 10.33$, $SE = 1.85$, 95% CI = [6.70, 13.96], $p < .001$, and the meta-dehumanization intervention, $M_{Diff(Adj)} = 6.76$, $SE = 1.86$, 95% CI = [3.12, 10.41], $p < .001$. Those receiving the meta-dehumanization intervention also expressed marginally less negative affect than the control condition, $M_{Diff(Adj)} = 3.56$, $SE = 1.93$, 95% CI = [−0.22, 7.35], $p = .065$.[13]

We then performed two analyses of variance (ANOVAs) to determine whether the interventions reduced partisans' anti-democratic hostility. We observed marginally significant differences between conditions on social distance, $F(2, 917) = 2.99$, $d = .10$, $p = .051$. The meta-dehumanization intervention led to marginally less desire for social distance than the control condition, $M_{Diff} = 4.47$, $SE = 2.38$, 95% CI = [−0.19, 9.14], $p = .06$. The meta-prejudice intervention also led to less desire for social distance than the control condition, $M_{Diff} = 5.31$, $SE = 2.31$, 95% CI = [0.77, 9.85], $p = .022$, and the intervention conditions did not differ in desire for social distance, $M_{Diff} = 0.84$, $SE = 2.34$, 95% CI = [−3.76, 5.43], $p = .721$. However, there were no significant differences among the conditions on anti-democratic spite, $F(2, 917) = 1.42$, $d = .05$, $p = .242$.

Finally, we conducted parallel mediations to test whether the interventions' effects on reducing desire for social distance were mediated by decreased dehumanization and negative affect. Indeed–relative to the control condition–the effect of the meta-dehumanization intervention on reduced social distance was mediated by decreased dehumanization, $\beta_{indirect} = -.21$, $SE = 0.03$, 95% CI = [−0.28, −0.15], and negative affect, $\beta_{indirect} = -.13$, $SE = 0.03$, 95% CI = [−0.20, −0.08]. Likewise, the effect of the meta-prejudice intervention on reduced social distance was mediated by decreased dehumanization, $\beta_{indirect} = -.14$, $SE = 0.03$, 95% CI = [−0.20, −0.07], and negative affect, $\beta_{indirect} = -.20$, $SE = 0.03$, 95% CI = [−0.27, −0.14].

Study 3

In Studies 1 and 2, correcting partisans' exaggerated meta-dehumanization reduced their own dehumanization, negative affect and desire for social distance from out-partisans.

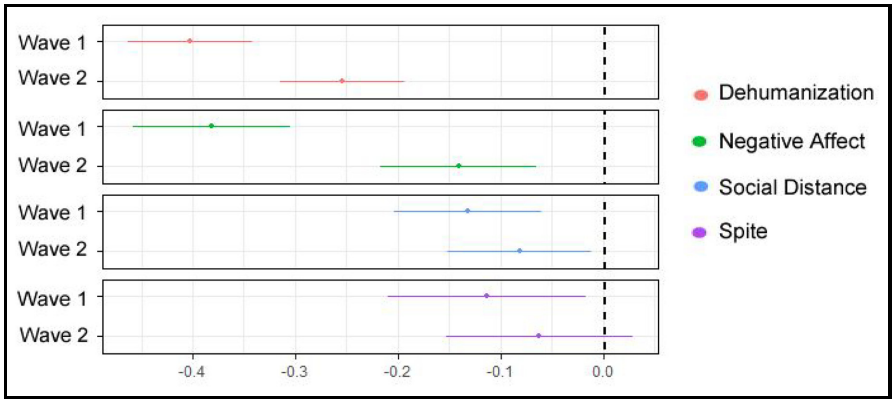


Figure 2. Estimated Effects of the Meta-Dehumanization Intervention on Outcome Variables.
Note. The X-axis represents the estimated effect of the intervention relative to the control condition. All scales were recoded to range from 0 to 1. 95% confidence intervals that overlap the dashed vertical line at zero represent a failure to reject the null hypothesis significance test of zero effect.

In Study 2 we observed similar effects for a meta-prejudice intervention, further demonstrating the promise of correcting overly-negative meta-perceptions. Nonetheless, our failure to find an effect of either intervention on anti-democratic spite in Study 2 warrants further investigation. The generalizability of our findings is also limited because we employed convenience samples in Studies 1 and 2. Most importantly, it is uncertain whether the effects of the meta-dehumanization intervention endure for any meaningful length of time–especially considering that the immediate impact of a similar intervention did not persist at a 1-week follow-up (Lees & Cikara, 2020). Therefore, we conducted Study 3 on a highly-powered, nationally-representative sample of American partisans and employed a longitudinal design to evaluate whether the meta-dehumanization intervention's effects endured 1 week after it was administered.

Method

Wave 1 of this study was preregistered at https://aspredicted.org/1XR_9PP. Wave 2 was preregistered at https://aspredicted.org/N52_TX2.

Participants. We recruited a representative sample of Democrats and Republicans through the ForthRight panel managed by Bovitz Inc (see Supplemental Material for information about this panel). To maximize power and account for attrition at Wave 2, we preregistered a Wave 1 target sample of 2,500 and initially recruited 3,589 participants. After applying our preregistered exclusion criteria (see Procedure), we were left with 2,513 participants. Of these, 84.6% completed Wave 2, leaving us with a final sample of 2,127 (M_{age} = 49.05, SD_{age} = 15.88; 48.6% male, 50.4% female, 1% nonbinary; 76.9% White, 10.5% Black, 12.1% Hispanic, 3.9% Asian, 3.2% other/prefer not to answer). The final sample approximated population demographics of the respective parties on gender, age, and race/ethnicity (see Supplemental Table S4 for a comparison to 2016 American National Election Survey demographic benchmarks). There was no differential attrition between the experimental conditions from Wave 1 to Wave 2. We were sensitive to detect an effect of f^2 = .01 with 95% power.

Measures. All measures were identical to those in Studies 1 and 2.

Procedure. At Wave 1 participants were first asked an attention agreement ("Do you agree to pay attention and carefully read all questions?") Those who did not agree (n = 23) were directed to the end of the survey. Participants then completed demographic information and were directed to the end of the survey if they did not report affiliating with, or leaning toward, the Democratic or Republican Party (n = 510). We then tested the reading comprehension of all participants with a mock vignette and directed them to the end of the survey if they failed a comprehension question (n = 543). Participants were then randomly assigned to receive the meta-dehumanization intervention[14] or to an empty control condition. They then completed the dependent measures as in Studies 1 and 2. At Wave 2 (M = 6.42 days later), all participants again completed the dependent measures in the same manner.

Results

Because we recruited a more demographically diverse sample in this study, we preregistered analyses using ordinary least squares (OLS) models controlling for party identification, political ideology, age, gender, race, and education. Figure 2 shows the estimated effects of the meta-dehumanization intervention—relative to the control

condition—on dehumanization, negative affect, desire for social distance, and anti-democratic spite. After receiving the intervention at Wave 1, partisans reported less dehumanization than those in the control condition, $\beta = -.27$, $d = .56$, $SE = 0.04$, $t = -12.87$, $p < .001$. The intervention also reduced partisans' Wave 1 negative affect, $\beta = -.21$, $d = .43$, $SE = 1.16$, $t = -9.77$, $p < .001$; desire for social distance, $\beta = -.08$, $d = .16$, $SE = 1.27$, $t = -3.63$, $p < .001$; and anti-democratic spite, $\beta = -.05$, $d = .10$, $SE = 0.06$, $t = -2.51$, $p = .012$.

Durability Test. We then tested the durability of these effects by examining the Wave 2 measures (administered $M = 6.42$ days later; $SD = 2.08$).[15] The intervention's effect on reduced dehumanization persisted at Wave 2, $\beta = -.16$, $d = .32$, $SE = 0.04$, $t = -7.64$, $p < .001$.[16] The intervention also reduced partisans' Wave 2 negative affect, $\beta = -.08$, $d = .16$, $SE = 1.21$, $t = -3.68$, $p < .001$, and desire for social distance, $\beta = -.05$, $d = .10$, $SE = 1.32$, $t = -2.21$, $p = .027$. However, the effect of the intervention on reduced anti-democratic spite did not persist at Wave 2, $\beta = -.03$, $d = .06$, $SE = 0.06$, $t = -1.52$, $p = .13$.

Mediation Analyses. Finally, we conducted parallel mediations to test whether the intervention's effects on reducing partisans' desire for social distance at both waves, and reducing their anti-democratic spite at Wave 1, were mediated by decreased dehumanization and negative affect. Indeed, the effect of the intervention on reduced Wave 1 social distance was mediated by reductions in Wave 1 dehumanization, $\beta_{indirect} = -.22$, $SE = 0.02$, 95% CI = [−0.26, −0.18], and negative affect, $\beta_{indirect} = -.11$, $SE = 0.01$, 95% CI = [−0.14, −0.08]. Similarly, the effect of the intervention on reduced Wave 2 social distance was mediated by reductions in Wave 2 dehumanization, $\beta_{indirect} = -.13$, $SE = 0.02$, 95% CI = [−0.17, −0.10], and negative affect, $\beta_{indirect} = -.04$, $SE = 0.01$, 95% CI = [−0.07, −0.02]. Moreover, the intervention's effect on reduced anti-democratic spite at Wave 1 was mediated by reduced dehumanization, $\beta_{indirect} = -.24$, $SE = 0.02$, 95% CI = [−0.29, −0.20], and negative affect, $\beta_{indirect} = -.04$, $SE = 0.01$, 95% CI = [−0.06, −0.02].

Intervention Effects by Party

In Studies 1 and 3, we also evaluated the efficacy of the intervention for Democrats and Republicans separately (we did not do so in Study 2 due to the small number of Republican participants). The intervention appeared to be particularly effective for Democrats, reducing their dehumanization, negative affect, desire for social distance, and anti-democratic spite in Study 1 and both waves of Study 3. However, while the intervention also reliably reduced Republicans' dehumanization, these effects were smaller than those observed for Democrats. Moreover, while the intervention decreased Republicans' negative affect and desire for social distance in Study 1 and Wave 1 of Study 3, these reductions did not persist to Wave 2 of Study 3. The intervention had no effect on Republicans' anti-democratic spite in either study. See Supplemental Material for detailed reporting of these results.

Internal Meta-Analysis

We conducted an internal meta-analysis following procedures specified by McShane and Böckenholt (2017). To test the aggregate effect of the meta-dehumanization intervention on reduced dehumanization, we analyzed the Ascent and animalistic trait measures separately because analyzing the standardized composite measure would prevent us from accurately assessing method heterogeneity across studies (B. McShane, personal communication, December 29, 2021). Across studies,[17] the intervention reduced dehumanization relative to a control condition on both the Ascent, *effect* = −15.95, $SE = 1.70$, 95% CI = [−19.28, −12.63], and animalistic trait measures, *effect* = −0.66, $SE = 0.08$, 95% CI = [−0.81, −0.52]. Heterogeneity across methods of studies was very high (Ascent: $I^2 = 97.48\%$, 95% CI = [95.94%, 98.43%]; animalistic trait: $I^2 = 93.48\%$, 95% CI = [87.69%, 96.54%]), suggesting that method factors accounted for almost all of the total variation observed across the data beyond that attributable to the experimental manipulation. See Figure 3 for the meta-analytic effect estimates of the intervention on reduced dehumanization.

In addition, the intervention reduced negative affect across studies, *effect* = −9.21, $SE = 1.51$, 95% CI = [−12.17, −6.24]. Method heterogeneity was high, $I^2 = 85.78\%$, 95% CI = [68.76, 93.53], although the relatively wide CI suggests that heterogeneity was not estimated precisely. Crucially, the intervention also reduced partisans' desire for social distance and anti-democratic spite across studies (social distance: *effect* = −5.53, $SE = 0.97$, 95% CI = [−7.44, −3.63]; anti-democratic spite: *effect* = −0.15, $SE = 0.04$, 95% CI = [−0.23, −0.06]). Method heterogeneity across studies for social distance was estimated imprecisely, $I^2 = 0\%$, 95% CI = [0%, 74.24%], suggesting that methods factors accounted for anything from a trivial to large amount of the variance observed across the data beyond that attributable to the experimental manipulation. Method heterogeneity for anti-democratic spite was high and estimated precisely, $I^2 = 95.86\%$, 95% CI = [92.79%, 97.62%].

Discussion

Explicit blatant dehumanization continues to mar contemporary intergroup relations (Kteily & Landry, 2022). For instance, a troubling number of American partisans

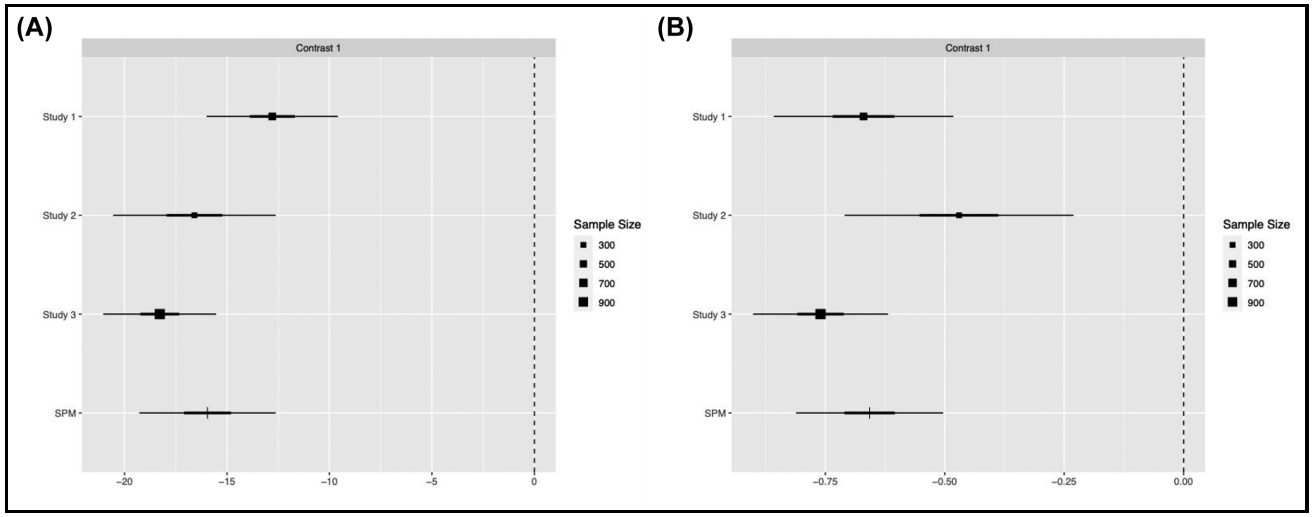


Figure 3. Effect of the Meta-Dehumanization Intervention on Reduced (A) Ascent and (B) Animalistic Trait Dehumanization Across Studies. *Note*. The point estimates for each study and the overall estimate for the Single Paper Meta-analysis (SPM) are given by the squares; 50% and 95% confidence intervals are given by the thick and thin lines, respectively; 95% confidence intervals that overlap the dashed vertical line at zero represent a failure to reject the null hypothesis significance test of zero effect.

explicitly, blatantly dehumanize one another, which has been linked to their anti-democratic hostility (e.g., Moore-Berg et al., 2020). We sought to reduce partisan dehumanization by integrating research demonstrating that (a) individuals who think an outgroup dehumanizes their own group (i.e., *meta-dehumanization*) respond with reciprocal dehumanization (Kteily et al., 2016; Landry, Ihm & Schooler, 2022) and (b) individuals attribute overly-negative attitudes to outgroup members (Lees & Cikara, 2021). We developed an intervention informing American partisans of their tendency to overestimate how much they are dehumanized by out-partisans (Landry, Ihm, Kwit & Schooler, 2021; Moore-Berg et al., 2020). This reduced partisans' own dehumanization of out-partisans across three studies–an effect that persisted at a 1-week follow-up.

Correcting partisans' meta-dehumanization also produced modest—yet reliable—reductions in their anti-democratic hostility. This is notable given recent work finding that interventions reducing negative affect do not influence anti-democratic attitudes (Broockman et al., 2020; Voelkel et al., 2021). Perhaps our dehumanization-focused intervention reduced anti-democratic attitudes when affect-focused interventions did not because dehumanization is more strongly linked to anti-democratic attitudes. Indeed, we observed particularly strong indirect effects of the intervention on reduced anti-democratic spite through dehumanization (average $\beta_{indirect} = -.23$, compared to an average $\beta_{indirect} = -.03$ for negative affect; see also Landry, Ihm & Schooler, 2022).[18] Although experimental tests of mediation are needed to confirm this cross-sectional indirect effect, future work attempting to bolster support for democratic norms should consider the promise of targeting dehumanization.

Limitations and Future Directions

Although we observed promising reductions in anti-democratic hostility, we do not wish to overstate their relatively small magnitude. Correcting exaggerated meta-dehumanization produced far stronger reductions in reciprocal dehumanization than these downstream outcomes, which is consistent with previous work suggesting the link between inaccurate meta-perceptions and attitudes is largely domain-specific (Mernyk et al., 2021; Voelkel et al., 2021). To bolster their effects on downstream outcomes, researchers may combine interventions correcting (meta-) misperceptions with those inducing prosocial emotions. For instance, combining an article correcting overly negative stereotypes about enemy soldiers with a video inducing empathy toward them not only reduced the dehumanization of these soldiers but also led to downstream reductions in acceptance of civilian casualties (Landry et al., 2022; see also Moore-Berg et al., 2022).

Encouragingly, the modest reductions in anti-democratic hostility we observed could perhaps be amplified through positive feedback because the relationships between meta-dehumanization, dehumanization, and hostility are mutually reinforcing (Kteily et al., 2016). For instance, the intervention's effect on decreasing partisans' desire for social distance could feed back to further dampen their (meta-)dehumanization toward out-partisans because intergroup contact reduces these attitudes (Bruneau et al., 2021; Capozza et al., 2017). Longitudinal research should examine

this possibility. Future work should also address a key limitation of our design: since we compared our intervention with an empty control condition, we can't definitively establish whether the effect of the intervention was due specifically to the meta-dehumanization correction or to the fact that (meta-)dehumanization was mentioned to participants at all (but see the comprehension check analysis in Supplemental Material for evidence that both played some role).

Most crucially, applied work needs to more widely disseminate the fact that Democrats and Republicans consider one another more human than the other side thinks. This could be done through popular media such as radio or television programs (e.g., Blair et al., 2019; Paluck, 2009), which may reverse the process whereby media contributes to political (meta-)misperceptions and inflames hostility (e.g., Garrett et al., 2019; Levendusky & Malhotra, 2016). In a similar vein, we consider it worthwhile to investigate boundary conditions to our intervention's efficacy (e.g., why the intervention was more effective for Democrats than Republicans) and examine whether these results generalize to other contexts marked by high degrees of (meta-)dehumanization (e.g., Hungarians who feel dehumanized by the Roma; Kteily et al., 2016).

Conclusion

We integrated research identifying (a) meta-dehumanization as a precursor to dehumanization and (b) overly-negative meta-perceptions as drivers of intergroup conflict. By correcting partisans' exaggerated meta-dehumanization, we durably reduced their own explicit blatant dehumanization. The intervention also spilled over to reduce anti-democratic hostility, suggesting dehumanization may be a particularly crucial driver of hostility (vis-à-vis negative affect; cf. Broockman et al., 2020; Voelkel et al., 2021). This brief, cost-effective intervention lays a promising foundation to temper the rising partisan enmity in America (Iyengar et al., 2019) and should inspire future work to develop scalable means of correcting exaggerated meta-dehumanization. Doing so may help to lift dehumanization's dark specter from intergroup relations.

Acknowledgment

We thank Bastian Moritz Weitz for improving this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

We thank the Stanford Center on Philanthropy and Civil Society for helping to fund this research.

ORCID iD

Alexander P. Landry https://orcid.org/0000-0003-1579-7326

Data Availability

Data and analysis code files have been deposited in Open Science Framework (https://osf.io/pcx53/?view_only = 0b205e62c37749d2b 8d1db74b445a027).

Supplemental Material

Supplemental material is available online with this article.

Notes

1. Regarding high levels of dehumanization, participants in the control condition of our studies scored 97% higher on a measure of outgroup dehumanization than those in the control condition of Pavetich and Stathi's (2021) studies. Regarding highly inaccurate meta-dehumanization—as noted in the main text—samples were drawn from our population of interest consistently believe the outgroup dehumanizes them 2 to 3 times more than is actually the case (Landry et al., 2021a; Moore-Berg et al., 2020).
2. Power and sensitivity analyses were performed with *G*Power* (Faul et al., 2007).
3. We attempted to collect 620 responses in line with our preregistration, but an error on behalf of our participant recruitment platform (CloudResearch; Litman & Robinson, 2020) resulted in a larger-than-anticipated sample size.
4. In this and Study 3, we also evaluated the efficacy of the intervention for Democrats and Republicans separately. The intervention showed similar effects across party lines, although it appeared to be somewhat more effective for Democratic participants. See Supplemental Material for detailed reporting of these results.
5. The decision to report our results using the composite, instead of each measure separately, was a deviation from our preregistrations made in the interests of concision and minimizing Type II error. However, in this and all subsequent studies, the same patterns of results emerged when examining the measures separately.
6. We initially used a six-item measure of what we referred to as "political violence" in Studies 1 and 2. However, upon further consideration, these items appeared to measure distinct constructs–namely general anti-democratic tendencies (Item 1: "[Ingroup] should do whatever is necessary to pass policies, even if it seems undemocratic"), support for violence (Item 2: "It is justified for [ingroup] to use violence in advocating their political goals"), and anti-democratic spite (Items 3–6). Therefore, we limited our primary analyses to the 4 items directly measuring spite. However, in both Studies 1 and 2, the same pattern of results emerged when using the six-item measure.
7. See Supplemental Table S1 for proportions of correct responses.
8. In this and Study 2, the same pattern of results emerged when including participants who failed the attention check and/or data quality question.

9. The intervention also reduced dehumanization without controlling for negative affect in an independent-samples t test, $M_{Diff} = 0.42$, $t(1,105) = 8.11$, $d = .48$, $SE = 0.05$, 95% confidence interval [CI] = [0.32, 0.53], $p < .001$.
10. We performed all mediation analyses with the PROCESS macro (Hayes, 2017)
11. See Supplemental Figures S1–S3 for path diagrams of all mediations reported in this article.
12. We also performed a one-way analysis of variance (ANOVA) on dehumanization without controlling for negative affect, $F(2, 917) = 24.70$, $d = .46$, $p < .001$. The same pattern of results emerged: The meta-dehumanization intervention reduced dehumanization relative to the control condition, $M_{Diff} = 0.49$, $SE = .07$, 95% CI = [.35, .63], $p < .001$, and the meta-prejudice intervention, $M_{Diff} = 0.18$, $SE = .07$, 95% CI = [.04, .31], $p = .011$. The meta-prejudice intervention also reduced dehumanization relative to the control condition, $M_{Diff} = 0.31$, $SE = .07$, 95% CI = [.18, .45], $p < .001$.
13. We also performed a one-way ANOVA on negative affect without controlling for dehumanization, $F(2, 917) = 25.14$, $d = .46$, $p < .001$. The same pattern of results emerged: The meta-prejudice intervention reduced negative affect relative to the control condition, $M_{Diff} = 13.96$, $SE = 2.00$, 95% CI = [10.05, 17.88], $p < .001$, and the meta-dehumanization intervention, $M_{Diff} = 4.71$, $SE = 2.02$, 95% CI = [0.75, 8.67], $p = .02$. The meta-dehumanization intervention also reduced negative affect relative to the control condition, $M_{Diff} = 9.26$, $SE = 2.05$, 95% CI = [5.23, 13.28], $p < .001$.
14. We made a minor change to the meta-dehumanization intervention in this study. The intervention used in Studies 1 and 2 informed participants that "Democrats and Republicans think each other are equally human" on the basis of results from Moore-Berg et al. (2020). However, subsequent research found asymmetries in partisan dehumanization (Landry et al., 2021a), so we removed this piece of information and the accompanying comprehension question.
15. None of the effects of the intervention on Wave 2 dehumanization, negative affect, social distance, or spite were moderated by time between waves. That is, the effects were not stronger among those who completed the follow-up sooner. See Supplemental Materials for these analyses.
16. The effects of the intervention on reduced dehumanization at both waves were robust when we entered negative affect in the model as an additional covariate (Wave 1: $\beta = -.20$, $SE = 1.33$, $t = -10.02$, $p < .001$; Wave 2: $\beta = -.15$, $SE = 1.34$, $t = -7.49$, $p < .001$).
17. We did not include Wave 2 of Study 3 in the analysis because the measurements at this later time point are not comparable to those administered immediately after the intervention (B. McShane, personal communication, December 29, 2021). Likewise, we only compare the meta-dehumanization intervention to the control condition in each study and not to the meta-prejudice intervention condition in Study 2.
18. In zero-order terms, dehumanization was also more strongly associated with anti-democratic spite than negative affect was (average rs = .49 vs. .30).

References

Albarello, F., Crisp, R., & Rubini, M. (2018). Promoting beliefs in the inalienability of human rights by attributing uniquely human emotions through multiple categorization. *The Journal of Social Psychology*, *158*(3), 309–321. https://doi.org/10.1080/00224545.2017.1346581

Ames, B. (2019, September 11). *The dehumanization of immigrants and the rise of the extreme right*. AICGS. https://www.aicgs.org/publication/the-dehumanization-of-immigrants-and-the-rise-of-the-extreme-right/

Blair, G., Littman, R., & Paluck, E. L. (2019). Motivating the adoption of new community-minded behaviors: An empirical test in Nigeria. *Science Advances*, *5*(3), 1–8. https://doi.org/10.1126/sciadv.aau5175

Bogardus, E. S. (1947). Measurement of personal-group relations. *Sociometry*, *10*(4), 306–311. https://doi.org/10.2307/2785570

Branscombe, N. R., & Wann, D. L. (1994). Collective self-esteem consequences of outgroup derogation when a valued social identity is on trial. *European Journal of Social Psychology*, *24*(6), 641–657. https://doi.org/10.1002/ejsp.2420240603

Broockman, D., Kalla, J., & Westwood, S. (2020). *Does affective polarization undermine democratic norms or accountability? Maybe not*. OSF Preprints. https://doi.org/10.31219/osf.io/9btsq

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., Jacoby, N., Kteily, N., & Saxe, R. (2018). Denying humanity: The distinct neural correlates of blatant dehumanization. *Journal of Experimental Psychology: General*, *147*(7), 1078–1093. https://doi.org/10.1037/xge0000417

Capozza, D., Bernardo, G. A. D., & Falvo, R. (2017). Intergroup contact and outgroup humanization: Is the causal relationship uni- or bidirectional? *PLOS ONE*, *12*(1), Article e0170554. https://doi.org/10.1371/journal.pone.0170554

Capozza, D., Falvo, R., Di Bernardo, G. A., Vezzali, L., & Visintin, E. (2014). Intergroup contact as a strategy to improve humanness attribution: A review of studies. *TPM—Testing, Psychometrics, Methodology in Applied Psychology*, *3*, 349–362. https://doi.org/10.4473/TPM21.3.9

Cassese, E. C. (2021). Partisan dehumanization in American politics. *Political Behavior*, *43*(1), 29–50. https://doi.org/10.1007/s11109-019-09545-w

Costello, K., & Hodson, G. (2014). Explaining dehumanization among children: The interspecies model of prejudice. *British Journal of Social Psychology*, *53*(1), 175–197. https://doi.org/10.1111/bjso.12016

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

Gallardo, R. A., Hameiri, B., Moore-Berg, S. L., & Bruneau, E. (2021). The collective praise intervention: A brief intervention highlighting prosocial behavior reduces hostility towards Muslims. *Group Processes & Intergroup Relations*, https://doi.org/10.1177/13684302211021665

Garrett, R. K., Long, J. A., & Jeong, M. S. (2019). From partisan media to misperception: Affective polarization as mediator. *Journal of Communication*, *69*(5), 490–512. https://doi.org/10.1093/joc/jqz028

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

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis, second edition: A regression-based approach*. Guilford Publications.

Heath, D. (2018, July 27). *Politics of dehumanization in the United States*. Human Rights Watch. https://www.hrw.org/news/2018/07/27/politics-dehumanization-united-states

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Jardina, A., & Piston, S. (2021). Hiding in plain sight: Dehumanization as a foundation of white racial prejudice. *Sociology Compass*, *15*(9), 1–15. https://doi.org/10.1111/soc4.12913

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

Landry, A., Protzko, J., Barr, N., Ariely, D., & Seli, P. (2022). Integrating theoretical perspectives on dehumanization to explain–and reduce–moral violence amidst the Russo-Ukrainian war. *PsyArXiv*. https://doi.org/10.31234/osf.io/jth5z

Landry, A. P., Ihm, E., Kwit, S., & Schooler, J. W. (2021). Metadehumanization erodes democratic norms during the 2020 presidential election. *Analyses of Social Issues and Public Policy*, *21*, 51–63. https://doi.org/10.1111/asap.12253

Landry, A. P., Ihm, E., & Schooler, J. W. (2021). Filthy animals: Integrating the behavioral immune system and disgust into a model of prophylactic dehumanization. *Evolutionary Psychological Science*. https://doi.org/10.1007/s40806-021-00296-8

Landry, A. P., 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

Lees, J., & Cikara, M. (2020). Inaccurate group meta-perceptions drive negative out-group attributions in competitive contexts. *Nature Human Behaviour*, *4*(3), 279–286. https://doi.org/10.1038/s41562-019-0766-4

Lees, J., & Cikara, M. (2021). Understanding and combating misperceived polarization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *376*(1822), 20200143. https://doi.org/10.1098/rstb.2020.0143

Levendusky, M., & Malhotra, N. (2016). Does media coverage of partisan polarization affect political attitudes? *Political Communication*, *33*(2), 283–301. https://doi.org/10.1080/10584609. 2015.1038455

Litman, L., & Robinson, J. (2020). *Conducting online research on Amazon mechanical Turk and beyond*. SAGE.

Markowitz, D. M., & Slovic, P. (2020). Social, psychological, and demographic characteristics of dehumanization toward immigrants. *Proceedings of the National Academy of Sciences*, *117*(17), 9260–9269. https://doi.org/10.1073/pnas.1921790117

McShane, B. B., & Böckenholt, U. (2017). Single-paper meta-analysis: Benefits for study summary, theory testing, and replicability. *Journal of Consumer Research*, *43*(6), 1048–1063. https://doi.org/10.1093/jcr/ucw085

Mernyk, J., Pink, S., Druckman, J., & Willer, R. (2021). *Correcting inaccurate metaperceptions reduces Americans' support for Partisan violence*. OSF Preprints. https://doi.org/10.31219/osf.io/8ftsx

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

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology*, *96*(3), 574–587. https://doi.org/10.1037/a0011989

Pavetich, M., & Stathi, S. (2021). Meta-humanization reduces prejudice, even under high intergroup threat. *Journal of Personality and Social Psychology*, *120*(3), 651–671. https://doi.org/10.1037/pspi0000259

Ruggeri, K., Većkalov, B., Bojanić, L., Andersen, T. L., Ashcroft-Jones, S., Ayacaxli, N., Barea-Arroyo, P., Berge, M. L., Bjørndal, L. D., Bursalıoğlu, A., Bühler, V., Čadek, M., Çetinçelik, M., Clay, G., Cortijos-Bernabeu, A., Damnjanović, K., Dugue, T. M., Esberg, M., Esteban-Serna, C., & Folke, T. (2021). The general fault in our fault lines. *Nature Human Behaviour*, *5*(10), 1369–1380. https://doi.org/10.1038/s41562-021-01092-x

Smith, D. L. (2020). *On inhumanity: Dehumanization and how to resist it*. Oxford University Press.

Tajfel, H., & Turner, J. (1986). The social identity theory of intergroup behavior. In S. Worchel, & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Voelkel, J. G., Chu, J., Stagnaro, M., Mernyk, J., Redekopp, C., Pink, S., Druckman, J., Rand, D., & Willer, R. (2021). *Interventions reducing affective polarization do not improve anti-democratic attitudes*. OSF Preprints. https://doi.org/10.31219/osf.io/7evmp

Author Biographies

Alexander Landry is a PhD student who studies the social psychology of intergroup conflict at Stanford's Graduate School of Business.

Jonathan Schooler is a distinguished professor of psychology at the University of California, Santa Barbara. He is a leader in the study of consciousness, memory, and mind-wandering.

Robb Willer is a professor of sociology at Stanford University. His research looks at the social and psychological forces shaping Americans' political attitudes, and the development of techniques for overcoming political polarization.

Paul Seli is a professor of psychology and Neuroscience at Duke University, where he studies political polarization, consciousness, mind-wandering, memory, and creativity.

Handling Editor: Danny Osborne