Social Cognition, Vol. 44, No. 2, 2026, pp. 150–169
© 2026 The Guilford Press. https://doi.org/10.1521/soco.2026.44.2.150

Mental Representations of Middle Eastern Versus Ethnically Unspecified Women and Men Reveal How Ethnicity and Gender Constrain Each Other

Matthew Elderkin
University of California, Davis

Samuel A.W. Klein
Columbia University

Stephanie Oliinyk, Emily Gelerian, and Andrew R. Todd
University of California, Davis

Mental representations of other people are informed by their intersecting social identities. In two experiments (N_{total} = 1,592 participants), we tested whether ethnicity and binarized gender interact to shape how Middle Eastern (vs. ethnically unspecified) women and men are mentally represented. Results revealed that spontaneously generated character and facial appearance representations of Middle Eastern women differed from those of Middle Eastern men (i.e., more trustworthy and less dominant in character and facial appearance) and ethnically unspecified women (i.e., less trustworthy in character and facial appearance, less dominant in character but more dominant in facial appearance). Notably, these results differed somewhat from self-reported impressions, suggesting some, but not complete, correspondence between spontaneous and deliberate reactions to these groups. This work complements growing evidence that mental representations of ethnicity and gender constrain each other.

Keywords: ethnicity, gender, person perception, reverse correlation, social categorization

Social categorization informs internal cognitive reflections, or mental representations, of oneself and others. Merely classifying someone as belonging to one or another, even arbitrary, social category affects how (un)favorably they are mentally represented (Dunham, 2018; Tajfel et al., 1971). Meaningful social identities

Address correspondence to Matthew Elderkin or Andrew R. Todd, Department of Psychology, University of California, Davis, One Shield Ave., Davis, CA 95816. E-mail: melderkin@ucdavis.edu or atodd@ucdavis.edu

carry pertinent information that further shapes such mental representations (Macrae & Bodenhausen, 2000). Women, for example, are often viewed as more communal but less agentic than men (Martin & Slepian, 2021). Moving beyond single social categories (e.g., binarized gender), accumulating research has explored how people with different identity combinations—based on age, gender, race/ethnicity, sexual orientation, etc.—are mentally represented (Hall et al., 2019; Kang & Bodenhausen, 2015; Lei et al., 2023).

Gendered race theorizing posits that gender and race/ethnicity are intimately linked, with mounting evidence indicating that Black women, for example, are construed as less prototypical of their constituent racial and gender categories than Black men and White women, respectively (Galinsky et al., 2013; Johnson et al., 2012; Schug et al., 2015).[1] In one study, "dangerous" and "criminal" were among the top traits describing Black people and Black men, but not Black women (Ghavami & Peplau, 2013). In another study, Black women and girls were more weakly associated with guns than Black men and boys but more strongly associated with guns than White women and girls (Thiem et al., 2019). Might Middle Eastern women also be construed differently than the prototypes of their constituent ethnic and gender categories?

Middle Eastern people are doubly burdened by the biases they endure and by their underrepresentation as targets of bias in research (Samari, 2023). Such invisibility may stem, in part, from the fact that Middle Eastern people are classified as White in the United States (U.S.) census, despite their not identifying as White nor being encoded as White by others (Chaney et al., 2021; Maghbouleh et al., 2022). As non-prototypical members of their constituent ethnic and gender categories (Ghavami & Peplau, 2013), Middle Eastern women may be particularly underrepresented in research and the public eye. Indeed, Middle Eastern women experience greater invisibility than Middle Eastern men and Middle Eastern people broadly (Samari et al., 2018). The current research examined whether and how mental representations of Middle Eastern women diverge from those of Middle Eastern men and ethnically unspecified women.

Particularly within the U.S., mental representations of Middle Eastern people are often negative, even dehumanizing (Kteily et al., 2015), in part because of alleged ties to terrorism (Saleem & Anderson, 2013). Although violent terrorist stereotypes derive primarily from Middle Eastern men, they typify Middle Eastern people broadly. Indeed, "terrorist" and "dangerous" are among the top traits describing Middle Eastern men and Middle Eastern people, but not Middle Eastern women (Ghavami & Peplau, 2013). The top traits describing Middle Eastern women include "oppressed" and "submissive," both of which are absent for Middle Eastern men and Middle Eastern people (Ghavami & Peplau, 2013). These same traits are also evident in media depictions of Middle Eastern women (Qutub, 2013).

Notably, most prior research on mental representations of Middle Eastern people has assessed self-reported impressions. In one exception, Petsko and colleagues (2021) used a reverse-correlation paradigm and found that Arab men's faces were

1. This evidence can also be accommodated by other theorizing about how intersectional social category prototypes shape person construal (see Lei et al., 2023, for a review).

mentally represented as more dehumanized than American men's faces, even among participants who disavowed such dehumanization in their self-reported impressions. This finding illustrates how indirect methods can capture aspects of mental representations that direct methods might miss, especially for stigmatized groups that evoke self-presentation concerns (Nosek, 2007). What remains untested is whether mental representations of Middle Eastern *women* differ from their constituent ethnic and gender categories when assessed *less directly*.

Unlike prior work that focused on specific groups in the Middle East and North Africa (MENA) geographical region (e.g., Arabs: Hong et al., 2024; Petsko et al., 2021; Moroccans: Dotsch et al., 2008, 2011), we focused on the broader category Middle Eastern. Although our focal interest was in Middle Eastern women versus Middle Eastern men and ethnically unspecified women, we used full factorial designs based on target ethnicity (Middle Eastern, unspecified) and binarized gender (women, men) in the two experiments reported here.

Experiment 1 used a two-phase approach to capture *character* representations (Gallagher & Bodenhausen, 2021; Petsko & Bodenhausen, 2019). In Phase 1, a first set of participants listed adjectives describing a person with a particular identity combination (e.g., Middle Eastern woman). These adjectives were treated as a profile reflecting each participant's mental representation of the target group's character (i.e., what group members *are* like). In Phase 2, a separate set of participants, with no knowledge about whom the profiles portrayed, reported their impressions of the profiled people. Experiment 2 used a different two-phase approach—a reverse-correlation paradigm—to capture *facial* representations (Brinkman et al., 2017). In Phase 1, a first set of participants repeatedly selected which of two noise-laden facial images looks more like a person with a particular identity combination (e.g., Middle Eastern woman). The selected images were combined with a base-face image to create a classification image reflecting each participant's mental representation of the target group's facial appearance (i.e., what group members *look* like). Phase 1 participants also directly reported their impressions of each target group. In Phase 2, an independent set of participants, with no knowledge about whom the images depicted, reported their impressions of the people in the images on the same traits, which afforded comparison with the direct reports from Phase 1.

Of primary interest was the trustworthiness ascribed to the character and facial representations. Related to communion, warmth, and "getting along", trustworthiness signals one's good or ill intentions and is part of the *horizontal* dimension of social evaluation (Abele et al., 2021; Koch et al., 2021). Given links between trustworthiness and valence in both character and facial representations (Tamir & Thornton, 2018; Todorov et al., 2008), assessing trustworthiness allows us to capture representation positivity. We also assessed the dominance ascribed to these character and facial representations. Related to agency, competence, and "getting ahead", dominance signals one's ability to enact one's intentions and is part of the *vertical* dimension of social evaluation (Abele et al., 2021; Koch et al., 2021). Crucially, because in neither experiment were Phase 1 participants prompted to consider their assigned group's trustworthiness or dominance, these two-phase

approaches afford indirect assessments of Phase 1 participants' character and facial representations along these dimensions via the trustworthiness and dominance ratings reported by Phase 2 participants.

Historically, women have been rated higher on horizontal traits and lower on vertical traits than men (Martin & Slepian, 2021; but see Eagly et al., 2020, for evidence that gender differences on some vertical traits have shifted). Thus, we expected that women would be viewed as more trustworthy but less dominant than men. We also expected that, given negative Middle Eastern sentiment (e.g., Kteily et al., 2015), Middle Eastern representations would be less trustworthy than ethnically unspecified representations. Insofar as Middle Eastern women are mentally represented differently than their constituent ethnic and gender categories, we reasoned that they would be viewed as more trustworthy than Middle Eastern men but less trustworthy than ethnically unspecified women. Finally, given stereotypes of Middle Eastern women's submissiveness (Ghavami & Peplau, 2013), we predicted that Middle Eastern women would be viewed as less dominant than Middle Eastern men and ethnically unspecified women.

Below, we describe our sample size rationale, and we report all manipulations, measures, and data exclusions.[2] These are our only experiments examining these research questions (i.e., there is no file drawer). Data, code, and materials are available at https://shorturl.at/FIcUn.

EXPERIMENT 1

Experiment 1 used an open-ended approach to assess character representations. Phase 1 participants (i.e., profile generators) listed adjectives describing a man or woman labeled Middle Eastern or whose ethnicity was unspecified. These adjectives were used to construct profiles. Phase 2 participants (i.e., profile raters), who had no knowledge about whom the profiles portrayed, rated a subset of the profiles from each condition on trustworthiness or dominance. Because Phase 1 participants were never prompted to consider their assigned group with respect to these traits, this procedure affords an indirect assessment of Phase 1 participants' character representations of the groups along these dimensions via Phase 2 participants' profile ratings. Accordingly, this approach may offer insight into spontaneously activated character stereotypes (Gallagher & Bodenhausen, 2021; Petsko & Bodenhausen, 2019).

METHOD

Phase 1: Profile Generation

Participants (Profile Generators). To our knowledge, there is no consensus about a priori power analysis calculations for Phase 1 in two-phase procedures like those used here (see also Hutchings et al., 2021, 2024). So, we set a target sample size using a heuristic of approximately 100 participants per condition, or approximately 400

2. All data exclusion decisions were made a priori before any data were analyzed.

TABLE 1. Participant Demographics (Experiments 1 and 2)

		Age		Gender (%)			Race/Ethnicity (%)						
Experiment	*N*	*M*	*SD*	Women	Men	Nb	A	B	L	Me	Mx	Na	W
1 (PG)	391	39.2	12.0	52.4	44.5	1.5	5.9	7.7	4.9	0.0	2.8	0.5	77.2
1 (PA)	535	39.9	12.2	57.9	40.2	1.2	4.8	10.8	5.8	0.0	1.7	0.6	75.1
2 (IG)	430	19.5	1.9	69.1	27.7	6.2	57.7	2.1	10.7	1.6	2.6	0.2	15.6
2 (IA)	265	39.8	12.5	55.5	43.0	1.1	6.0	12.8	4.9	0.0	3.7	1.5	69.4

Note. PG = profile-generation phase, PA = profile-assessment phase, IG = image-generation phase, IA = image-assessment phase, *N* = number of participants included in analyses, Nb = nonbinary, A = Asian, Asian American, or Pacific Islander, B = Black or African American, L = Latina/o/e/x or Hispanic, Me = Middle Eastern or North African, Mx = reported another or more than one race/ethnicity, Na = Native American, W = White or European American. Some participants did not report their gender or race/ethnicity.

participants in Phase 1's 2 × 2 between-subjects design. Anticipating exclusions, we recruited 500 participants. In total, 511 Amazon's Mechanical Turk (MTurk) workers[3] located in the U.S. participated for pay. We excluded data from participants who failed an attention check ($n = 63$). Two coders flagged profiles that failed to follow task instructions (e.g., listing numbers), which we also excluded ($n = 57$). The final sample comprised 391 participants. Table 1 displays participant demographics for all experiments.

Procedure. Participants were randomly assigned to one of the 2 (target gender: man, woman) × 2 (target ethnicity: unspecified, Middle Eastern) between-subjects conditions. They completed a profile-generation task, listing the first five adjectives that came to mind when thinking of a man, a woman, a Middle Eastern man, or a Middle Eastern woman (Gallagher & Bodenhausen, 2021; Petsko & Bodenhausen, 2019). For illustrative purposes, Table 2 displays the top five adjectives listed in each condition. In both phases here and in Experiment 2, participants also completed an attention check (i.e., selecting "strongly agree" from several options) and reported demographic information.

Phase 2: Profile Assessment

Participants (Profile Raters). We set a target sample size of 250 participants each for the trustworthiness and dominance ratings. In total, 570 MTurk workers located in the U.S participated for pay. We excluded data from participants who failed an attention check ($n = 30$) or who rated >90% of the profiles identically ($n = 5$). The final sample comprised 535 participants, which afforded ≥80% power to detect main effects as small as $b = 0.23$ and interactions as small as $b = 0.44$.[4]

3. All MTurk workers were recruited from CloudResearch (Litman et al., 2017). To maximize data quality, we sampled CloudResearch-approved workers (Hauser et al., 2023) with approval ratings ≥95%.

4. For the sensitivity power analyses here and in Experiment 2, we used the *simr* package (Green & MacLeod, 2016).

TABLE 2. Most Frequently Listed Adjectives by Target Ethnicity and Gender (Experiment 1)

Target Ethnicity and Gender			
Unspecified		Middle Eastern	
Men	**Women**	**Men**	**Women**
Strong (58.9%)	Strong (37.6%)	Religious (19.8%)	Religious (23.6%)
Tall (20.0%)	Beautiful (26.7%)	Smart (12.3%)	Beautiful (19.8%)
Smart (12.6%)	Kind (26.6%)	Loud (9.8%)	Quiet (18.9%)
Aggressive (12.6%)	Caring (22.9%)	Strong (9.8%)	Oppressed (17.0%)
Masculine (11.6%)	Smart (19.3%)	Muslim (8.6%)	Smart (16.0%)

Note. Values in parentheses reflect the percentage of profile generators in that condition who listed the trait.

Procedure. Participants viewed multiple profiles and formed an impression of each person based only on the traits in their profile, which appeared in the exact order listed by profile generators. Of the 391 Phase 1 profiles, Phase 2 participants, who were blind to the group for which the profiles were generated, rated 39 or 40 randomly selected profiles (evenly distributed across conditions) on trustworthiness (n = 240 participants) or dominance (n = 266 participants; 1 = *extremely untrustworthy/submissive*, 7 = *extremely trustworthy/dominant*).

RESULTS

Analysis Strategy

Using the *lme4* (Bates et al., 2015) and *lmerTest* packages (Kuznetsova et al., 2017), we fit separate linear mixed-effects models (LMEMs) for profile raters' trustworthiness and dominance impressions. The LMEMs included fixed effects for target gender, target ethnicity, and their interaction.[5] We began with the maximal random-effects structure (Barr et al., 2013) and downsized to solve problems of nonconvergence and singularity. The final random-effects structure for both LMEMs included by-profile and by-rater random intercepts. For each analysis, we report both main effects and, when the interaction was significant, all underlying simple effects. Figure 1 displays the pattern of means for trustworthiness (left panel) and dominance (right panel). Descriptive statistics by condition for both experiments appear in Table 3.

Trustworthiness

Profiles of women were ascribed more trustworthiness than profiles of men, $b = 0.90$, $CI_{95\%}$ [0.69, 1.12], $t(350.98) = 8.32$, $p < .001$, and Middle Eastern profiles were

5. Preliminary exploratory analyses that included participant ethnicity as a moderator revealed negligible effects in both experiments (see the Supplemental Materials).

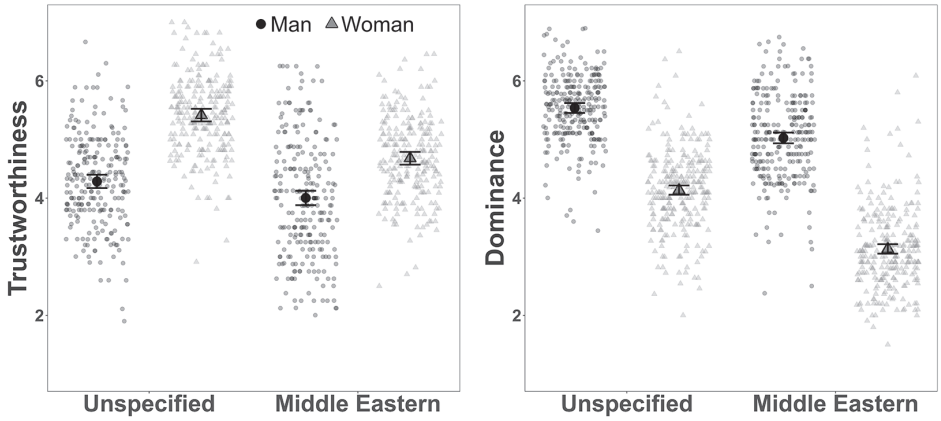

FIGURE 1. Profile raters' trustworthiness and dominance impressions of participant-level profiles (i.e., character representations) by target ethnicity and gender (Experiment 1). Markers reflect profile ratings on trustworthiness and dominance. Semitransparent markers reflect each participant's rating averaged across the profiles from that condition; filled shapes and their error bars reflect estimated marginal means and their 95% confidence intervals, respectively, from the linear mixed-effects models. The x-axis displays whether the profile portrayed an ethnically unspecified or Middle Eastern person. Shading and shape of markers reflect whether the profile portrayed a man or woman.

ascribed less trustworthiness than ethnically unspecified profiles, $b = –0.51$, $CI_{95\%}$ [–0.72, –0.29], $t(350.90) = –4.68$, $p < .001$. A significant target gender × target ethnicity interaction, $b = –0.45$, $CI_{95\%}$ [–0.88, –0.02], $t(365.82) = –2.08$, $p = .038$, further indicated that the gender difference (i.e., women more trustworthy than men) emerged for Middle Eastern profiles, $b = 0.68$, $CI_{95\%}$ [0.36, 0.99], $t(350.92) = 4.30$, $p < .001$, and ethnically unspecified profiles, $b = 1.13$, $CI_{95\%}$ [0.83, 1.43], $t(350.91) = 7.55$, $p < .001$,

TABLE 3. Mean Trustworthiness and Dominance Impressions by Target Ethnicity and Gender (Experiments 1 and 2)

	Target Ethnicity and Gender			
	Unspecified		Middle Eastern	
Trait	**Men**	**Women**	**Men**	**Women**
Experiment 1				
Trustworthiness	4.30 (0.77)	5.41 (0.66)	3.99 (1.02)	4.69 (0.69)
Dominance	5.53 (0.57)	4.14 (0.69)	5.03 (0.73)	3.13 (0.70)
Experiment 2 (Image-Generation Phase)				
Trustworthiness	3.67 (1.41)	5.17 (1.23)	4.05 (1.38)	5.08 (1.21)
Dominance	5.39 (0.97)	3.86 (1.08)	5.46 (1.11)	3.31 (1.30)
Experiment 2 (Image-Assessment Phase)				
Trustworthiness	4.10 (0.94)	4.59 (0.82)	4.12 (0.86)	4.43 (0.86)
Dominance	4.55 (0.72)	3.84 (0.91)	4.45 (0.68)	4.12 (0.78)

Note. Values in parentheses are standard deviations

but it was larger for the latter. Additionally, profiles of Middle Eastern women were ascribed less trustworthiness than profiles of ethnically unspecified women, b = –0.74, $CI_{95\%}$ [–1.03, –0.44], t(350.88) = -5.04, $p < .001$, whereas profiles of Middle Eastern men and ethnically unspecified men did not significantly differ, b = -0.28, $CI_{95\%}$ [-0.61, 0.04], t(350.87) = –1.75, $p = .080$.

Dominance

Profiles of women were ascribed less dominance than profiles of men, b = –1.65, $CI_{95\%}$ [–1.81, –1.49], t(389.41) = –20.04, $p < .001$, and Middle Eastern profiles were ascribed less dominance than ethnically unspecified profiles, b = –0.76, $CI_{95\%}$ [–0.92, –0.60], t(389.35) = -9.23, $p < .001$. A significant target gender × target ethnicity interaction, b = –0.49, $CI_{95\%}$ [–0.81, –0.17], t(389.30) = –2.99, $p = .002$, further indicated that the gender difference (i.e., women less dominant than men) emerged for Middle Eastern profiles, b = –1.89, $CI_{95\%}$ [–2.14, –1.64], t(389.34) = –15.90, $p < .001$, and ethnically unspecified profiles, b = –1.40, $CI_{95\%}$ [–1.64, –1.16], t(389.37) = –12.37, $p < .001$, but it was larger for the former. Also, consistent with stereotypes linking Middle Eastern women with submissiveness, profiles of Middle Eastern women were ascribed less dominance than profiles of ethnically unspecified women, b = –1.01, $CI_{95\%}$ [–1.23, –0.78], t(389.29) = –9.13, $p < .001$. A similar, albeit smaller ethnic difference emerged on the profiles of men, with profiles of Middle Eastern men being ascribed less dominance than those of ethnically unspecified men, b = –0.51, $CI_{95\%}$ [–0.75, –0.27], t(389.36) = –4.19, $p < .001$.

DISCUSSION

Character representations of women and ethnically unspecified people were more trustworthy than those of men and Middle Eastern people, respectively. These findings replicate established findings of positive sentiment toward women (Eagly & Mladinic, 1989) and negative sentiment toward Middle Eastern people (Kteily et al., 2015) on horizontal traits. Additionally, Middle Eastern women were viewed as more trustworthy than Middle Eastern men but less trustworthy than ethnically unspecified women. That women were viewed as less dominant than men aligns with historical gender differences (Martin & Slepian, 2021), suggesting that dominance may be a vertical trait where the gender gap remains (cf. Eagly et al., 2020). Furthermore, that character representations of Middle Eastern women were less dominant than those of Middle Eastern men and ethnically unspecified women aligns with stereotypes about Middle Eastern women's submissiveness (Ghavami & Peplau, 2013).

EXPERIMENT 2

People form impressions of others based solely on what they *look* like, with little to no knowledge about what they *are* like (Zebrowitz, 2017). Such facial representations often recapitulate character representations (e.g., Axt et al., 2023;

Brown-Iannuzzi et al., 2017; Dotsch et al., 2008, 2011; Klein et al., 2021), but not always (see, e.g., Hutchings et al., 2024; Petsko et al., 2021). To assess facial representations, we used a reverse-correlation paradigm. Phase 1 participants (i.e., image generators) constructed images reflecting their facial representation of a man or woman labeled Middle Eastern or whose ethnicity was unspecified. These participants also rated the four target groups on trustworthiness and dominance, which afforded a *direct* assessment of their character representations of these same groups. Phase 2 participants (i.e., image raters), who had no knowledge about whom the images depicted, rated a subset of the images from each condition on trustworthiness and dominance. Because Phase 1 participants were not prompted to consider their assigned group with respect to these traits until *after* generating the images, this procedure affords an indirect assessment of Phase 1 participants' facial representations of the groups along these dimensions via Phase 2 participants' image ratings. Accordingly, this approach may offer insight into spontaneously activated facial stereotypes (Brinkman et al., 2017).

METHOD

Phase 1: Image Generation

Participants (Image Generators). We again aimed for approximately 100 participants per condition, or approximately 400 participants in Phase 1's 2 × 2 between-subjects design. In total, 456 undergraduates from a university located in the Western U.S. participated for course credit. We excluded data from participants whose median response time on the image-classification task trials exceeded 4 s or whose response time on any trial exceeded 60 s (n = 7) or who responded identically on >90% of trials (n = 19). The final sample comprised 430 participants.[6]

Procedure. Participants were randomly assigned to one of the 2 (target gender: man, woman) × 2 (target ethnicity: unspecified, Middle Eastern) between-subjects conditions. They completed a two-image forced-choice image-classification task (Brinkman et al., 2017), generating a facial representation of a man, a woman, a Middle Eastern man, or a Middle Eastern woman. On each of 400 trials, they selected which of two adjacent noise-laden images looked more like someone from the target group. If participants did not respond within 4 s, a prompt ("Please answer the question") appeared. Blocks of 50 trials were separated by short breaks.

We created the stimuli using the *rcicr* package (Dotsch, 2014) by superimposing noise onto a base-face image. The base face (Gallagher & Bodenhausen, 2021) was the morphed average of 10 faces rated as the most gender-typical and race/ethnicity-typical faces of Asian, Black, Latina/o/e/x, and White women and men in the Chicago Face Database (Ma et al., 2015). A unique noise pattern[7] was

6. Degrees of freedom for some of the subsequent analyses differed from this total because of missing data.

7. The noise patterns were 4,092 superimposed truncated sinusoid patches in all possible combinations of two cycles in 6 orientations (0°, 30°, 60°, 90°, 120°, 150°) × 5 spatial frequencies (1, 2,

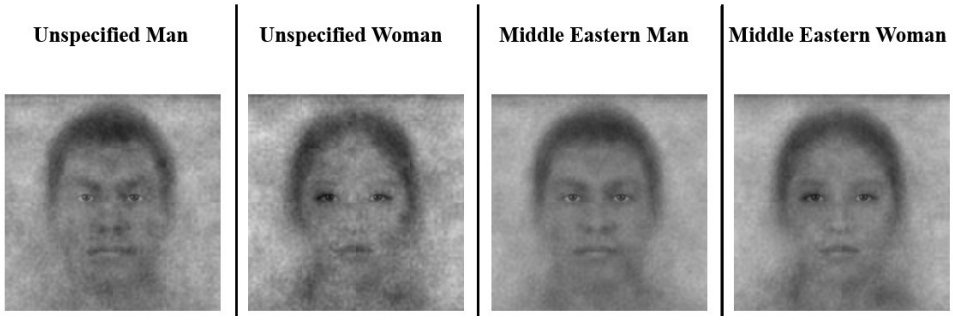

FIGURE 2. Condition-level classification images by target ethnicity and gender (Experiment 2). These condition-level classification images are displayed for illustrative purposes only. All analyses were conducted on the participant-level images.

generated for each trial and placed atop the base-face image to create one stimulus; the other stimulus was the inverse noise pattern atop the same base-face image, which maximizes between-face contrast.

Next, participants directly reported their impressions of all four groups on trustworthiness and dominance (1 = *extremely untrustworthy/submissive*, 7 = *extremely trustworthy/dominant*). The groups appeared in randomized order.

Image Processing. Again using the *rcicr* package, we created classification images (hereafter, 'images') by averaging each participant's selected noise patterns and superimposing it onto the base-face image. This procedure produced 430 participant-level images, one each from the 430 Phase 1 participants whose data were retained. Condition-level images, created by averaging the images generated by all participants in each condition, appear in Figure 2 (see the Supplemental Materials for image validation using cluster analysis); however, to minimize Type-I error inflation from analyzing condition-level images (Cone et al., 2021), we conducted all Phase 2 analyses on participant-level images.

Phase 2: Image Assessment

Participants (Image Raters). Using the same logic as before, we set a target sample size of 250 participants, who rated participant-level images on trustworthiness *and* dominance. In total, 293 MTurk workers located in the U.S. participated for pay. We excluded data from participants who failed the attention check (n = 16), who rated >90% of images identically (n = 5), or who requested that their data be excluded (n = 5). The final sample comprised 267 participants, which afforded ≥80% power to detect main effects of b = 0.10 and interactions of b = 0.21. The preregistration for Phase 2 can be found at https://aspredicted.org/vd9k-dgfh.pdf.

4, 8, 16 patches per image) × 2 phases (0, $\pi/2$) with random contrasts.

Procedure. The procedure followed Phase 2 in Experiment 1, except participants rated images instead of profiles. Of the 430 Phase 1 images, Phase 2 participants rated 43 randomly selected images (evenly distributed across conditions) on trustworthiness and dominance (1 = *extremely untrustworthy/submissive*, 7 = *extremely trustworthy/dominant*). Trait order was counterbalanced across participants.

RESULTS

Image-Generation Phase

*Analysis Strategy***.** We conducted separate 2 (target gender) × 2 (target ethnicity) repeated-measures ANOVAs for image generators' *direct* ratings of trustworthiness and dominance. For each analysis, we report both main effects and, when the interaction was significant, all underlying simple effects. Figure 3 displays the pattern of means for trustworthiness (left panel) and dominance (right panel).

Trustworthiness. Like Experiment 1's indirectly assessed character representations, women (M = 5.13, SD = 1.16) were ascribed more trustworthiness than men (M = 3.86, SD = 1.06), $F(1, 425) = 382.90$, $p < .001$, $\eta_p^2 = .47$, $CI_{90\%}$ [.42, .52]. Unlike Experiment 1's indirectly assessed character representations, however, Middle Eastern targets (M = 4.57, SD = 1.09) were ascribed *more* trustworthiness than ethnically unspecified targets (M = 4.42, SD = 0.95), $F(1, 425) = 8.92$, $p = .003$, $\eta_p^2 = .02$, $CI_{90\%}$ [.004, .05].

A significant target gender × target ethnicity interaction, $F(1, 425) = 27.24$, $p < .001$, $\eta_p^2 = .06$, $CI_{90\%}$ [.03, .10], further indicated that, like Experiment 1's indirectly assessed character representations, the gender difference (i.e., women more trustworthy than men) emerged for Middle Eastern targets, $t(425) = 15.05$, $p < .001$, d = 0.73, $CI_{95\%}$ [0.62, 0.84], and ethnically unspecified targets, $t(425) = 17.21$, $p < .001$, d = 0.83, $CI_{95\%}$ [0.72, 0.94], but it was larger for the latter. Also, unlike Experiment 1's indirectly assessed character representations, Middle Eastern women and ethnically unspecified women did not significantly differ, $t(425) = -1.47$, $p = .142$, d = -0.07, $CI_{95\%}$ [–0.17, 0.02], whereas Middle Eastern men were ascribed *more* trustworthiness than ethnically unspecified men, $t(425) = 5.10$, $p < .001$, d = 0.25, $CI_{95\%}$ [0.15, 0.34].

Dominance. Like Experiment 1's indirectly assessed character representations, women (M = 3.58, SD = 1.00) were ascribed less dominance than men (M = 5.42, SD = 0.84), $F(1, 427) = 679.12$, $p < .001$, $\eta_p^2 = .61$, $CI_{90\%}$ [.57, .65], and Middle Eastern targets (M = 4.39, SD = 0.76) were ascribed less dominance than ethnically unspecified targets (M = 4.63, SD = 0.67), $F(1, 427) = 31.87$, $p < .001$, $\eta_p^2 = .07$, $CI_{90\%}$ [.04, .11].

A significant target gender × target ethnicity interaction, $F(1, 427) = 48.65$, $p < .001$, $\eta_p^2 = .10$, $CI_{90\%}$ [.06, .14], further indicated that, like Experiment 1's indirectly assessed character representations, the gender difference (i.e., women less dominant than men) emerged for Middle Eastern targets, $t(427) = -23.67$, $p < .001$, d = -1.15, $CI_{95\%}$ [–1.27, –1.02], and ethnically unspecified targets, $t(427) = -20.50$,

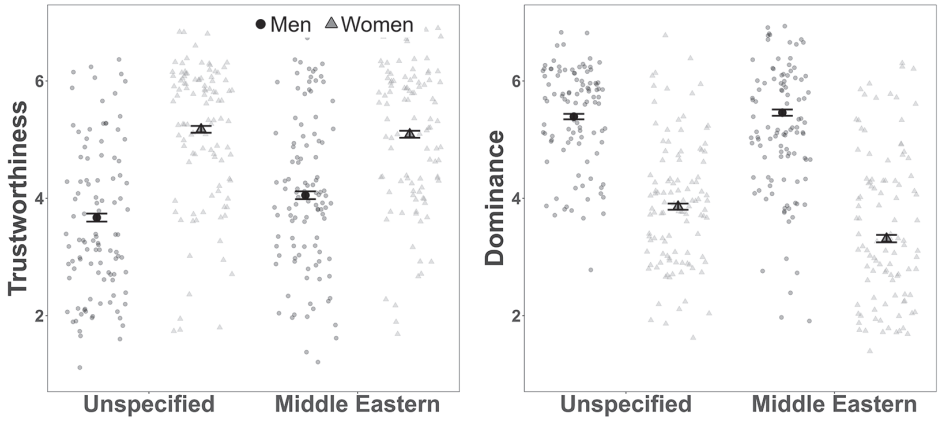

FIGURE 3. Image generators' trustworthiness and dominance impressions of groups (i.e., character representations) by target ethnicity and gender (Experiment 2). Markers reflect group ratings on trustworthiness and dominance. Semitransparent markers reflect each participant's rating for that condition; filled shapes and their error bars reflect estimated marginal means and their 95% confidence intervals, respectively, from the repeated-measures analyses of variance. The x-axis displays whether the group was ethnically unspecified or Middle Eastern. Shading and shape of markers reflect whether the group was men or women.

$p < .001$, $d = –0.99$, $CI_{95\%}$ [–1.11, –0.88], but it was larger for the former. Also, like Experiment 1's indirectly assessed character representations and consistent with stereotypes about Middle Eastern women's submissiveness, Middle Eastern women were ascribed less dominance than ethnically unspecified women, $t(427) = -8.70$, $p < .001$, $d = –0.42$, $CI_{95\%}$ [–0.52, –0.32]. Middle Eastern men and ethnically unspecified men, by contrast, did not significantly differ, $t(427) = 1.15$, $p = .253$, $d = 0.06$, $CI_{95\%}$ [–0.04, 0.15].

Image-Assessment Phase

Analysis Strategy. We fit separate LMEMs for image raters' trustworthiness and dominance impressions of the image generators' images. The models included fixed effects for target gender, target ethnicity, and their interaction. The final random-effects structure for both models included by-image and by-rater random intercepts. Figure 4 displays the pattern of means for trustworthiness (left panel) and dominance (right panel).

Trustworthiness. Like Experiment 1's indirectly assessed character representations and Experiment 2's directly assessed character representations, images of women were ascribed more trustworthiness than images of men, $b = 0.47$, $CI_{95\%}$ [0.40, 0.53], $t(420.72) = 13.54$, $p < .001$. Also, like Experiment 1's indirectly assessed character representations but unlike Experiment 2's directly assessed character representations,

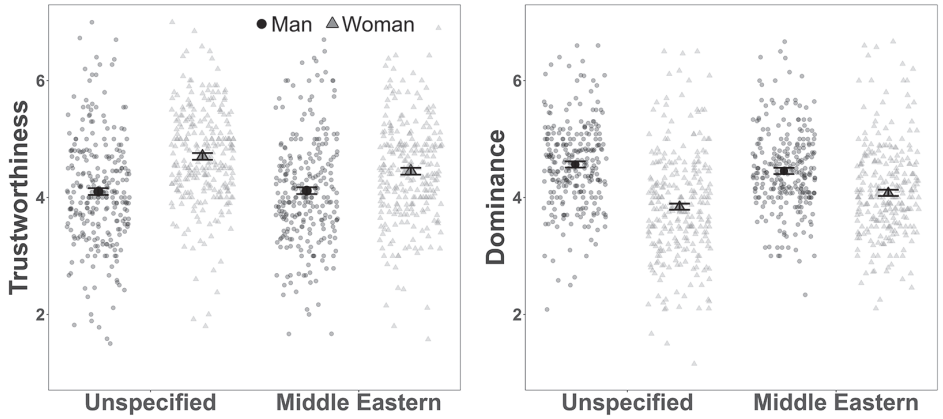

FIGURE 4. Image raters' trustworthiness and dominance impressions of participant-level classification images (i.e., facial representations) by target ethnicity and gender (Experiment 2). Markers reflect image ratings on trustworthiness and dominance. Semitransparent markers reflect participant-level data; filled shapes and their error bars reflect estimated marginal means and their 95% confidence intervals, respectively, from the linear mixed-effects models. The x-axis displays whether the image depicted an ethnically unspecified or Middle Eastern person. Shading and shape of markers reflect whether the image depicted a man or woman.

Middle Eastern images were ascribed less trustworthiness than ethnically unspecified images, b = -0.12, $CI_{95\%}$ [–0.19, –0.05], $t(419.09) = –3.39$, $p < .001$.

A significant target gender × target ethnicity interaction, b = –0.27, $CI_{95\%}$ [–0.41, –0.13], $t(423.33) = –3.85$, $p < .001$, further indicated that, like Experiment 1's indirectly assessed character representations and Experiment 2's directly assessed character representations, the gender difference (i.e., women more trustworthy than men) emerged for Middle Eastern images, b = 0.33, $CI_{95\%}$ [0.22, 0.44], $t(424.34) = 6.71$, $p < .001$, and ethnically unspecified images, b = 0.60, $CI_{95\%}$ [0.49, 0.71], $t(419.66) = 12.38$, $p < .001$, but it was larger for the latter. Also, like Experiment 1's indirectly assessed character representations but unlike Experiment 2's directly assessed character representations, images of Middle Eastern women were ascribed less trustworthiness than images of ethnically unspecified women, b = –0.26, $CI_{95\%}$ [–0.35, –0.16], $t(422.41) = –5.19$, $p < .001$. Images of Middle Eastern men and ethnically unspecified men, by contrast, did not significantly differ, b = 0.01, $CI_{95\%}$ [–0.08, 0.11], $t(420.09) = 0.32$, $p = .761$.

Dominance. Like Experiment 1's indirectly assessed character representations and Experiment 2's directly assessed character representations, images of women were ascribed less dominance than images of men, b = –0.55, $CI_{95\%}$ [–0.62, –047], $t(426.27) = –14.59$, $p < .001$. Unlike Experiment 1's indirectly assessed character representations and Experiment 2's directly assessed character representations, however, the target ethnicity main effect was not significant, b = 0.06, $CI_{95\%}$ [–0.01, 0.14], $t(423.97) = 1.68$, $p = .093$.

A significant target gender × target ethnicity interaction, $b = 0.35$, $CI_{95\%}$ [0.20, 0.50], $t(430.21) = 4.56$, $p < .001$, further indicated that the gender difference (i.e., women less dominant than men) emerged for Middle Eastern images, $b = –0.37$, $CI_{95\%}$ [–0.50, –0.25], $t(431.68) = -6.95$, $p < .001$, and ethnically unspecified images, $b = –0.72$, $CI_{95\%}$ [–0.85, –0.60], $t(424.73) = –13.63$, $p < .001$, but it was larger for the latter. Also, unlike Experiment 1's indirectly assessed character representations and Experiment 2's directly assessed character representations and contradicting stereotypes about Middle Eastern women's submissiveness, images of Middle Eastern women were ascribed *more* dominance than images of ethnically unspecified women, $b = 0.24$, $CI_{95\%}$ [0.13, 0.36], $t(428.56) = 4.43$, $p < .001$. Images of Middle Eastern men and ethnically unspecified men, by contrast, did not significantly differ, $b = –0.10$, $CI_{95\%}$ [–0.21, 0.01], $t(425.07) = –1.98$, $p = .052$.

Correspondence Between Character Representations and Facial Representations

Finally, we observed medium-sized positive correlations between the Phase 1 image generators' impressions of their assigned group and the Phase 2 image raters' impressions of the generated images on both traits—trustworthiness: $r(429) = .28$, $p < .001$; dominance: $r(430) = .31$, $p < .001$. The magnitude of these correlations suggests moderate correspondence between directly assessed character representations and indirectly assessed facial representations of the groups examined here (cf. Axt et al., 2023).

DISCUSSION

Many of Experiment 1's findings replicated in at least one of the phases in Experiment 2. The overarching gender differences on trustworthiness (i.e., women more trustworthy than men) and dominance (i.e., women less dominant than men) replicated across the image generators' direct reports and the assessments of their generated images. Middle Eastern facial images were ascribed less trustworthiness than ethnically unspecified facial images, just as the character profiles were in Experiment 1. Importantly, however, image generators explicitly rated Middle Eastern people, especially Middle Eastern men, as *more* trustworthy than ethnically unspecified people, potentially because they regulated their responses when responding directly.

Furthermore, image generators' explicit dominance ratings replicated Experiment 1's findings of Middle Eastern women being ascribed less dominance than ethnically unspecified women, but image raters' impressions of the generated facial images revealed the opposite pattern. The differences in patterns among the three assessment types, along with the moderate correlations between direct ratings and image ratings, suggest that indirectly assessed facial representations of Middle Eastern men and women may be capturing something that is distinct from indirectly assessed and directly assessed character representations of these groups.

GENERAL DISCUSSION

Two experiments examined mental representations of Middle Eastern versus ethnically unspecified women and men. In both experiments, women were viewed as more trustworthy and less dominant than men. Also, Middle Eastern people were viewed as less trustworthy than ethnically unspecified people when mental representations were assessed indirectly, but the opposite pattern emerged when they were assessed directly. Specifically, directly assessed character representations of Middle Eastern men were more trustworthy than those of ethnically unspecified men, which might reflect participants' attempts to regulate their responses. Directly assessed and indirectly assessed character representations, as well as indirectly assessed facial representations, of Middle Eastern women were more trustworthy than those of Middle Eastern men but less trustworthy than those of ethnically unspecified women.

Furthermore, character representations of Middle Eastern women—both directly assessed and indirectly assessed—were less dominant than those of both Middle Eastern men and ethnically unspecified women, which aligns with stereotypes of Middle Eastern women's submissiveness (Ghavami & Peplau, 2013). Facial representations of Middle Eastern women, by contrast, were less dominant than those of Middle Eastern men but *more* dominant than those of ethnically unspecified women. This latter effect was unexpected and might be explained by the absence of a key symbol of Middle Eastern women's submissiveness: the hijab (Ahmed, 1992). Middle Eastern ethnicity is often confounded with Islam (Lajevardi & Oskooii, 2018), likely because most people in the MENA geographical region are Muslim (Pew Research Center, 2015). Indeed, "religious" was the top adjective listed for both Middle Eastern men and Middle Eastern women in Experiment 1. This ethnic–religious conflation, symbolized in part by the hijab, might help explain the submissiveness stereotype of Middle Eastern women (Byng, 2010).

Prior work on MENA facial representations has used male targets exclusively (Dotsch et al., 2008, 2011; Hong et al., 2024; Petsko et al., 2021). Despite there being religious symbols worn by Muslim men (e.g., the taqiyah), ethnic-religious conflation might not have arisen in this past work because the most evocative appearance-based symbols of Islam are arguably those worn by Muslim women (Ahmed, 1992). To address this issue, future work could independently manipulate target religion alongside target ethnicity during the image-generation phase. Additionally, using a base-face image with (vs. without) a hijab (cf. Unkelbach et al., 2010) could determine if its presence prompts the generation of more submissive facial representations.

Notably, the category Middle Eastern is multifaceted, comprising multiple religions and nationalities. Future work could explore if more prototypical Middle Eastern subgroups (e.g., Iraqi Muslims) are represented differently than less prototypical subgroups (e.g., Armenian Christians), or if an overarching representation is applied regardless of religion or nationality.

This work has several limitations, which suggest additional avenues for future research. For example, our base-face image in Experiment 2 was a morph of Asian, Black, Latina/o/e/x, and White men's and women's faces (Gallagher &

Bodenhausen, 2021). Thus, the facial representations generated here—and indeed in prior work on MENA facial representations that used a White male base-face image (Dotsch et al., 2008, 2011; Hong et al., 2024; Petsko et al., 2021)—were necessarily constrained in how much they could vary. Future work could explore whether the current findings replicate when using a different base-face image (e.g., a morph that includes Middle Eastern men's and women's faces).

Furthermore, analyzing participant-level images in Experiment 2 maximized variation while minimizing Type-I error (Cone et al., 2021), but there is a limit to how much an image can reflect what is in an image generator's mind (Eisenstadt et al., 2025). It is also unknown if the image raters spontaneously encoded the faces in terms of the same identity combination (e.g., Middle Eastern woman) that each image generator intended. Future research could test this possibility.

Similarly, although Phase 1 participants in Experiment 1 could list the exact character attributes that came to mind for their assigned target group, Phase 2 participants might not have interpreted them as Phase 1 participants intended. For example, "strong" was the top attribute listed for both men and women, but "strong" might have different meanings (e.g., physical vs. emotional strength) depending on whom it describes. Phase 2 participants had to use the context of the other words in the profile to infer any difference in meaning (Asch, 1946). Furthermore, these attributes may not capture Middle Eastern women's experiences, especially in the U.S. where Middle Eastern women report misfit between how they view themselves and how (American) society views them (Abdel-Salam et al., 2019).

These experiments are also limited by our convenience samples of participants. Although the participant demographics were typical for psychological research, our participants were exclusively recruited in the U.S., where salient events (e.g., the September 11, 2001, terrorist attacks) might inform mental representations of Middle Eastern people more than they do elsewhere. Also, except for Phase 1 of Experiment 2, our samples were majority White. Thus, our findings may not generalize to mental representations generated by participants in other countries or with other racial/ethnic backgrounds. Future work could explore how mental representations of Middle Eastern (vs. ethnically unspecified) women and men vary in different cultural contexts and in more racially/ethnically diverse participant samples.

Our findings indicate that Middle Eastern women are mentally represented differently from Middle Eastern men (i.e., more trustworthy and less dominant in character and facial appearance) and women more broadly (i.e., less trustworthy in character and facial appearance, less dominant in character but more dominant in facial appearance). This work illuminates the views held about Middle Eastern women—views that are not fully captured by their constituent ethnic and gender categories.

Disclosure of interest statement. The authors declare no conflicts of interest.

Funding statement. The authors received no specific funding for this work.

Ethics statement. This study was reviewed and approved by the University of California, Davis's Institutional Review Board, and the procedures followed were in accordance with the Helsinki Declaration as revised in 2013. Participants provided their informed consent.

Supplementary materials. Supplementary materials are available online at https://doi.org/10.1521/soco.2026.44.2.150

REFERENCES

Abdel-Salam, L., Rifkin, R., Smith, L., & Zaki, S. (2019). Experiences of gender among Arab American women: A qualitative study. *Journal of Counseling Psychology, 66*(3), 255–268. https://doi.org/10.1037/cou0000329

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review, 128*(2), 290–314. https://doi.org/10.1037/rev0000262

Ahmed, L. (1992). *Women and gender in Islam: Historical roots of a modern debate*. Yale University Press.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258–290. https://doi.org/10.1037/h0055756

Axt, J., Siemers, N., Discepola, M. N., Martinez, P., Xiao, Z., & Wehrli, E. (2023). The mind's "aye"? Investigating overlap in findings produced by reverse correlation versus self-report. *Journal of Experimental Social Psychology*, *107*, 104473. https://doi.org/10.1016/j.jesp.2023.104473

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, *28*(1), 333–361. https://doi.org/10.1080/10463283.2017.1381469

Brown-Iannuzzi, J. L., Dotsch, R., Cooley, E., & Payne, B. K. (2017). The relationship between mental representations of welfare recipients and attitudes toward welfare. *Psychological Science, 28*(1), 92–103. https://doi.org/10.1177/0956797616674999

Byng, M. D. (2010). Symbolically Muslim: Media, hijab, and the West. *Critical Sociology, 36*(1), 109–129. https://doi.org/10.1177/089692050934713

Chaney, K. E., Sanchez, D. T., & Saud, L. (2021). White categorical ambiguity: Exclusion of Middle Eastern Americans from the White racial category. *Social Psychological and Personality Science, 12*(5), 593–602. https://doi.org/10.1177/1948550620930546

Cone, J., Brown-Iannuzzi, J. L., Lei, R., & Dotsch, R. (2021). Type I error is inflated in the two-phase reverse correlation procedure. *Social Psychological and Personality Science*, *12*(5), 760–768. https://doi.org/10.1177/1948550620938616

Dotsch, R. (2014). *rcicr: Reverse correlation image classification toolbox* (R package). http://ron.dotsch.org/rcicr

Dotsch, R., Wigboldus, D. H., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science*, *19*(10), 978–980. https://doi.org/10.1111/j.1467-9280.2008.02186.x

Dotsch, R., Wigboldus, D. H. J., & van Knippenberg, A. (2011). Biased allocation of faces to social categories. *Journal of Personality and Social Psychology, 100*(6), 999–1014. https://doi.org/10.1037/a0023026

Dunham, Y. (2018). Mere membership. *Trends in Cognitive Sciences, 22*(9), 780–793. https://doi.org/10.1016/j.tics.2018.06.004

Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin*, *15*(4), 543–558. https://doi.org/10.1177/0146167289154008

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Eisenstadt, B. E., Spagna, A., & Stroessner, S. J. (2025). In the mind's eye: Exploring the relationship between visual mental imagery and stereotyping. *Personality and Social Psychology Review*, *29*(3), 250–272. https://doi.org/10.1177/10888683241273350

Galinsky, A. D., Hall, E. V., & Cuddy, A. J. (2013). Gendered races: Implications for interracial marriage, leadership selection, and athletic participation. *Psychological Science*, *24*(4), 498–506. https://doi.org/10.1177/0956797612457783

Gallagher, N. M., & Bodenhausen, G. V. (2021). Gender essentialism and the mental representation of transgender women and men: A multimethod investigation of stereotype content. *Cognition*, *217*, 104887. https://doi.org/10.1016/j.cognition.2021.104887

Ghavami, N., & Peplau, L. A. (2013). An intersectional analysis of gender and ethnic stereotypes: Testing three hypotheses. *Psychology of Women Quarterly*, *37*(1), 113–127. https://doi.org/10.1177/0361684312464203

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Hall, E. V., Hall, A. V., Galinsky, A. D., & Phillips, K. W. (2019). MOSAIC: A model of stereotyping through associated and intersectional categories. *Academy of Management Review, 44*(3), 643–672. https://doi.org/10.5465/amr.2017.0109

Hauser, D. J., Moss, A. J., Rosenzweig, C., Jaffe, S. N., Robinson, J., & Litman, L. (2023). Evaluating CloudResearch's approved group as a solution for problematic data quality on MTurk. *Behavior Research Methods*, *55*(8), 3953–3964. https://doi.org/10.3758/s13428-022-01999-x

Hong, Y., Maitner, A. T., & Ratner, K. G. (2024). Psychological effects of anti-Arab politics on American and Arab peoples' views of each other. *PLoS One*, *19*(5), e0301282. https://doi.org/10.1371/journal.pone.0301282

Hutchings, R. J., Morgan, I., Sherman, J. W., & Todd, A. R. (2024). Intergroup evaluative bias in facial representations of immigrants and citizens in the United States. *PLoS One, 19*(7), e0306872. https://doi.org/10.1371/journal.pone.0306872

Hutchings, R. J., Simpson, A. J., Sherman, J. W., & Todd, A. R. (2021). Perspective taking reduces intergroup bias in visual representations of faces. *Cognition*, *214*, 104808. https://doi.org/10.1016/j.cognition.2021.104808

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology, 102*(1), 116–131. https://doi.org/10.1037/a0025335

Kang, S. K., & Bodenhausen, G. V. (2015). Multiple identities in social perception and interaction: Challenges and opportunities. *Annual Review of Psychology 66*, 547–574. https://doi.org/10.1146/annurev-psych-010814-015025

Klein, S. A. W., Hutchings, R. J., & Todd, A. R. (2021). Revising mental representations of faces based on new diagnostic information. *Cognition, 217*, 104916. https://doi.org/10.1016/j.cognition.2021.104916

Koch, A., Yzerbyt, V., Abele, A., Ellemers, N., & Fiske, S. T. (2021). Social evaluation: Comparing models across interpersonal, intragroup, intergroup, several-group, and many-group contexts. *Advances in Experimental Social Psychology, 63*, 1–68. https://doi.org/10.1016/bs.aesp.2020.11.001

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–32. https://doi.org/10.18637/jss.v082.i13

Lajevardi, N., & Oskooii, K. A. R. (2018). Old-fashioned racism, contemporary Islamophobia, and the isolation of Muslim Americans in the age of Trump. *Journal of Race, Ethnicity, and Politics*, *3*(1), 112–152. https://doi.org/10.1017/rep.2017.37

Lei, R. F., Foster-Hanson, E., & Goh, J. X. (2023). A sociohistorical model of intersectional social category prototypes. *Nature Reviews Psychology*, *2*, 297–308. https://doi.org/10.1038/s44159-023-00165-0

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods*, *49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93–120. https://doi.org/10.1146/annurev.psych.51.1.93

Maghbouleh, N., Schachter, A., & Flores, R. D. (2022). Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. *Proceedings of the National Academy of Sciences, 119*(7), e2117940119. https://doi.org/10.1073/pnas.2117940119

Martin, A. E., & Slepian, M. L. (2021). The primacy of gender: Gendered cognition underlies the big two dimensions of social cognition. *Perspectives on Psychological Science*, *16*(6), 1143–1158. https://doi.org/10.1177/1745691620904961

Nosek, B. A. (2007). Implicit–explicit relations. *Current Directions in Psychological Science*, *16*(2), 65–69. https://doi.org/10.1111/j.1467-8721.2007.00477.x

Petsko, C. D., & Bodenhausen, G. V. (2019). Racial stereotyping of gay men: Can a minority sexual orientation erase race? *Journal of Experimental Social Psychology, 83*, 37–54. https://doi.org/10.1016/j.jesp.2019.03.002

Petsko, C. D., Lei, R. F., Kunst, J. R., Bruneau E., & Kteily, N. (2021). Blatant dehumanization in the mind's eye: Prevalent even among those who explicitly reject it? *Journal of Experimental Psychology: General, 150*(6), 1115–1131. https://doi.org/10.1037/xge0000961

Pew Research Center. (2015, April). *The future of world religions: Population growth projections, 2010–2050.* https://www.pewresearch.org/religion/2015/04/02/middle-east-north-africa/

Qutub, A. (2013). Harem girls and terrorist men: Media misrepresentations of Middle Eastern cultures. *Colloquy*, *9*, 139–155.

Saleem, M., & Anderson, C. A. (2013). Arabs as terrorists: Effects of stereotypes within violent contexts on attitudes, perceptions, and affect. *Psychology of Violence, 3*(1), 84–99. https://doi.org/10.1037/a0030038

Samari, G. (2023). Anti-Muslim racism: The double burden of racism and invisibility. In E. Tuck, K. W. Yang, & J. Nixon (Eds.), *New approaches to inequality research with youth* (pp. 78–90). Routledge.

Samari, G., Alcalá, H. E., & Sharif, M. Z. (2018). Islamophobia, health, and public health: A systematic literature review. *American Journal of Public Health, 108*(6), e1–e9. https://doi.org/10.2105/AJPH.2018.304402

Schug, J., Alt, N. P., & Klauer, K. C. (2015). Gendered race prototypes: Evidence for the non-prototypicality of Asian men and Black women. *Journal of Experimental Social Psychology*, *56*, 121–125. https://doi.org/10.1016/j.jesp.2014.09.012

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149–178. https://doi.org/10.1002/ejsp.2420010202

Tamir, D. I., & Thornton, M. A. (2018). Modeling the predictive social mind. *Trends in Cognitive Sciences, 22*(3), 201–212. https://doi.org/10.1016/j.tics.2017.12.005

Thiem, K. C., Neel, R., Simpson, A. J., & Todd, A. R. (2019). Are Black women and girls associated with danger? Implicit racial bias at the intersection of target age and gender. *Personality and Social Psychology*

Bulletin, *45*(10), 1427–1439. https://doi.org/10.1177/0146167219829182

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460. https://doi.org/10.1016/j.tics.2008.10.001

Unkelbach, C., Schneider, H., Gode, K., & Senft, M. (2010). A turban effect, too: Selection biases against women wearing Muslim headscarves. *Social Psychological and Personality Science*, *1*(4), 378–383. https://doi.org/10.1177/1948550610378381

Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science*, *26*(3), 237–242. https://doi.org/10.1177/0963721416683996