

## Table of Contents

Supplement to Image Generation and Image Rating Methods Section-----	p. 2
Study Designs and Power for Main Effects and Interactions across All Models-----	p. 2
Evidence that Raters' Anti-Arab Biases Influence Their Ratings-----	p. 3
Supplement to Link Between Self-Reported and Visual Dehumanization-----	p. 4
Group Images by Generators' Reported Levels of Dehumanization-----	p. 4
From MTurk sample-----	p. 4
From Qualtrics Panels sample-----	p. 5
A Full Analysis from Group Image Rating Studies-----	p. 5
From MTurk sample-----	p. 6
From Qualtrics Panels sample-----	p. 6
Supplement to Link Between Political Leanings and Visual Dehumanization-----	p. 7
Group Images by Generators' Political Ideology Levels-----	p. 7
From MTurk sample-----	p. 7
From Qualtrics Panels sample-----	p. 8
Self-Reported Dehumanization among Liberals and Conservatives-----	p. 8
From MTurk sample-----	p. 8
From Qualtrics Panels sample-----	p. 9
Decomposition of Marginal Three-Way Interaction-----	p. 9
Supplement to Exploratory Analyses from the Qualtrics Panels Sample-----	p. 10
Exploratory Analyses on Moderators of Visual Dehumanization-----	p. 10
Exploratory Analyses on Pixel Correlations-----	p. 12

## Supplement to Image Generation and Image Rating Methods Sections

### Study Designs and Power for Main Effects and Interactions across All Models

Our manuscript reports the results from seven rating studies that have different experimental designs (see Table S1), but all of which allow us to test for a) a main effect of target (Arab, American) on how dehumanizing mental representations are rated, and b) whether this main effect interacts with other features of the people who generated them (e.g., whether generators came from the MTurk or Qualtrics Panels sample, generators' political leanings, generators' explicit dehumanization levels of Arabs). Because we did not want to overload readers with too much information at once, we did not discuss the statistical power of each model for detecting main effects and interactions up front, but instead presented this information as each model became relevant to our research questions. Below, we present the smallest main effects and 2-way interaction effects that we had sufficient power to detect 80% of the time, given each study's design and sample size (which, again, are referenced in Table S1). Of note, these power estimates were conducted using Monte Carlo simulations (see text, for model specifications).

Table S1.

#### *List of All Rating Studies, Including Sample Sizes, Experimental Designs, and Type of Images Rated*

#	Composite Image Type	Image Source(s)	Rating Study Design
1	Group ( <i>N</i> = 105)	Both Samples	2 (target) × 2 (image source) within-person experiment.
2	Group ( <i>N</i> = 195)	MTurk Sample	2 (target) × 2 (explicit dehumanization level) × 2 (dehumanization measure) experiment with repeated measures on first two factors.
3	Group ( <i>N</i> = 200)	Panels Sample	2 (target) × 2 (explicit dehumanization level) × 2 (dehumanization measure) experiment with repeated measures on first two factors.
4	Group ( <i>N</i> = 295)	MTurk Sample	2 (target) × 2 (left vs. right wing) × 3 (ideology measure) experiment with repeated measures on first two factors.
5	Group ( <i>N</i> = 289)	Panels Sample	2 (target) × 2 (left vs. right wing) × 3 (ideology measure) experiment with repeated measures on first two factors.
6	Individual ( <i>N</i> = 397)	MTurk Sample	2 (target) within-person experiment (each rater saw 4 randomly selected 'Arab' and 4 randomly selected 'American' images)
7	Individual ( <i>N</i> = 820)	Panels Sample	2 (target) within-person experiment (each rater saw 6 randomly selected 'Arab' and 6 randomly selected 'American' images)

*Note.* (target) = whether images were of Arabs or of Americans; (image source) = whether images were based on generators in the MTurk or Qualtrics Panels samples; (dehumanization measure) = whether images were grouped by generators' Ascent- or trait-assessed dehumanization levels; (ideology measure) = whether images were aggregated by generators' ideological leanings on the one-item liberalism-conservatism measure, the SDO measure, or the RWA measure.

Table S2.

*List of All Rating Studies, Including Sample Sizes, Type of Images Rated, and Statistical Power*

#	Composite Image Type	Image Source(s)	Smallest Detectable Main Effect	Smallest Detectable Interaction
1	Group ( <i>N</i> = 105)	Both Samples	$\beta = 0.20$	$\beta = 0.26$
2	Group ( <i>N</i> = 195)	MTurk Sample	$\beta = 0.15$	$\beta = 0.20$
3	Group ( <i>N</i> = 200)	Panels Sample	$\beta = 0.15$	$\beta = 0.20$
4	Group ( <i>N</i> = 295)	MTurk Sample	$\beta = 0.12$	$\beta = 0.17$
5	Group ( <i>N</i> = 289)	Panels Sample	$\beta = 0.12$	$\beta = 0.17$
6	Individual ( <i>N</i> = 397)	MTurk Sample	$\beta = 0.09$	$\beta = 0.10$
7	Individual ( <i>N</i> = 820)	Panels Sample	$\beta = 0.038$	$\beta = 0.04$

*Note.* All power estimates were computed by z-standardizing outcomes, contrast coding categorical predictors (such that all contrasts have a range of 1, are centered around zero, and are orthogonal to each other), z-standardizing continuous predictors (including covariates), and then running Monte Carlo simulations on each specified main effect or interaction beta. Every effect listed here can be detected at least 80% of the time, given each study's sample size and modeling strategy.

### **Evidence that Raters' Anti-Arab Biases Influence Their Ratings**

Throughout our manuscript, we control for raters' own anti-Arab biases (that is, their dehumanization of and prejudice against "Arabs" relative to "Americans"), where applicable. The reason for doing this is to preclude the possibility that differences between Arab and American CIs would be attributable to the biases in the minds of *raters* rather than *generators*.

In general, there was clear evidence that these were reasonable controls to include in our models. Data from the rating study of overall composite images serves as a nice proof of concept for this idea. In that study (described in Row 1 of Table S1), all raters indicated how dehumanizing they perceived faces to be in a 2 (target: Arab, American)  $\times$  2 (image source: MTurk Sample, Panels Sample) within-person experiment. Running that analysis with raters' *prejudice* included in the model—both its main effect and its interaction with the other two factors—revealed that the main effect of target (whereby composite images of Arabs are rated as more dehumanizing than composite images of Americans:  $\beta = 1.24$ , 95% CI[1.14, 1.34],  $F(1, 313) = 622.50$ ,  $p < .001$ ) was indeed moderated by raters' own prejudices against Arabs relative to Americans: interaction

$\beta = 0.33$   $p < .001$ . The nature of this interaction is what one might expect. Raters' who were a standard deviation higher than the sample mean on prejudice against Arabs were indeed likely to rate Arab CIs as more dehumanized relative to American CIs [ $\beta = 1.57$ , 95% CI[1.43, 1.71],  $F(1, 313) = 492.40$ ,  $p < .001$ ] than were raters who were a standard deviation lower than the sample mean on prejudice against Arabs [ $\beta = 0.91$ , 95% CI[0.78, 1.05],  $F(1, 313) = 168.38$ ,  $p < .001$ ]. Thus, there was clear reason to believe that raters' prejudices against Arabs can and did influence the extent to which they rated Arab CIs as more dehumanizing than American CIs, thus bolstering our decision to include rater prejudice as a covariate in our models.

Running the analysis described above with raters' *dehumanization* included in the model—both its main effect and its interaction with the other two factors—revealed a similar pattern. That is, the main effect of target was indeed moderated by raters' own dehumanization of Arabs relative to Americans: interaction  $\beta = 0.33$   $p < .001$ . The nature of this interaction was that raters who were a standard deviation higher than the sample mean on Arab dehumanization were indeed likely to rate Arab CIs as more dehumanized relative to American CIs [ $\beta = 1.58$ , 95% CI[1.44, 1.72],  $F(1, 314) = 482.11$ ,  $p < .001$ ] than were raters who were a standard deviation lower than the sample mean on Arab dehumanization [ $\beta = 0.92$ , 95% CI[0.78, 1.05],  $F(1, 313) = 166.77$ ,  $p < .001$ ]. Thus, there was clear reason to believe that raters' dehumanization of Arabs—in addition to their prejudices against Arabs—can and did influence the extent to which they rate Arab CIs as more dehumanizing than American CIs.

### **Supplement to Link Between Self-Reported and Visual Dehumanization**

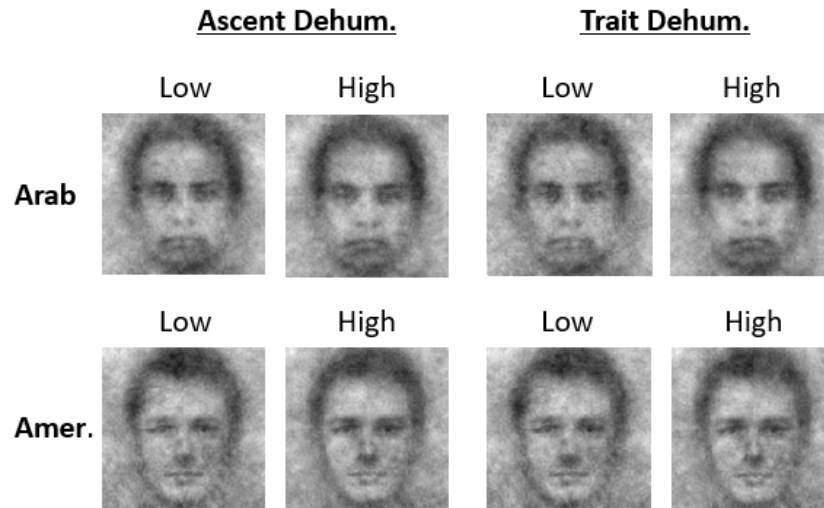
Our manuscript describes two group-image rating studies (rows 2 and 3 of Table S1) that focus on the question of whether generators who are higher in explicit dehumanization of Arabs (vs. Americans) likewise are higher in visual dehumanization of Arabs (vs. Americans). Below, we show what the group-aggregated images from these rating studies look like side by side. In addition, we present the full model that we ran on the data from these two rating studies (as opposed to collapsing across one of the factors, as we did for simplicity in the paper).

#### **Group Images by Generators' Reported Levels of Dehumanization**

As noted in footnote 11, the group images from the MTurk sample contained the data of approximately  $n = 25$  generators per image, whereas the group images from the Qualtrics Panels sample contained the data of approximately  $n = 42$  generators per image. Below, we show what group-aggregated images from these studies look like among generators who were below vs. above the median on each of our two measures of explicit Arab dehumanization (in the case of the MTurk sample) or among generators who were in the bottom vs. top quartile on each of our two explicit measures of Arab dehumanization (in the case of the Qualtrics Panels sample).

#### **From MTurk sample**

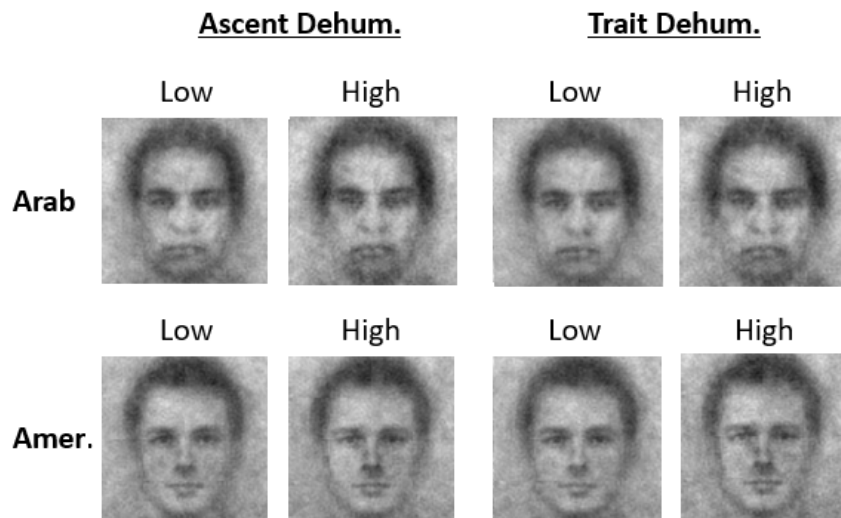
Below are images that were generated from the MTurk sample of generators. These photos were rated in the rating study described in row 2 of Table S1.



*Figure S1.* Group composite images of Arabs (top row) and Americans (bottom row) as a function of whether generators were low vs. high on each of two explicit Arab dehumanization measures. Generators were considered “low” if they were below the median; “high” if they were above.

#### **From Qualtrics Panels sample**

Below are images that were generated from the Qualtrics Panels sample of generators. These photos were rated in the rating study described in row 3 of Table 1. Note that these images are also on display in Figure 4.



*Figure S2.* Group composite images of Arabs (top row) and Americans (bottom row) as a function of whether generators were high vs. low on each of two explicit Arab dehumanization measures. Generators were considered “low” if they were in the bottom quartile; “high” if they were in the top quartile.

#### **A Full Analysis from Group Image Rating Studies**

The images depicted in Figures S1 and S2, respectively, were rated in separate rating studies (described in rows 2 and 3 of Table 1, respectively) that had a 2 (target: Arab, American)  $\times$  2 (generator explicit dehumanization: low, high)  $\times$  2 (measure type: trait, ascent) design with repeated measures on the first two factors. The third factor, measure type, was a between-person factor that described whether raters saw images broken down by generators' explicit dehumanization levels on the trait measure of dehumanization, or on the ascent measure of dehumanization. Because this third factor was not theoretically interesting, we did not discuss its main effects or interactions with the other factors from our full 2  $\times$  2  $\times$  2 model in our paper. Here, we report these main effects and interactions. As in our paper, we conducted these analyses by constructing two multilevel models, one for each rating study. In these models we regressed ratings of how dehumanizing representations appear onto contrast codes representing the full 2  $\times$  2 mixed factorial design, as well as onto the mean-centered covariates referenced in the paper (i.e., ratings of how favorable the mental representations appear; raters' anti-Arab prejudice and dehumanization). These models also included a random effect of rater intercept.

### **From MTurk sample**

In our paper, we reported from this analysis that there was a main effect of target (Arab vs. American) such that Arab images were rated as more visually dehumanized than American images ( $p < .001$ ). In addition, we reported in our paper that the magnitude of this effect (in the MTurk sample) was not moderated by whether the people who generated the images were high vs. low in explicit dehumanization of Arabs ( $p = .92$ ). This two-way interaction was not contingent on whether raters had been assigned images that were broken down by the ascent measure, or whether they had instead been assigned images that were broken down by the trait measure (three-way interaction  $\beta = -0.01$ ,  $p = .95$ ). In addition, which set of images raters had been assigned did not have a main effect on dehumanization ratings ( $\beta = -0.10$ ,  $p = .12$ ) and it also did not interact with either target (Arab vs. American: interaction  $\beta = -0.01$ ,  $p = .92$ ) or generator dehumanization level (low vs. high: interaction  $\beta = 0.04$ ,  $p = .55$ ). Thus, there was no evidence that any of our reported effects depended on whether explicit dehumanization was operationalized as ascent dehumanization, or instead as trait dehumanization.

### **From Qualtrics Panels sample**

Analyses from the Qualtrics Panel sample also suggest that which image set raters had been assigned had no impact on how visually dehumanized they perceived images to be. In the Qualtrics Panels sample, we did find that the main effect of target (Arab vs. American) on visual dehumanization was moderated by whether generators were high vs. low on explicit dehumanization of Arabs (two-way interaction:  $p = .046$ ). This two-way interaction was not dependent on whether raters saw images that were broken down by generators' dehumanization levels on the ascent measure, or instead by generators' dehumanization levels on the trait measure of explicit dehumanization (three-way interaction:  $\beta = 0.04$ ,  $p = .76$ ). In addition, measure type (i.e., ascent vs. trait) did not exert a main effect influence on dehumanization ratings ( $\beta = -0.04$ ,  $p = .58$ ) and it likewise did not interact with target (Arab vs. American;  $\beta = -0.07$ ,  $p = .31$ ) or with generators' dehumanization levels (low vs. high;  $\beta = 0.08$ ,  $p = .21$ ). Thus, we again found that our conclusions do not depend on how generators' explicit dehumanization of Arabs (vs. Americans) was operationalized.

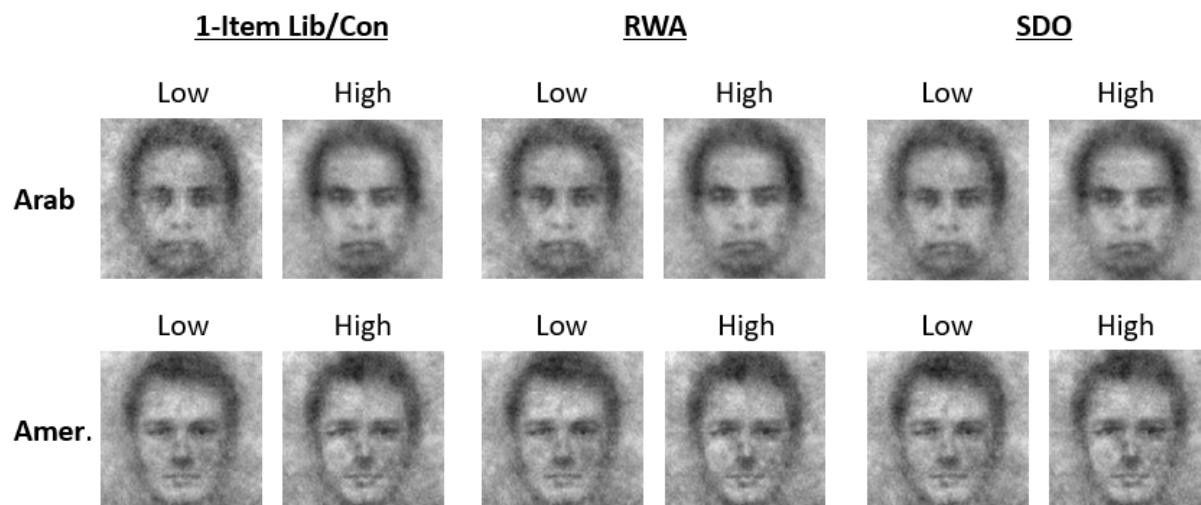
## Supplement to Link Between Political Leanings and Visual Dehumanization

Our paper reports two rating studies (described in rows 4 and 5 of Table S1, respectively) in which participants rated four images in a 2 (target)  $\times$  2 (left vs. right wing)  $\times$  3 (ideology measure) design with repeated measures on first two factors. The third factor was a between-subjects factor that describes whether images were broken down by generators' political ideology levels on either the 1-item liberalism-conservatism measure of political ideology, the right-wing authoritarianism measure of political ideology, or the social dominance orientation measure of political ideology. Group images that were considered "left wing" were those that contained the data of generators who were below the median on each of these measures (in the case of the MTurk sample) or who were in the bottom quartile on each of these measures (in the case of the Qualtrics Panels sample). Group images that were considered "right wing" were those that contained data of generators who were above the median or in the top quartile, respectively. Below, we show what these group-aggregated images look like, and we discuss the average levels of explicit Arab dehumanization among those grouped as "left wing" vs. "right wing." Finally, in our paper we reported a marginal three-way interaction that suggested, perhaps, that the tendency for generators' political ideology levels to influence their visual representations of Arabs might depend on which measure of political ideology we used to create the group-composite images. Here, we decompose that marginal three-way interaction.

### Group Images by Generators' Political Ideology Levels

#### From MTurk sample

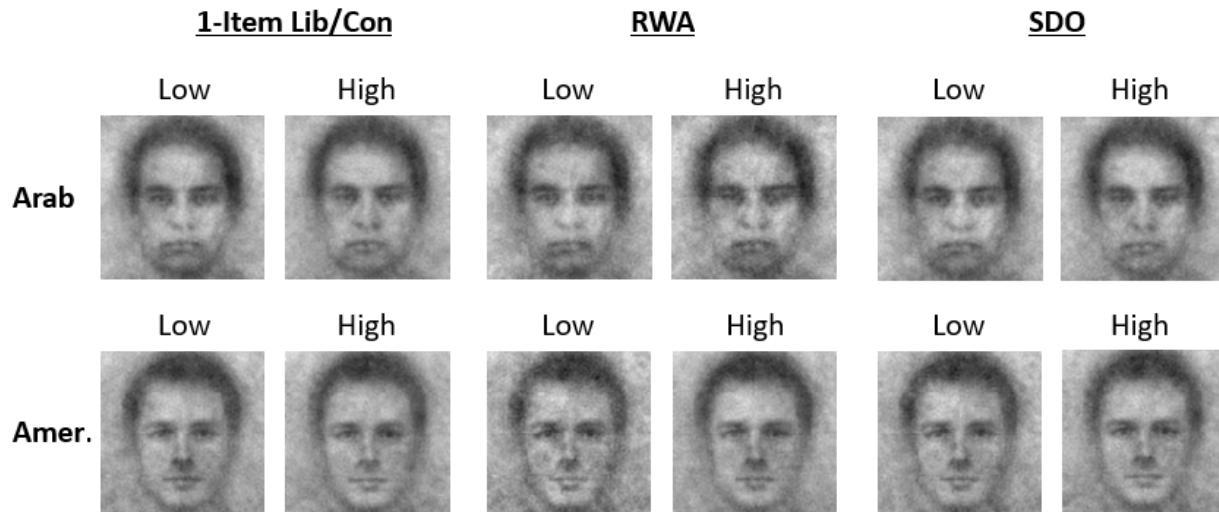
Below are images that were generated from the MTurk sample of generators. These photos were rated in the rating study described in row 4 of Table S1.



*Figure S3.* Group composite images of Arabs (top row) and Americans (bottom row) as a function of whether generators were high vs. low on each of three right-wing ideology measures. Generators were considered "low" if they were below the median; "high" if they were above.

### From Qualtrics Panels sample

Below are images that were generated from the Qualtrics Panels sample of generators. These photos were rated in the rating study described in row 5 of Table 1. Note that these images are exactly the same as those we depict in Figure 5 of the manuscript.



*Figure S4.* Group composite images of Arabs (top row) and Americans (bottom row) as a function of whether generators were high vs. low on each of three right-wing ideology measures. Generators were considered “low” if they were in the bottom quartile; “high” if they in the top quartile.

### Self-Reported Dehumanization among Liberals and Conservatives

As noted in our manuscript, those who were grouped as “left wing” dehumanized Arabs only to a very small degree (if at all). Below we report the descriptive statistics on explicit dehumanization among those groups as “left” vs. “right” wing in both the MTurk and Qualtrics Panels samples of generators.

### From MTurk sample

In the MTurk sample of generators, those who were grouped as left wing reported substantially lower levels of blatant dehumanization of Arabs (vs. Americans) than those who were grouped as right-wing (all  $ps < .001$ ). Here, we report the descriptive amounts by which generators in both groups (that is, those who were below vs. above the median on each of our three political ideology measures) dehumanized Arabs more than Americans on both of our explicit measures of dehumanization. An  $M_{diff}$  of 0 would indicate that the group did not dehumanize Arabs any more or less than Americans; scores greater than 0 indicate that Arabs were explicitly dehumanized more than Americans on the measure.



Table S3.

*Blatant Dehumanization Among MTurk Generators Grouped as Left vs. Right Wing*

	Left-Wing Generators	Right-Wing Generators
<b>RWA</b>	Trait: $M_{\text{diff}} = 0.42, p = .13$ Ascent: $M_{\text{diff}} = 9.12, p = .02$	Trait: $M_{\text{diff}} = 2.00, p < .001$ Ascent: $M_{\text{diff}} = 32.54, p < .001$
<b>SDO</b>	Trait: $M_{\text{diff}} = 0.63, p = .03$ Ascent: $M_{\text{diff}} = 12.41, p = .002$	Trait: $M_{\text{diff}} = 1.72, p < .001$ Ascent: $M_{\text{diff}} = 28.00, p < .001$
<b>1-item Lib/Con</b>	Trait: $M_{\text{diff}} = 0.39, p = .15$ Ascent: $M_{\text{diff}} = 12.65, p = .002$	Trait: $M_{\text{diff}} = 2.09, p < .001$ Ascent: $M_{\text{diff}} = 17.84, p < .001$

*Note.* This table describes the amount by which left-wing and right-wing generators dehumanize Arabs relative to Americans on the trait and ascent measures of explicit dehumanization.

**From Qualtrics Panels sample**

In the Qualtrics Panels sample of generators, those who were grouped as left wing reported substantially lower levels of blatant dehumanization of Arabs (vs. Americans) than those who were grouped as right-wing (all  $ps < .001$ ). Here, we report the descriptive amounts by which generators in both groups (that is, those who were in the bottom vs. top quartiles on each of our three political ideology measures) dehumanized Arabs more than Americans on both of our explicit measures of dehumanization. An  $M_{\text{diff}}$  of 0 would indicate that the group did not dehumanize Arabs any more or less than Americans; scores greater than 0 indicate that Arabs were explicitly dehumanized more than Americans on the measure.

Table S4.

*Blatant Dehumanization Among Qualtrics Panels Generators Grouped as Left vs. Right Wing*

	Left-Wing Generators	Right-Wing Generators
<b>RWA</b>	Trait: $M_{\text{diff}} = -0.08, p = .62$ Ascent: $M_{\text{diff}} = 7.44, p = .013$	Trait: $M_{\text{diff}} = 1.02, p < .001$ Ascent: $M_{\text{diff}} = 20.33, p < .001$
<b>SDO</b>	Trait: $M_{\text{diff}} = 0.01, p = .93$ Ascent: $M_{\text{diff}} = 4.57, p = .12$	Trait: $M_{\text{diff}} = 0.80, p < .001$ Ascent: $M_{\text{diff}} = 20.53, p < .001$
<b>1-item Lib/Con</b>	Trait: $M_{\text{diff}} = 0.15, p = .32$ Ascent: $M_{\text{diff}} = 7.74, p = .006$	Trait: $M_{\text{diff}} = 0.96, p < .001$ Ascent: $M_{\text{diff}} = 19.11, p < .001$

*Note.* This table describes the amount by which left-wing and right-wing generators dehumanize Arabs relative to Americans on the trait and ascent measures of explicit dehumanization.

**Decomposition of Marginal Three-Way Interaction**

In the rating study described in row 4, there was no three-way interaction between which measure of political ideology generators' images were grouped by (SDO, RWA, or 1-item liberalism/conservatism), who was targeted in the image (i.e., Arabs vs. Americans), and whether the images were generated by people who were left vs. right wing (three-way interaction  $ps > .17$ ). However, in the rating study described in row 5—that is, in the rating study of images generated from the Qualtrics Panels sample—there *was* a marginal three-way interaction between these three variables:  $\beta = 0.21$ ,  $F(1, 850) = 2.87$ ,  $p = .09$ . The nature of this three-way interaction was that the target-by-ideology interaction was marginally stronger when the photos were broken down by generator RWA levels (two-way interaction  $\beta = 0.23$ ,  $F(1, 850) = 5.29$ ,  $p = .022$ ,  $R^2 < .01$ ) than when photos were broken down by generator SDO or Liberalism-Conservatism levels (two-way interaction  $\beta = 0.02$ ,  $F(1, 850) = 0.10$ ,  $p = .75$ ,  $R^2 < .01$ ). When the images were broken down by whether generators were left vs. right wing on the RWA dimension specifically, those who were right-wing visually dehumanized Arabs slightly more ( $\beta = 0.63$ , 95% CI[0.48, 0.79],  $F(1, 889) = 66.57$ ,  $p < .001$ ) than those who were left-wing ( $\beta = 0.40$ , 95% CI[0.25, 0.54],  $F(1, 876) = 27.63$ ,  $p < .001$ ). However, when the images were broken down by the other two dimensions (generators' placement on the 1-item liberalism/conservatism measure or the SDO measure), there was no such tendency for visual dehumanization to be greater among right-wing generators ( $\beta = 0.34$ , 95% CI[0.22, 0.45],  $F(1, 913) = 33.78$ ,  $p < .001$ ) than among left-wing generators ( $\beta = 0.31$ , 95% CI[0.19, 0.43],  $F(1, 924) = 27.44$ ,  $p < .001$ ). Thus, although we found some evidence in the Qualtrics Panels study that right-wing generators on RWA specifically did indeed visually dehumanize Arabs to a greater level than left-wing generators, this effect was not reliably different from the null effects we found when the photos were broken down either by generators' SDO levels, or generators levels of liberalism vs. conservatism on the 1-item measure of political ideology.

### Exploratory Analyses from the Qualtrics Panels Sample

As we discuss in our paper, the Qualtrics Panels sample of generators completed exploratory measures that were meant help shed light on the question of how accurate they are at anticipating how their own mental representations appear. Here, we present the results from additional exploratory analyses that we did that did not make it into our paper, but that are nevertheless helpful in beginning to address this research question.

#### Exploratory Analyses on Moderators of Visual Dehumanization

In general, we found evidence that positive intergroup contact was associated with harboring less dehumanized visual representations of Arabs. However, we found that positive media exposure to Arabs was unrelated to visual dehumanization of Arabs. In addition, we found that those who had more exposure to terroristic depictions of the Arabs in the media harbored visual representations of Arabs that were *less* dehumanized.

Positive intergroup contact. Generators in the Panels sample answered the question, “How often do you interact with Arab people?” on a scale from 0 = *never* to 6 = *daily*. In addition, we asked generators three questions about positive their intergroup encounters with Arabs were (e.g., “When you meet Arab people, do you find the contact pleasant? Anchored at 0 = *not at all* and 6 = *very much*), which were aggregated into a positivity index. Finally, we multiplied generators'

interaction frequency by the positivity of their intergroup encounters to obtain a continuous index ranging from  $0 = \text{no positive interaction with Arabs}$  to  $36 = \text{daily positive interaction with Arabs}$ . To investigate whether positive intergroup contact with Arabs influenced the extent to which generators visually dehumanized Arabs (vs. Americans), we constructed a multilevel model (using data from the study described in Row 7 of Table 1). In this model, we regressed visual dehumanization ratings onto a contrast representing whether the image was of Arabs vs. Americans; onto generators' levels of positive intergroup contact with Arabs; as well as onto a contrast representing the cross-level interaction between these two variables. As in our manuscript, we included mean-centered covariates in this model (image positivity, raters' prejudice and dehumanization levels), as well as random effects of raters' intercepts as well as stimulus intercepts. This analysis revealed an interaction between generators' levels of positive contact with Arabs and how dehumanized their visual representations of Arabs (relative to Americans) were (interaction  $\beta = -0.03, p = .031$ ). The nature of this interaction was that those who reported high levels of positive intergroup contact (that is, those who were a standard deviation above the mean) indeed exhibited lower levels of visual dehumanization of Arabs ( $\beta = 0.08, p < .001$ ) than those who reported low levels of positive intergroup contact (that is, those who were a standard deviation below the mean:  $\beta = 0.15, p < .001$ ).

Positive media exposure. Generators in the Panels sample also responded to the question, "How often are Arabs depicted, discussed, or talked about in any of the media outlets you consume?" on a scale from  $0 = \text{never}$  to  $6 = \text{daily}$ . In addition, generators answered the question, "When you see Arabs depicted, discussed, or talked about in any of these outlets, how negative vs. positive does it seem?" on a scale from  $-3 = \text{very negative}$  to  $3 = \text{very positive}$ . As before, frequency of (vicarious) contact was multiplied by the valence of that contact into a composite index ranging from  $-18 = \text{daily negative contact}$  to  $18 = \text{daily positive contact}$ . To investigate whether positive media exposure to Arabs influenced the extent to which generators visually dehumanized Arabs (vs. Americans), we constructed a multilevel model (see study in Row 7 of Table 1). In this model, we regressed visual dehumanization ratings onto a contrast representing whether the image was of Arabs vs. Americans; onto generators' levels of positive media exposure to Arabs; as well as onto a contrast representing the cross-level interaction between these two variables. As above, we included mean-centered covariates in this model (image positivity, raters' prejudice and dehumanization levels), as well as random effects of raters' intercepts as well as image intercepts. This analysis revealed no interaction between generators' levels of positive media exposure to Arabs and how dehumanized their visual representations of Arabs (vs. Americans) were rated (interaction  $\beta = -0.02, p = .10$ ). It is worth noting, though, that this non-significant interaction was in the direction of more positive media exposure leading to less dehumanized mental representations of Arabs.

Exposure to terroristic depictions of Arabs. Finally, participants completed a 4-item measure of how often they reported seeing terroristic depictions of Arabs in the media they consumed (measure adapted from Saleem, Wojcieszak, Hawkins, Li, & Ramasubramanian, 2019: e.g., "How often have you seen news stories or media reports about terrorism perpetrated by Arabs?" from  $0 = \text{never}$  to  $6 = \text{daily}$ ). This measure was then subjected to the same analysis as described for the other positive media exposure measures. Results from this analysis revealed that greater exposure to terroristic depictions of Arabs in the media was actually associated with having *less* dehumanized representations of Arabs (vs. Americans: interaction  $\beta = -0.02, p = .044$ ). That is,

visual dehumanization of Arabs was lower among generators high (+1 SD) on exposure to terroristic depictions ( $\beta = 0.08, p < .001$ ) than it was among generators low (−1 SD) on exposure to terroristic depictions ( $\beta = 0.14, p < .001$ ). This finding was unexpected and broadly inconsistent with the two findings noted above. Although undoubtedly speculative, this may have to do with differences in the threshold individuals use to define terroristic depictions. For example, individuals with more positive mental representations of Arabs might be faster to judge a media depiction as rising to the level of ‘terroristic’.

### **Exploratory Analyses on Pixel Correlations**

Although it is not noted in our manuscript, we were also able to assess generators’ objective accuracy at inferring what their own mental representation would look like by conducting luminance pixel correlations between a) pixels in the image they chose on the image-selection measure, and b) pixels in the images that they generated during the reverse-correlation procedure. To conduct these analyses, we used the “jpeg” package in R to save generators’ selected images and actual images (that is, those they produced during the reverse-correlation procedure) as vector strings that contain the luminance of each pixel. After applying a mask to generators’ chosen images and mental representations (mask downloaded from OSF page associated with Brown-Iannuzzi, Dotsch, Cooley, and Payne, 2017), we then conducted correlations between these pixel vectors for each generator.

In a set of exploratory analyses, we examined whether there were differences between left-wing and right-wing generators in terms of how objectively accurate they were at inferring how their own mental representations appeared (that is, whether the strength of generators’ pixel correlations depended on their own political leanings). To investigate these questions, we conducted three separate linear regressions in which we predicted z-standardized pixel correlation strength from generators’ z-standardized scores on SDO, RWA, and the 1-item measure of liberalism-conservatism, respectively. These analyses revealed that objective accuracy (indexed here by pixel correlation strength) was not correlated with how right-leaning generators were on SDO ( $\beta = -0.01, 95\% \text{ CI}[-0.12, 0.09], t(334) = -0.24, p = .81$ ), on RWA ( $\beta = 0.01, 95\% \text{ CI}[-0.09, 0.12], t(334) = 0.20, p = .84$ ), or on the 1-item measure of political ideology ( $\beta = -0.01, 95\% \text{ CI}[-0.12, 0.10], t(334) = -0.16, p = .88$ ). Thus, according to these exploratory analyses, generators on the left are no more or less accurate than those on the political right when it comes to projecting how their mental representations appear. Of note, these analyses accord with what we wrote in our manuscript’s exploratory analysis section.

Finally, we were able to use pixel correlations to assess the extent to which mental representations of Arabs were objectively similar to each other in each of our samples: the MTurk sample, and the Panels sample. This information is reported in Tables S5 and S6, which reveal high objective overlap in representations of Arabs regardless of one’s political leanings or levels of explicit dehumanization.

Table S5

*Pixel Correlations Between All Group-Level Arab Composite Images Generated by MTurk Sample*

	<b>Overall</b>	<b>Lib.</b>	<b>Con.</b>	<b>Low SDO</b>	<b>High SDO</b>	<b>Low RWA</b>	<b>High RWA</b>	<b>Low Ascent</b>	<b>High Ascent</b>	<b>Low Trait</b>	<b>High Trait</b>
<b>Overall</b>	---	.96	.99	.98	.99	.98	.99	.98	.98	.97	.99
<b>Lib.</b>		---	.93	.98	.94	.98	.94	.97	.94	.97	.94
<b>Con.</b>			---	.96	.99	.96	.99	.97	.98	.95	.99
<b>Low SDO</b>				---	.95	.99	.96	.98	.96	.97	.96
<b>High SDO</b>					---	.96	.99	.97	.98	.95	.98
<b>Low RWA</b>						---	.95	.98	.96	.97	.96
<b>High RWA</b>							---	.97	.98	.95	.99
<b>Low Ascent</b>								---	.94	.98	.96
<b>High Ascent</b>									---	.94	.99
<b>Low Trait</b>										---	.93

*Note.* Overall composite image (column 1) of Arabs drawn from Figure 2 of the manuscript; next six columns describe the composite images of Arabs drawn from Figure S3; final four columns describe the composite images of Arabs drawn from Figure S1. All  $ps < .0001$ .

Table S6.

*Pixel Correlations Between All Group-Level Arab Composite Images Generated by Panels Sample*

	<b>Overall</b>	<b>Lib.</b>	<b>Con.</b>	<b>Low SDO</b>	<b>High SDO</b>	<b>Low RWA</b>	<b>High RWA</b>	<b>Low Ascent</b>	<b>High Ascent</b>	<b>Low Trait</b>	<b>High Trait</b>
<b>Overall</b>	---	.99	.98	.99	.98	.98	.98	.99	.99	.98	.99
<b>Lib.</b>		---	.97	.98	.97	.98	.97	.98	.98	.98	.98
<b>Con.</b>			---	.97	.99	.95	.97	.97	.97	.97	.98
<b>Low SDO</b>				---	.97	.97	.97	.99	.98	.98	.97
<b>High SDO</b>					---	.96	.97	.97	.98	.97	.97
<b>Low RWA</b>						---	.95	.97	.97	.97	.96
<b>High RWA</b>							---	.96	.98	.97	.98
<b>Low Ascent</b>								---	.97	.98	.97
<b>High Ascent</b>									---	.97	.99
<b>Low Trait</b>										---	.97

*Note.* Overall composite image (column 1) of Arabs drawn from Figure 2 of the manuscript; next six columns describe the composite images of Arabs drawn from Figure S4; final four columns describe the composite images of Arabs drawn from Figure S2. All  $ps < .0001$ .