UC Davis
UC Davis Electronic Theses and Dissertations

Title
Perspective Taking Reduces Intergroup Bias in Visual Representations of Faces

Permalink
https://escholarship.org/uc/item/06t331zg

Author
Hutchings, Ryan

Publication Date
2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Perspective Taking Reduces Intergroup Bias in Visual Representations of Faces

By

RYAN HUTCHINGS
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Psychology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

____

Jeffrey Sherman, Chair

____

Alison Ledgerwood

____

Andrew Todd

Committee in Charge

2021

Abstract

Intergroup biases shape most aspects of person construal, including lower-level visual representations of group members' faces. Specifically, ingroup members' faces tend to be represented more positively than outgroup members' faces. Here, we used a reverse-correlation paradigm to test whether engaging in perspective taking (i.e., actively imagining another person's mental states) can reduce these biased visual representations. In an initial image-generation experiment, participants were randomly assigned to a minimal group and then composed a narrative essay about an ingroup or an outgroup target person, either while adopting the person's perspective or while following control instructions. Afterward, they generated an image of the person's face in a reverse-correlation image-classification task. Subsequent image-assessment experiments using an explicit rating task, a sequential priming task, and an economic trust game with separate samples of participants revealed that ingroup faces elicited more likability and trustworthiness than did outgroup faces. Importantly, this pattern of intergroup bias was consistently weaker in faces created by perspective takers. Additional image-assessment experiments identified the mouth (i.e., smiling cues) as a critical facial region wherein the interactive effects of group membership and perspective taking emerged. These findings provide initial evidence that perspective taking may be an effective strategy for attenuating, though not for eliminating, intergroup biases in visual representations of what group members look like.

Perspective Taking Reduces Intergroup Bias in Visual Representations of Faces

A venerable literature spanning multiple areas of psychology has identified the developmental origins, the cognitive and neurobiological underpinnings, and the many judgmental and behavioral consequences of classifying other people as "us" versus "them" (Cikara & Van Bavel, 2014; Liberman et al., 2017; Macrae & Bodenhausen, 2000). Emerging from this research are countless demonstrations that ingroup members are commonly preferred over outgroup members (Hewstone et al., 2002)—even when the basis for group membership is arbitrary, as in the minimal-group paradigm (Dunham, 2018; Tajfel et al., 1971). So deep-seated is such intergroup bias that it can shape most aspects of person construal, including initial impressions of group members' faces (Hehman et al., 2019; Kawakami et al., 2017).

Importantly, these initial face impressions can have substantial outcomes for targets (see Todorov et al., 2015, for a review). For instance, CEOs and politicians with competent-looking faces enjoy more financial success (Rule & Ambady, 2008, 2009) and electoral success (Ballew & Todorov, 2007; Todorov et al., 2005), respectively, whereas convicted murderers with untrustworthy-looking faces are more likely to receive death sentences (Wilson & Rule, 2015, 2016). Documenting intergroup biases in the visual representations underlying perceivers' face impressions when they exist—and identifying potential strategies for de-biasing these representations—thus may provide useful insights for pressing social issues. In the research reported here, we tested the efficacy of perspective taking as one such bias-mitigation strategy.

Intergroup Bias in Visual Representations of Faces

Accumulating evidence using reverse-correlation techniques has revealed robust patterns of intergroup bias in how people visually represent group members' faces (Brinkman et al., 2017). The reverse-correlation image-classification paradigm (Mangini & Biederman, 2004), which has its roots in techniques for identifying the stimulus information underlying low-level perceptual classification judgments (Ahumada & Lovell, 1971), is a data-driven approach for visualizing the specific facial features underlying face classifications. Social cognition researchers have increasingly used this paradigm to elucidate the facial features underlying face classifications of different social categories, which affords visual estimates of the prototypical face of particular social groups in perceivers' minds (Brinkman et al., 2017).

In a typical two-image forced-choice (2IFC) reverse-correlation image-classification task (Dotsch & Todorov, 2012), participants in an *image-generation* phase complete hundreds of trials in which they select which of two noise-laden face images more closely resembles their internal representation of a person according to a dimension of interest. For example, participants may view several hundred pairs of face images and select which image looks more like a Moroccan person (Dotsch et al., 2008). The noise patterns selected by a given participant are then combined with the underlying base face to create a classification image (CI) corresponding to that participant's visual depiction of Moroccan faces.[1] Then, in a separate *imagine-assessment* phase, a naïve sample of participants rates the CIs on one or more dimensions of interest (e.g., a trait, such as trustworthiness). The responses of these image raters provide insights into the minds of the image generators—for example, that their visual representations of Moroccan faces look untrustworthy (Dotsch et al., 2008).

Experiments using this paradigm have documented a variety of intergroup biases in visual representations of social groups. Specifically, representations of ingroup faces tend to look more positive than do representations of outgroup faces. For example, Europeans generate untrustworthy-looking visual representations of Moroccan and Chinese faces (Dotsch et al., 2008), and Republicans generate more favorable representations of Republican politicians than Democrats do (Young et al., 2014). Such biases even extend to minimal group members' faces. In one study, participants were randomly assigned to a minimal group (dot over-estimators or under-estimators) and then generated a representation of an ingroup member's face or an outgroup member's face (Ratner et al., 2014). Separate samples of participants explicitly ascribed more positive (and fewer negative) traits to ingroup versus outgroup faces, implicitly evaluated ingroup faces more positively than outgroup faces, and displayed more trusting behavior toward ingroup versus outgroup faces in an economic trust game. These and other findings (e.g., Bjornsdottir et al., 2019; Brown-Iannuzzi et al., 2017; Hong & Ratner, 2021; Lei & Bodenhausen, 2017; Paulus et al., 2016; Petsko et al., 2020) attest to the prevalence of intergroup bias in visual representations of faces.

[1] Because such *individual* CIs can be noisy, a common procedure entails aggregating across all the individual CIs in a particular experimental condition to generate a *group* CI reflecting the average visual representation of the participants in that condition. Most social-cognitive research on visual representations of social groups focuses on these group CIs (Brinkman et al., 2017).

Perspective Taking and Intergroup Bias Reduction

Abundant research has explored strategies for reducing various forms of intergroup bias (Paluck & Green, 2009). To our knowledge, however, no studies have examined strategies for reducing intergroup bias in visual representations of group members' faces, making it unclear whether strategies known to reduce biases in *conceptual* representations of a group also produce shifts in *visual* representations of that group and its members. Indeed, few studies have examined situational influences of any kind on visual representations of social groups (for an exception, see Krosch & Amodio, 2014). Here, we examined the viability of perspective taking (i.e., actively imagining another person's mental states) as a strategy for attenuating intergroup biases in visual representations of group members' faces.

Evidence from multiple studies suggests that perspective taking can temper both overt and covert expressions of intergroup bias (e.g., Batson et al., 1997; Broockman & Kalla, 2016; Galinsky & Moskowitz, 2000; Shih et al., 2009; Todd, Bodenhausen et al., 2011; see Ku et al., 2015; Todd & Galinsky, 2014, for reviews). In one set of studies, for example, implicit evaluative biases toward Turkish people, Black people, and older adults were weaker among participants who, in a preceding task, had actively adopted the perspective of an individual Turkish, Black, or older adult target, respectively (Todd & Burgmer, 2013). Insofar as these positive effects of perspective taking on conceptual representations of group members are also apparent in visual prototypes that guide initial impression formation (Brinkman et al., 2017), the potential impact of perspective taking may be wide-ranging on downstream forms of intergroup bias that are built upon those visual prototypes.

One reason for perspective taking's effectiveness in reducing intergroup bias is that, much like other de-biasing strategies that prompt deviation from default modes of information processing (e.g., Kleiman et al., 2014; Lord et al., 1984), actively imagining the world from another person's point of view encourages perceivers to step outside their usual mental routines and to represent and process social information differently than they typically do (Todd et al., 2012, 2016). By default, outgroup members tend to be represented and processed in ways that accentuate their differences from the ingroup and the self (Ames, 2004a, 2004b; Gawronski et al., 2005; Todd, Hanko et al., 2011), commonly resulting in evaluative biases whereby outgroup members (via their relatively weaker connection to the self) are viewed less favorably than ingroup members (Otten, 2003). Perspective taking can undermine this default tendency, leading perceivers to represent and process outgroup targets more favorably than they otherwise might (Ku et al., 2015; Todd & Galinsky, 2014).

If perspective taking can produce positive shifts in conceptual representations of what outgroup members are like, might it also be capable of positively altering visual representations of what they *look* like? We investigated this possibility in the current research.

Perspective Taking and Processing of Facial Features

How might perspective taking alter visual representations of group members' faces? Prior evidence indicates that, in the absence of perspective taking, visual representations of ingroup faces look more positive and trustworthy than do visual representations of outgroup faces. Indeed, when compared pixel-by-pixel, visualizations of "trustworthy" faces are more physically similar to visualizations of ingroup versus outgroup faces (Ratner et al., 2014). Of particular relevance here is research linking facial trustworthiness to specific emotion expressions, including positive links to subtle happiness expressions (e.g., upward-turned mouths and open eyes resembling a smile) and negative links to subtle anger expressions (e.g., downward-turned mouths and lowered brows resembling a frown/scowl; Dotsch & Todorov, 2012; Oosterhof & Todorov, 2008, 2009; Sutherland et al., 2013). We, therefore, explored the possibility that the visualizations of outgroup faces generated by perspective takers look more trustworthy by virtue of their increased appearance of smiling and/or their decreased appearance of frowning/scowling.

The Current Research

The current research examined whether and how perspective taking alters visual representations of ingroup and outgroup members' faces. Our investigation comprised an image-generation experiment followed by eight image-assessment experiments. In the image-generation experiment, participants learned about a minimal ingroup or outgroup target person while actively adopting the person's perspective or while following control instructions. Then, in a reverse-correlation image-classification task, participants constructed a visual representation of the (ingroup or outgroup) target person. The image-assessment experiments, described in more detail below, assessed these visual representations along various dimensions (e.g., likability, trustworthiness) in new samples of participants who were naïve as to how the visual representations were constructed in the image-generation experiment.

Overall, we expected that ingroup faces would be visually represented more favorably than would outgroup faces (e.g., Ratner et al., 2014). Based on prior evidence that engaging in perspective taking can reduce intergroup biases in conceptual representations (Todd & Galinsky, 2014), we further predicted that perspective taking would temper this pattern of intergroup bias in visual representations. Specifically, we anticipated that adopting an outgroup member's perspective would increase the likability and trustworthiness of outgroup representations. Moreover, we hypothesized that perspective taking would affect subtle facial expressions (i.e., smiling and frowning/scowling) related to approachability, with changes in facial expression representations potentially mediating changes in trustworthiness impressions.

Image-Generation Experiment

Method

Participants

To our knowledge, no formal power analysis procedures exist for the image-generation phase in reverse-correlation paradigms (see also Brown-Iannuzzi et al., 2021). Thus, we set our target sample size for the image-generation experiment using a heuristic of at least 60 participants per cell in a 2 × 2 between-subjects design; data were collected until this target number was surpassed. Undergraduates (N = 257) participated for course credit. We excluded data from participants (n = 3) who did not finish the reverse-correlation task. The final sample comprised 254 White participants (159 women, 95 men; M_{age} = 19.28, SD_{age} = 1.25).

Procedure

In this image-generation experiment and in all subsequent image-assessment experiments, we obtained informed consent from participants prior to their participation. The data for all experiments are available on the Open Science Framework: https://osf.io/5tqv7/.

The first task manipulated minimal-group membership (Bernstein et al., 2007). Participants rated their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with items from the Ten-Item Personality Inventory (Gosling et al., 2003). After a short delay during which the computer purportedly calculated their scores, participants learned that they had either a "green" personality type or an "orange" personality type (in actuality, group membership was randomly assigned). Although no description of the personality types was provided, participants were informed that the personality measure predicts social and financial success and that businesses and scientists use it in hiring and research. In this way, group membership was minimal but likely meaningful for participants. Participants also received a green or an orange wristband to wear for the remainder of the experimental session as a reminder of their group membership (for a similar minimal-group induction, see Simpson & Todd, 2017).

Next, participants wrote a short essay about a typical day in the life of an unknown, gender-matched target person who, depending on condition, had a green or an orange personality type. Thus, the target was an ingroup member for participants with the same personality type and an outgroup member for participants with the other personality type. Participants received one of two sets of instructions for writing their essay (Todd et al., 2016). Instructions in the *perspective-taking* condition encouraged active consideration of the person’s mental states:

> As you’re writing, we ask that you take [his or her] perspective. In your mind’s eye, visualize clearly and vividly what [he or she] might be thinking and feeling, and what [his or her] intentions and goals are.

Instructions in the *control* condition encouraged a more objective and detached focus:

> As you’re writing, we ask that you adopt an objective focus. Try not to get caught up in what [he or she] might be thinking or feeling, or what [his or her] intentions or goals are. Just write as if you were a casual observer.

Participants then completed a 2IFC reverse-correlation image-classification task (Brinkman et al., 2017) during which they generated a visual representation of the target person. On each of 300 trials, participants viewed two adjacent images of degraded faces (211 × 270 pixels) and selected which face looked more like the target person from the essay-writing task.

The face stimuli were created with the *rcicr* 0.3.4.1 package (Dotsch, 2014) in R (Version 3.6.0; R Core Team 2019) by superimposing noise patterns onto 2 base faces (1 male, 1 female). The base faces were the neutral male and female faces from the Averaged Karolinska Directed Emotion Face Database (Lundqvist & Litton, 1998). The noise patterns were 4,092 superimposed truncated sinusoid patches in all possible combinations of 2 cycles in 6 orientations (0°, 30°, 60°, 90°, 120°, 150°) × 5 spatial frequencies (1, 2, 4, 8, 16 patches per image) × 2 phases (0, π/2), with random contrasts. For each trial, we generated a unique noise pattern and combined it with the base face to create one of the two face stimuli. The other face stimulus was the inverse of the noise pattern and the same base face. This technique maximizes the contrast between the two faces presented on each trial (Dotsch & Todorov, 2012). The order of the stimulus pairs and the left-right position of the stimuli in each pair were randomized.

Image Processing

Also using the *rcicr* 0.3.4.1 package (Dotsch, 2014), we created 3 types of classification images (CIs) that were later evaluated in the image-assessment experiments: group CIs (used in image-assessment Experiment 1), individual CIs (used in image-assessment Experiments 2B and 3–5), and subgroup CIs (used in image-assessment Experiments 2A and 6). We created group CIs by superimposing onto the base face the average noise patterns of all the selected images across all participants in each condition, resulting in 4 group CIs of each gender based on group membership and perspective taking. We created separate male and female group CIs because male and female participants completed the image-classification task with different, gender-matched base faces (Fig. 1 displays the group CIs by condition). By averaging across all the classification judgments from all the image generators within a condition, group CIs have the advantage of amplifying the shared signal in the CIs due to the effect of the manipulation while reducing idiosyncratic differences between the CIs that might reflect random responding (Lei et al., 2020). However, the group CI procedure has clear disadvantages. Using group CIs necessarily eliminates all the variance among the image generators by combining all the classification judgments from an experimental condition into a single CI, and thus this procedure is only appropriate when there is sufficient agreement in the classification judgments of the image generators (Brinkman et al., 2017; Dotsch & Todorov, 2012). Moreover, recent research suggests that two-phase reverse-correlation procedures that rely exclusively on group CIs can artificially inflate differences between conditions, thereby increasing Type I error (Cone et al., 2020).

Importantly, two-phase procedures that preserve at least some of the variation among the image generators do not inflate Type I error (Cone et al., 2020; Lei et al., 2020). Thus, in all the image-assessment experiments except for Experiment 1, we used procedures that maintain some of the variation (i.e., subgroup CIs) or all the variation (i.e., individual CIs) among the image generators. We generated male and female individual CIs by averaging all the selected noise patterns for each participant and superimposing this noise pattern onto the base face. We completed this procedure separately for female and male participants. Next, we created subgroup CIs by following the procedure outlined in Lei et al. (2020). This procedure entails averaging random subsets of individual CIs in each experimental condition, resulting in multiple subgroup CIs per condition. The subgroup-CI approach was developed in response to Cone et al.'s (2020) findings concerning the inflation of Type I error in two-phase reverse correlation procedures. Preliminary evidence indicates that the subgroup-CI approach, in which multiple aggregated CIs are generated in each condition, controls Type I error when at least 2 subgroup CIs are created per condition (Lei et al., 2020).

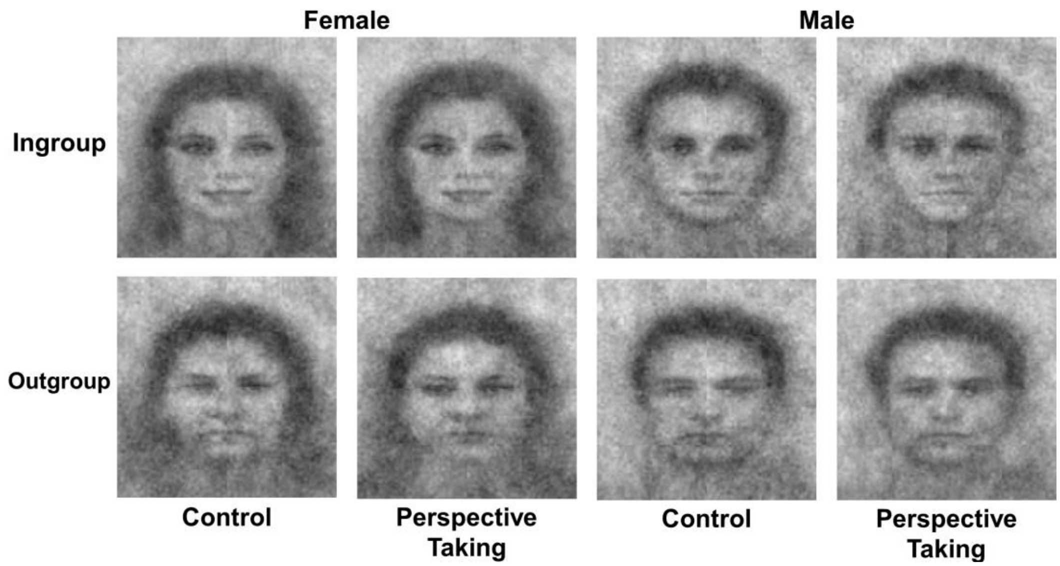

Figure 1. Group classification images of ingroup and outgroup targets in the control and perspective-taking conditions generated by female participants (left panel) and male participants (right panel).

We generated male and female subgroup CIs by averaging the selected noise patterns from random subsets of participants, without replacement, in each condition and superimposing this averaged noise pattern onto the base face. Because we had more data for female (n = 159) than male participants (n = 95), we created 6 subgroup CIs per condition for female participants but 4 subgroup CIs per condition for male participants. Each subgroup CI averaged randomly selected noise patterns from about 6 participants, resulting in a total of 40 subgroup CIs. We settled on this number of subgroup CIs to balance our goals of maximizing the number of stimuli while limiting rater fatigue. We completed this procedure separately for image-assessment Experiments 2A and 6 to ensure that our results cannot be explained as an artifact of the particular randomization of individual CIs in a given set of subgroup CIs.

Image-Assessment Experiments

To test our predictions about the impact of perspective taking on visual representations of ingroup and outgroup faces, we conducted several image-assessment experiments. In Experiment 1, a new sample of participants rated the group CIs on several traits known to be observable in face. Experiments 2A and 2B conceptually replicated Experiment 1 using subgroup and individual CIs, respectively, and focused on the trait of likability. Experiments 3A and 3B conceptually replicated Experiments 2A and 2B using a sequential priming task to examine the interactive effects of group membership and perspective taking on the spontaneously evoked likability and trustworthiness of the individual CIs. Experiment 4 tested whether these effects extend to cooperative behavior in an economic trust game. Experiment 5 explored which of several facial regions (i.e., eyes, mouth) are critical for understanding how perspective taking alters visual representations of group members' faces. Finally, Experiment 6 tested whether subtle differences in smiling or frowning/scowling expressions underlie the effects of perspective taking on the trustworthiness of visual representations of ingroup and outgroup members' faces.

Experiment 1

Method

Participants. Undergraduates (N = 153; 119 women, 34 men; 76 Asian, 38 White, 27 Latinx, 2 Black, 10 reporting other or multiple races/ethnicities; M_{age} = 19.96, SD_{age} = 2.21) participated for course credit. A sensitivity analysis indicated that this sample size afforded >80% power to detect a two-way interaction with a small-to-medium effect size (η_p^2 = 0.050; Faul et al., 2007). In this and all subsequent image-assessment experiments, participants were naïve as to how the facial images were constructed in the image-generation experiment.

Procedure. Participants used a slider scale (0 = *not at all*, 100 = *very*) to rate the gender-matched group CIs from each of the 4 conditions in the image-generation experiment on the following traits (order was randomized) known to be observable in faces: *trustworthy*, *attractive*, *dominant*, *caring*, *sociable*, *confident*, *emotionally stable*, *responsible*, *intelligent*, *aggressive*, *mean*, *weird*, and *unhappy* (Oosterhof & Todorov, 2008). The ratings appeared in two blocks (order was counterbalanced). In one block, participants rated the ingroup and outgroup CIs generated by participants in the control condition; in the other block, they rated the ingroup and outgroup CIs generated by participants in the perspective-taking condition. In each block, the ingroup and outgroup CIs appeared side-by-side, in counterbalanced positions, to facilitate comparison of the faces (Ratner et al., 2014). In all subsequent image-assessment experiments, the CIs appeared one at a time.

Results

Multivariate analyses. We first conducted a repeated-measures multivariate analysis of variance (rMANOVA) testing the effects of Group Membership, Perspective Taking, and their interaction on the average ratings across traits (see Hong & Ratner, 2021, for a similar analytic approach).[2] There were significant main effects of Group Membership, Pillai's Trace = 0.69, $F(13, 140) = 23.45$, $p < .001$, and Perspective Taking, Pillai's Trace = 0.22, $F(13, 140) = 3.06$, $p = .001$. More important for the current work, the predicted Group Membership × Perspective Taking interaction was also significant, Pillai's Trace = 0.36, $F(13, 140) = 5.95$, $p < .001$.

To decompose this interaction, we examined the simple effect of group membership separately on the CIs generated in the perspective-taking and control conditions. Although sizable intergroup bias emerged in both conditions, the effect of group membership was weaker on the CIs generated by perspective takers, Pillai's Trace = 0.46, $F(13, 140) = 9.00$, $p < .001$, than on the CIs generated by control participants, Pillai's Trace = 0.70, $F(13, 140) = 25.00$, $p < .001$. Approaching the interaction differently, we tested the simple effect of perspective taking separately on the ingroup and outgroup CIs. A significant effect of perspective taking emerged on the outgroup CIs, Pillai's Trace = 0.37, $F(13, 140) = 6.25$, $p < .001$, but not on the ingroup CIs, Pillai's Trace = 0.11, $F(13, 140) = 1.32$, $p = .208$.

Univariate analyses. We next conducted analogous univariate analyses of variance (ANOVAs) separately for each trait. A significant Group Membership main effect emerged on every trait except dominance. The ingroup CIs were rated as more attractive, caring, confident, emotionally stable, intelligent, responsible, sociable, and trustworthy, and as less aggressive, mean, unhappy, and weird than were the outgroup CIs. A significant Perspective Taking main effect also emerged on multiple traits. The CIs generated by perspective takers were rated as more attractive, caring, confident, emotionally stable, intelligent, responsible, sociable, and trustworthy, and as less unhappy and weird than were the CIs generated by control participants. Finally, the Group Membership × Perspective Taking interaction was significant for every trait (Table 1 displays descriptive statistics for all conditions and *F* values for both main effects and the interaction).

[2] Across experiments, we conducted preliminary analyses that included participant/target gender; these analyses appear in the Appendix.

To understand these interactions, we first tested the impact of group membership separately on the CIs created by perspective takers and control participants. Although significant intergroup bias emerged on every trait except dominance in both conditions, it was weaker on perspective takers' CIs than on control participants' CIs. Next, we examined the impact of perspective taking separately on the ingroup and outgroup CIs. The outgroup CIs generated by perspective takers were rated as more attractive, caring, confident, emotionally stable, intelligent, responsible, sociable, and trustworthy, and as less aggressive, mean, unhappy, and weird than were the outgroup CIs generated by control participants. The ingroup CIs generated by perspective takers were also rated as more trustworthy and as less unhappy than were the ingroup CIs generated by control participants; for the remaining traits, none of the simple effects of perspective taking on the ingroup CIs was significant.

Discussion

The results of Experiment 1 revealed robust intergroup bias in visual representations of minimal group members' faces, replicating prior work (Ratner et al., 2014) but with a different minimal-group induction. The ingroup visualizations were rated higher on multiple positive traits (e.g., attractive, intelligent, trustworthy) and lower on multiple negative traits (e.g., aggressive, mean) than were the outgroup visualizations. Importantly, perspective taking dampened, but did not eliminate, this intergroup bias. Relative to control participants, perspective takers generated more positive (and less negative) outgroup visual representations; however, perspective taking had negligible effects on impressions of the ingroup visual representations.

Table 1

Univariate ANOVA Results for Trait Impression Ratings in Experiment 1

Trait	Control			Perspective Taking			*F* values		
	Ingroup Mean (*SD*)	Outgroup Mean (*SD*)	*d*	Ingroup Mean (*SD*)	Outgroup Mean (*SD*)	*d*	Group	PT	Group × PT
Aggressive	36.93 (24.53)	61.25 (25.28)	-0.68	40.06 (24.17)	52.26 (22.31)	-0.42	81.71^{***}	2.86	12.63^{***}
Attractive	64.90 (23.27)	29.02 (21.63)	1.25	65.32 (23.07)	43.69 (23.14)	0.80	234.54^{***}	34.68^{***}	32.35^{***}
Caring	63.50 (19.65)	38.88 (20.52)	0.96	60.73 (19.49)	47.93 (20.20)	0.51	137.45^{***}	3.97^{*}	21.20^{***}
Confident	64.52 (19.14)	41.16 (23.56)	0.77	65.05 (19.62)	48.93 (22.97)	0.56	120.00^{***}	5.98^{*}	5.30^{***}
Dominant	51.58 (20.60)	57.14 (25.95)	-0.18	55.05 (20.63)	54.13 (22.97)	0.03	1.73	0.02	4.45^{***}
Emotionally stable	59.60 (19.87)	37.29 (21.41)	0.79	56.25 (20.92)	52.05 (21.63)	0.17	70.95^{***}	10.37^{**}	37.60^{***}
Intelligent	62.99 (16.05)	47.23 (18.96)	0.70	62.56 (17.06)	56.67 (16.74)	0.30	63.70^{***}	11.53^{***}	22.32^{***}
Mean	38.27 (22.42)	62.80 (24.54)	-0.74	41.56 (24.45)	53.37 (21.26)	-0.40	85.64^{***}	2.99	15.78^{***}
Responsible	61.61 (18.44)	46.73 (21.19)	0.55	61.05 (18.33)	55.49 (19.41)	0.23	41.78^{***}	7.08^{**}	11.62^{***}
Sociable	64.98 (21.10)	30.99 (20.22)	1.17	63.80 (22.87)	42.75 (22.46)	0.69	222.96^{***}	11.38^{***}	17.27^{***}
Trustworthy	59.69 (19.43)	39.63 (20.09)	0.72	55.18 (19.97)	50.56 (19.00)	0.18	55.78^{***}	4.23^{*}	29.55^{***}
Unhappy	35.37 (22.52)	72.84 (23.16)	-1.17	40.63 (23.36)	57.25 (23.11)	-0.55	192.96^{***}	10.45^{**}	42.16^{***}
Weird	39.27 (23.97)	59.72 (21.74)	-0.70	40.80 (22.87)	50.53 (21.77)	-0.36	75.04^{***}	6.10^{*}	13.16^{***}

Note. Group = Group Membership. PT = Perspective Taking. $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$

Experiments 2A and 2B

Experiments 2A and 2B aimed to replicate Experiment 1's findings using subgroup and individual CIs rather than group CIs. As previously noted, individual and subgroup CIs preserve at least some of the variation among the image generators and, more importantly, do not inflate Type I error (Cone et al., 2020; Lei et al., 2020). We focused specifically on the trait of likability. We predicted that the ingroup CIs would be rated as more likable than the outgroup CIs, but that this group difference would be weaker on the CIs generated by perspective takers than on the CIs generated by control participants.

Method

Participants. Amazon's Mechanical Turk (MTurk) workers (Experiment 2A: N = 154; Experiment 2B: N = 584) participated for modest pay (Experiment 2A: $0.55; Experiment 2B: $0.75). We excluded data from participants who reported not taking the experiment seriously (Experiment 2A: n = 8; Experiment 2B: n = 47) and who did not finish the entire experiment (Experiment 2A: n = 9; Experiment 2B: n = 6). The final samples comprised 137 participants in Experiment 2A (75 women, 59 men, 2 non-binary, 1 unreported; 94 White, 14 Black, 12 Asian, 8 Latinx, 2 Native American, 3 reporting other or multiple races/ethnicities, 1 unreported; M_{age} = 35.76, SD_{age} = 10.92) and 531 participants in Experiment 2B (301 women, 226 men, 2 non-binary, 2 unreported; 402 White, 44 Black, 33 Asian, 22 Latinx, 4 Native American, 10 reporting other or multiple races/ethnicities, 9 unreported; M_{age} = 39.88, SD_{age} = 12.89). These samples afforded >80% power to detect two-way interactions with small-to-medium effect sizes (Experiment 2A: η_p^2 = 0.056; Experiment 2B: η_p^2 = 0.015; Faul et al., 2007).

Procedure. Participants in both experiments judged the likability (1 = *not at all likable*, 7 = *very likable*) of CIs from each of the four experimental conditions from the image-generation experiment. Instructions encouraged participants to provide their "gut-level" impressions. In Experiment 2A, female participants rated female subgroup CIs (n = 24), and male participants rated male subgroup CIs (n = 16). In Experiment 2B, male participants rated all 95 individual male CIs. Because there were many more individual female CIs (n = 159), we had female participants rate 95 randomly selected female CIs from this larger set to keep task duration comparable for male and female participants.

Results

Subgroup CIs (Experiment 2A). A 2 (Group Membership) × 2 (Perspective Taking) repeated-measures ANOVA[3] on the likability of the subgroup CIs in each condition revealed a significant Group Membership main effect, $F(1, 136) = 196.91$, $p < .001$, $\eta_p^2 = 0.59$, $CI_{90\%}$ [0.51, 0.66],[4] whereby the ingroup CIs ($M = 4.29$, $SD = 1.13$) were rated as more likable than were the outgroup CIs ($M = 3.23$, $SD = 1.18$). A significant Perspective Taking main effect, $F(1, 136) = 31.62$, $p < .001$, $\eta_p^2 = 0.19$, $CI_{90\%}$ [0.10, 0.28], indicated that the CIs generated by perspective takers ($M = 3.86$, $SD = 1.23$) were rated as more likable than were the CIs generated by control participants ($M = 3.66$, $SD = 1.31$). More important for the current research, the predicted interaction was also significant, $F(1, 136) = 9.29$, $p = .003$, $\eta_p^2 = 0.06$, $CI_{90\%}$ [0.01, 0.14] (see Fig. 2).

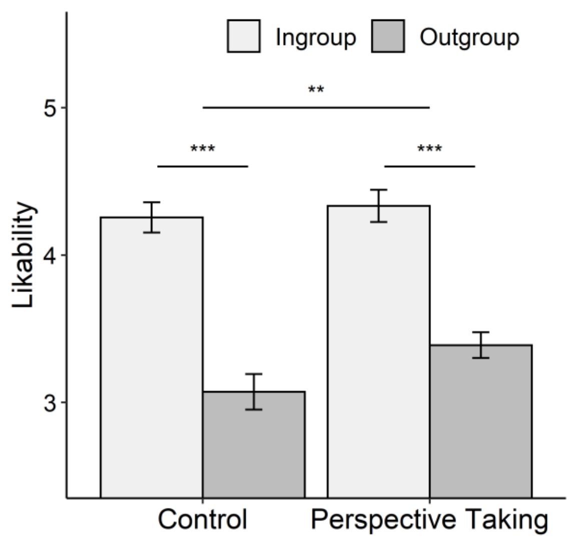

Figure 2. Likability explicitly ascribed to subgroup classification images by group membership and perspective taking in Experiment 2A. Error bars reflect 95% confidence intervals. ***$p < .001$, **$p < .01$.

Simple-effects tests indicated that, although robust intergroup bias emerged in both conditions, it was smaller for the CIs generated by perspective takers, $t(136) = 11.95$, $p < .001$, $d = 1.02$, $CI_{95\%}$ [0.82, 1.23], than for the CIs generated by control participants, $t(136) = 12.98$, $p < .001$, $d = 1.11$, $CI_{95\%}$ [0.90,

---

[3] In all remaining image-assessment experiments, we conducted analyses at the level of the participant (i.e., image rater) by collapsing across CIs, rather than at the level of the CI by collapsing across image raters. This analysis strategy is the only one reported in studies using subgroup CIs (Brown-Iannuzzi et al., 2021; Cone et al., 2020); however, both analysis strategies have been reported in studies using individual CIs. Notably, Cone et al. (2020) conducted their analyses at the level of the individual CI. We also report results at the level of the individual CI for Experiments 2B–5 in the Appendix.
[4] Following Steiger (2004), we report 90% confidence intervals for effect sizes accompanying *F* tests. We used the *effectsize* package (Ben-Shachar et al., 2020) to generate both the 90% confidence intervals for the η_p^2 estimates and the 95% confidence intervals for the Cohen's *d* estimates.

1.33]. Decomposing the interaction differently, the outgroup CIs generated by perspective takers (*M* = 3.39, *SD* = 1.15) were rated as more likable than were the outgroup CIs generated by control participants (*M* = 3.07, *SD* = 1.20), $t(136) = -6.11$, $p < .001$, $d = -0.52$, $CI_{95\%}$ [-0.70, -0.34]. In contrast, the difference in likability of the ingroup CIs generated by perspective takers (*M* = 4.33, *SD* = 1.12) and control participants (*M* = 4.26, *SD* = 1.14) was not significant, $t(136) = -1.47$, $p = .143$, $d = -0.13$.

Individual CIs (Experiment 2B). An identical 2 (Group Membership) × 2 (Perspective Taking) repeated-measures ANOVA on the likability of the individual CIs[5] in each condition revealed that the ingroup CIs (*M* = 3.90, *SD* = 0.83) were rated as more likable than were the outgroup CIs (*M* = 3.30, *SD* = 0.79), $F(1, 530) = 1{,}046.98$, $p < .001$, $\eta_p^2 = 0.66$, $CI_{90\%}$ [0.63, 0.69], and that the CIs generated by perspective takers (*M* = 3.62, *SD* = 0.85) were rated as more likable than were the CIs generated by control participants (*M* = 3.57, *SD* = 0.88), $F(1, 530) = 16.07$, $p < .001$, $\eta_p^2 = 0.03$, $CI_{90\%}$ [0.01, 0.06]. Once again, the predicted interaction was also significant, $F(1, 530) = 97.79$, $p < .001$, $\eta_p^2 = 0.15$, $CI_{90\%}$ [0.10, 0.19] (see Fig. 3).

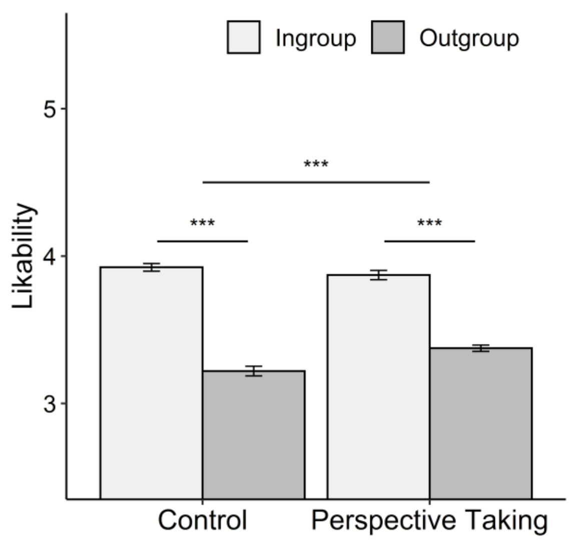

Figure 3. Likability explicitly ascribed to individual classification images by group membership and perspective taking in Experiment 2B. Error bars reflect 95% confidence intervals. $^{***}p < .001$.

[5] Collecting ratings on the individual CIs afforded the use of more conservative linear mixed-effects models, which can simultaneously account for participant (i.e., image raters) and stimulus (i.e., image generators' CIs) variance. The results of these analyses for Experiments 2B–5, which generally aligned with the results reported in the main text, appear in the Appendix. Significant Group Membership × Perspective Taking interactions emerged in Experiments 3A, 4, and 5; however, this interaction was non-significant (but in the predicted direction) in Experiments 2B and 3B. Overall, the magnitude of the Group Membership × Perspective Taking interaction was weaker when using analyses that simultaneously model random effects of participants *and* stimuli.

Simple-effects tests again indicated that, although sizable intergroup bias emerged in both conditions, it was weaker for the CIs generated by perspective takers, $t(530) = 24.69$, $p < .001$, $d = 1.07$, $CI_{95\%}$ [0.97, 1.18], than for the CIs generated by control participants, $t(530) = 30.95$, $p < .001$, $d = 1.34$, $CI_{95\%}$ [1.23, 1.46]. Approaching this interaction differently, perspective takers' outgroup CIs ($M = 3.38$, $SD = 0.78$) were rated as more likable than were control participants' outgroup CIs ($M = 3.22$, $SD = 0.80$), $t(530) = -9.69$, $p < .001$, $d = -0.42$, $CI_{95\%}$ [-0.51, -0.33], whereas perspective takers' ingroup CIs ($M = 3.87$, $SD = 0.84$) were rated as less likeable than were control participants' ingroup CIs ($M = 3.92$, $SD = 0.81$), $t(530) = 2.99$, $p = .003$, $d = 0.13$, $CI_{95\%}$ [0.04, 0.22].

Discussion

The results of Experiments 2A and 2B replicate those from Experiment 1, but this time using subgroup CIs and individual CIs rather than group CIs. We observed a robust pattern of intergroup bias whereby participants generated more likable ingroup visual representations than outgroup visual representations. Importantly, perspective taking attenuated, but did not eliminate, this pattern of intergroup bias. In Experiment 2A, perspective taking increased the likability ascribed to the outgroup representations but had no discernible impact on the likability ascribed to the ingroup representations. In Experiment 2B, however, perspective taking increased the likability of the outgroup representations and decreased the likability of the ingroup representations, relative to the control condition.

Experiments 3A and 3B

The first set of image-assessment experiments revealed that perspective taking reduced biases in the positivity ascribed to visual representations of ingroup faces versus outgroup faces. All of these image-assessment experiments, however, used direct measures of trait impressions wherein the image raters explicitly judged the likability of the CIs. Evidence indicates that people form consensual trait impressions of faces after minimal exposure (Willis & Todorov, 2006). People also spontaneously infer traits from faces, even when these inferences are irrelevant to their focal task goal (Klapper et al., 2016). Insofar as face impressions are formed rapidly and spontaneously, similar effects should emerge on indirect assessments of such impressions.

The next two image-assessment experiments examined rapid and spontaneous impressions of likability and trustworthiness using a variant of the stereotype misperception task (SMT; Krieglmeyer &

Sherman, 2012), a well-validated sequential priming task that assesses the biasing effect of semantic content (e.g., trait impressions) evoked by prime faces on judgments of unrelated target faces. We tested whether the individual ingroup and outgroup CIs generated by perspective takers and control participants (which served as prime faces in the SMT) evoke differential likability (Experiment 3A) and trustworthiness (Experiment 3B).

Method

Participants. Undergraduates (Experiment 3A: N = 165; Experiment 3B: N = 250) participated for course credit. We excluded data from participants (n = 2) who pressed the same key on every trial in the SMT in Experiment 3B. The final samples comprised 165 participants in Experiment 3A (122 women, 43 men)[6] and 248 participants in Experiment 3B (178 women, 70 men; 124 Asian, 50 Latinx, 40 White, 7 Black, 22 reporting other or multiple races/ethnicities, 4 unreported; M_{age} = 20.37, SD_{age} = 2.35), which afforded >80% power to detect two-way interactions with small-to-medium effect sizes (Experiment 3A: η_p^2 = .047; Experiment 3B: η_p^2 = .032; Faul et al., 2007).

Procedure. Participants in both image-assessment experiments completed a variant of the SMT (Krieglmeyer & Sherman, 2012) with the following properties. On each trial, a prime face and a target face appeared in quick succession. The prime faces were the gender-matched individual CIs. In addition to the CI primes, we also included neutral primes consisting of a facial outline, as is customary in the SMT. For ease of presentation, however, we do not include the neutral primes in the SMT analyses reported below.[7] The target faces were degraded versions of computer-generated facial morphs (Oosterhof & Todorov, 2008). There were 24 unique facial identities that were morphed to be high or low in likability (Experiment 3A) or high or low in trustworthiness (Experiment 3B).

After completing two short practice blocks (5 trials and 10 trials, respectively), participants completed two testing blocks (120 trials each). Each trial began with a fixation cross (500 ms), followed by

[6] Because of a programming error, participant age and race/ethnicity were not recorded in Experiment 3A.
[7] A typical finding in research using the SMT is that the neutral primes are not evaluatively neutral (Rees et al., 2019; Rivers et al., 2020), and thus they do not easily serve as a baseline condition. This was also the case in our data. In Experiment 3A, the neutral primes (M = .47, SD = .24) spontaneously evoked less likability than did the CI primes generated in each of the experimental conditions (ts > 5.57, ps < .001). Similarly, in Experiment 3B, the neutral primes (M = .42, SD = .24) spontaneously evoked less trustworthiness than did the CI primes in each of the experimental conditions (ts > 9.72, ps < .001). Thus, overall, the neutral primes were spontaneously evaluated as less likable (Experiment 3A) and less trustworthy (Experiment 3B) than were all the CI primes.

a prime face (i.e., an individual CI; 150 ms), a blank screen (50 ms), and then a target face (100 ms). Finally, a gray pattern mask appeared until participants judged the target face as "more likable" or "less likable" (Experiment 3A) or as "more trustworthy" or "less trustworthy" (Experiment 3B). Trials appeared in random order, and prime CIs of each type appeared equally often with target faces of each type.

Instructions asked participants to attend to and remember the prime faces for a later task, but to avoid letting these prime faces affect their judgments of the target faces. Instructions also urged participants to respond quickly and to use their "gut" feelings when judging the target faces. The proportion of "more likable" or "more trustworthy" judgments after each prime face served as an indirect index of the likability/trustworthiness of those faces. Specifically, spontaneous prime impressions are revealed by their unintended influence on judgments of target likability/trustworthiness (Uleman, 1999).

Results

Spontaneous Likability (Experiment 3A). A 2 (Group Membership Prime) × 2 (Perspective Taking Prime) repeated-measures ANOVA on the proportion of individual CIs in each condition that prompted a "more likable" response revealed that the ingroup CI primes (M = 0.61, SD = 0.19) elicited more likability than did the outgroup CI primes (M = 0.57, SD = 0.20), $F(1, 164) = 36.35$, $p < .001$, $\eta_p^2 = 0.18$, $CI_{90\%}$ [0.10, 0.27].[8] The predicted interaction was also significant, $F(1, 164) = 18.54$, $p < .001$, $\eta_p^2 = 0.10$, $CI_{90\%}$ [0.04, 0.18] (see Fig. 4).

[8] Preliminary analyses that included Target Face (high likability/trustworthiness vs. low likability/trustworthiness) as a factor indicated that this variable did not moderate the key findings in either Experiment 3A or 3B, as is typical in research using the SMT (Krieglmeyer & Sherman, 2012).

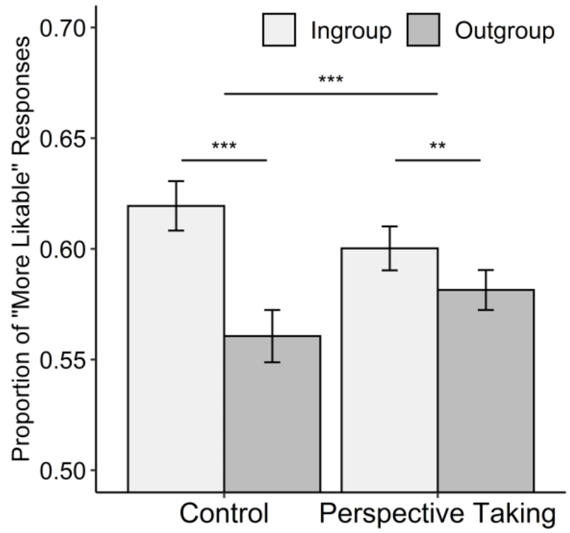

Figure 4. Proportion of “more likable” responses spontaneously elicited by individual classification images by group membership and perspective taking in Experiment 3A. Error bars reflect 95% confidence intervals. $^{***}p < .001$, $^{**}p < .01$.

Simple-effects tests indicated that the ingroup CIs elicited more likability than did the outgroup CIs in both conditions, but this group difference was weaker for the CIs generated by perspective takers, $t(164) = 2.75$, $p = .007$, $d = 0.21$, $CI_{95\%}$ [0.06, 0.37], than for the CIs generated by control participants, $t(164) = 6.61$, $p < .001$, $d = 0.51$, $CI_{95\%}$ [0.35, 0.68]. Approaching the interaction differently, perspective takers’ outgroup CIs ($M = 0.58$, $SD = 0.19$) elicited more likability than did control participants’ outgroup CIs ($M = 0.56$, $SD = 0.20$), $t(164) = -3.16$, $p = .002$, $d = -0.25$, $CI_{95\%}$ [-0.40, -0.09,], whereas perspective takers’ ingroup CIs ($M = 0.60$, $SD = 0.19$) elicited less likability than did control participants’ ingroup CIs ($M = 0.62$, $SD = 0.18$), $t(164) = 2.88$, $p = .004$, $d = 0.22$, $CI_{95\%}$ [0.07, 0.38].

Spontaneous Trustworthiness (Experiment 3B). An identical 2 × 2 ANOVA on the proportion of individual CIs in each condition that prompted a “more trustworthy” response revealed that the ingroup CI primes ($M = 0.60$, $SD = 0.18$) elicited more trustworthiness than did the outgroup CI primes ($M = 0.56$, $SD = 0.19$), $F(1, 247) = 49.30$, $p < .001$, $\eta_p^2 = 0.17$, $CI_{90\%}$ [0.10, 0.24]. Once again, the predicted interaction was also significant, $F(1, 247) = 10.70$, $p = .001$, $\eta_p^2 = 0.04$, $CI_{90\%}$ [0.01, 0.09] (see Fig. 5).

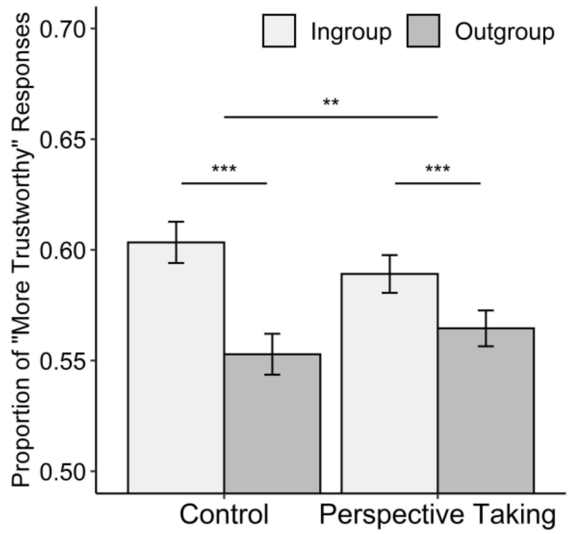

Figure 5. Proportion of "more trustworthy" responses spontaneously elicited by individual classification images by group membership and perspective taking in Experiment 3B. Error bars reflect 95% confidence intervals. $^{***}p < .001$, $^{**}p < .01$.

Simple-effects tests revealed that the ingroup CIs elicited more trustworthiness than did the outgroup CIs in both conditions, but again this group difference was weaker for the CIs generated by perspective takers, $t(274) = 4.00$, $p < .001$, $d = 0.25$, $CI_{95\%}$ [0.13, 0.38], than for the CIs generated by control participants, $t(247) = 7.08$, $p < .001$, $d = 0.45$, $CI_{95\%}$ [0.32, 0.58]. Approaching the interaction differently, perspective takers' outgroup CIs ($M = 0.56$, $SD = 0.19$) elicited more trustworthiness than did control participants' outgroup CIs ($M = 0.55$, $SD = 0.19$), $t(247) = -2.14$, $p = .033$, $d = -0.14$, $CI_{95\%}$ [-0.26, -0.01], whereas perspective takers' ingroup CIs ($M = 0.59$, $SD = 0.18$) elicited less trustworthiness than did control participants' ingroup CIs ($M = 0.60$, $SD = 0.18$), $t(247) = 2.48$, $p = .014$, $d = 0.16$, $CI_{95\%}$ [0.03, 0.28].

Discussion

These results align with and extend those reported in the previous image-assessment experiments by documenting that the effects of perspective taking extend to spontaneously evoked impressions of intergroup visual representations. The ingroup visual representations evoked more likability and trustworthiness than did the outgroup visual representations, and engaging in perspective taking dampened this pattern of intergroup bias. Specifically, perspective takers' outgroup representations evoked more likability and trustworthiness, but their ingroup representations evoked less likeability and trustworthiness, relative to the respective visual representations generated in the control condition.

Experiment 4

Experiment 4 examined whether the effects of perspective taking on intergroup bias in visual representations observed thus far have downstream consequences for overt behavior in an economic trust game (Berg et al., 1995; Chang et al., 2010; van ’t Wout & Sanfey, 2008). Participants completed a series of modified single-shot trust games with interaction partners represented by the individual CIs. Prior work indicates that people trust ingroup members more than outgroup members in such trust games (Evans & Krueger, 2009), and that this intergroup bias in trust behavior extends to visual representations of minimal group members (Ratner et al., 2014). Accordingly, we expected that participants would trust partners represented by ingroup CIs more than they would trust partners represented by outgroup CIs, and that perspective taking would attenuate this intergroup bias in trust behavior.

Method

Participants. MTurk workers (*N* = 226) participated for modest pay ($0.70). All participants also received a bonus ($0.30) to use in the trust game. We excluded data from participants (*n* = 49) who answered more than one of the three trust game comprehension questions incorrectly.[9] The final sample comprised 177 participants (99 women, 78 men; 124 White, 20 Black, 14 Latinx, 10 Asian, 9 reporting other or multiple races/ethnicities; M_{age} = 36.71, SD_{age} = 11.39), which afforded 80% power to detect a two-way interaction with a small-to-medium effect size (η_p^2 = 0.044; Faul et al., 2007).

Procedure. Participants completed a series of modified single-shot trust games (Berg et al., 1995) with 95 distinct interaction partners represented by the gender-matched individual CIs. The design and instructions for the trust game were adapted from materials used in prior work (Everett et al., 2018). Participants learned that they would play multiple rounds of a decision-making game with numerous different interaction partners, described as MTurk workers who had purportedly provided responses in a previous session. Importantly, participants were told that they would receive a bonus payment based on the responses of one randomly selected interaction partner. In actuality, all participants received a $0.30 bonus after the experiment. Participants learned that both players were given a $0.30 bonus to use during the game, and they could transfer any portion of the bonus (from $0.00 to $0.30) to their partner. Any

[9] Retaining these participants’ data did not change the statistical inferences drawn from the results.

portion of the bonus shared would be doubled and sent to the partner, after which the partner could transfer any portion of their total bonus back to the participant.

We informed participants that their partners had already indicated in a previous session, for every possible amount of money transferred to them, how much money they wanted to return to the participant. Before playing the game, participants completed three comprehension questions. If participants cooperate and transfer money to their partner, then both players can earn additional money. Importantly, cooperating also leaves participants vulnerable, in that partners could keep all the money for themselves. Accordingly, the amount of money transferred to the partner captures the extent to which participants trust their partner.

On each round of the trust game, participants saw a photo of their partner, who was represented by a gender-matched individual CI from one of the four conditions based on group membership and perspective taking. We informed participants that the photos had been visually edited for identity protection. Participants completed five practice trials to gain familiarity with the task. As in the previous image-assessment experiments, because there were many more female (n = 159) than male (n = 95) individual CIs, female participants played a trust game with 95 randomly selected CIs, whereas male participants played the game with all 95 male CIs.

Results

A 2 (Group Membership) × 2 (Perspective Taking) repeated-measures ANOVA on the average amount of money transferred to the individual CIs in each condition revealed that more money was entrusted to interaction partners represented by the ingroup CIs (M = 19.87 cents, SD = 9.07) than to partners represented by the outgroup CIs (M = 19.23 cents, SD = 9.31), $F(1, 176) = 25.77$, $p < .001$, $\eta_p^2 = 0.13$, $CI_{90\%}$ [0.06, 0.20]. The predicted interaction was also significant, $F(1, 176) = 14.47$, $p < .001$, $\eta_p^2 = 0.08$, $CI_{90\%}$ [0.02, 0.14] (see Fig. 6).

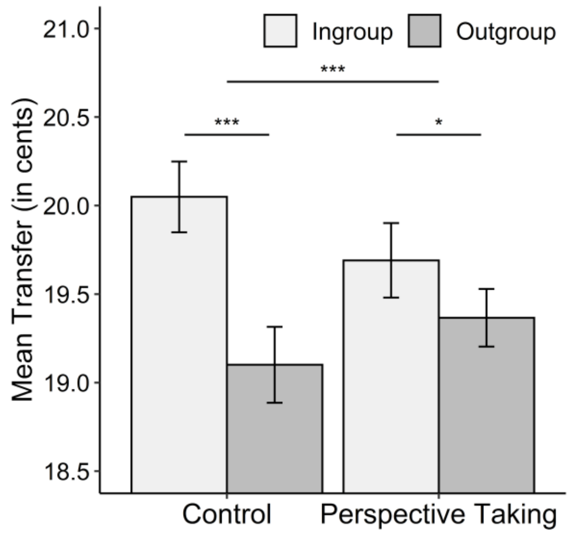

Figure 6. Mean transfers (in cents) to interaction partners (i.e., individual classification images) by group membership and perspective taking in Experiment 4. Error bars depict 95% confidence intervals. ***p < .001, *p < .05.

Simple-effects tests indicated that, although partners represented by ingroup CIs were entrusted with more money than were partners represented by outgroup CIs in both conditions, this pattern of intergroup bias was weaker for CIs generated by perspective takers, $t(176) = 2.32$, $p = .021$, $d = 0.17$, $CI_{95\%}$ [0.03, 0.32], than for CIs generated by control participants, $t(176) = 5.95$, $p < .001$, $d = 0.45$, $CI_{95\%}$ [0.29, 0.60]. Approaching the interaction differently, perspective takers' outgroup CIs (M = 19.37 cents, SD = 9.29) received more money than did control participants' outgroup CIs (M = 19.10 cents, SD = 9.36), $t(176) = -2.40$, $p = .018$, $d = -0.18$, $CI_{95\%}$ [-0.33, -0.03], whereas perspective takers' ingroup CIs (M = 19.69 cents, SD = 9.18) received less money than did control participants' ingroup CIs (M = 20.05 cents, SD = 8.98), $t(176) = 2.77$, $p = .006$, $d = 0.21$, $CI_{95\%}$ = [0.06, 0.36].

Discussion

These results dovetail with those of the previous image-assessment experiments by showing that the effects of perspective taking on intergroup bias in visual representations extend to cooperative behavior in an economic trust game. Although participants transferred more money to interaction partners represented by ingroup CIs than to partners represented by outgroup CIs, perspective taking attenuated this intergroup bias. In line with the previous image-assessment experiments, participants transferred more money to partners represented by perspective takers' outgroup CIs, but less money to partners represented by perspective takers' ingroup CIs, relative to control participants' respective CIs.

Experiment 5

In the image-assessment experiments reported thus far, perspective taking consistently reduced, but never eliminated, intergroup bias in visual representations; these results emerged in both deliberate and spontaneous impressions of and cooperative behavior with those visual representations. Next, we aimed to localize the critical facial region(s) underlying these effects. Facial trustworthiness has been linked to subtle expressions of approach-related emotions (i.e., happiness) and avoidance-related emotions (i.e., anger) in the eye and mouth regions (Oosterhof & Todorov, 2008, 2009). Intergroup differences in representations of both face regions have also been reported in prior research (Paulus et al., 2016; Ratner et al., 2014). To gauge the importance of these face regions in the visual representations under investigation here, we tested whether and how occluding the eyes, the mouth, or both features in the individual CIs would affect trustworthiness impressions (see Rule et al., 2008, for a similar empirical approach).

Method

Participants. MTurk workers (N = 401, 204 women, 197 men; 262 White, 77 Black, 28 Asian, 26 Latinx, 8 reporting other or multiple races/ethnicities; M_{age} = 36.08, SD_{age} = 10.68) participated for modest pay ($0.70). This sample afforded >80% power to detect a three-way interaction with a small-to-medium effect size (η_p^2 = 0.020; Faul et al., 2007).

Procedure. We modified each of the individual CIs to occlude either the eyes (84 × 248 pixels), the mouth (66 × 188 pixels), or both regions, resulting in three different versions of the CIs (eyes-occluded CIs, mouth-occluded CIs, or eyes-and-mouth-occluded CIs; see Fig. 7).[10] Participants were randomly assigned to one of the three occlusion conditions. Instructions urged participants to use their “gut-level” impressions to judge the faces as “more trustworthy” or “less trustworthy” than the average face in the task by pressing one of two response keys. They completed 10 practice trials to gain familiarity with the average trustworthiness of the faces. As before, because there were many more female CIs (n =

[10] We created occlusion boxes in MATLAB using the *createMask* function, and we saved the output as a 512 × 512 matrix of logical values.

159) than male CIs (n = 95), female participants rated 95 randomly selected female CIs from the larger set, whereas male participants rated all 95 male CIs.[11]

Eyes-occluded **Mouth-occluded** **Eyes-and-mouth-occluded**

Figure 7. Examples of occluded classification images in Experiment 5. Each occlusion box was the same size and placed in the same location for every modified individual classification image.

Results

A 3 (Occlusion) × 2 (Group Membership) × 2 (Perspective Taking) mixed ANOVA, with repeated measures on the last two factors, on the proportion of individual CIs in each condition that elicited a “more trustworthy” response revealed main effects of Group Membership, $F(1, 398) = 86.54$, $p < .001$, $\eta_p^2 = 0.18$, $CI_{90\%}$ [0.13, 0.23], and Occlusion, $F(2, 398) = 5.82$, $p = .003$, $\eta_p^2 = 0.03$, $CI_{90\%}$ [0.01, 0.06]. The ingroup CIs ($M = 0.39$, $SD = 0.27$) were judged as more trustworthy than were the outgroup CIs ($M = 0.34$, $SD = 0.25$). Additionally, the CIs with both the eyes and mouth occluded ($M = 0.30$, $SD = 0.27$) were judged as less trustworthy than were the CIs with just the eyes occluded ($M = 0.40$, $SD = 0.25$), $t(1051.60) = -5.86$, $p < .001$, $d = -0.36$, $CI_{95\%}$ [-0.48, -0.24], and the CIs with just the mouth occluded ($M = 0.39$, $SD = 0.26$), $t(1044.65) = -5.09$, $p < .001$, $d = -0.31$, $CI_{95\%}$ [-0.44, -0.19]. Trustworthiness impressions for the eyes-occluded CIs and the mouth-occluded CIs did not significantly differ, $t(1072) = 0.67$, $p = .505$, $d = 0.04$.

As in the previous image-assessment experiments, the Group Membership × Perspective Taking interaction was significant, $F(1, 398) = 27.94$, $p < .001$, $\eta_p^2 = 0.07$, $CI_{90\%}$ [0.03, 0.11]. Simple-effects tests indicated that the ingroup CIs were judged as more trustworthy than were the outgroup CIs, but once again this pattern of intergroup bias was weaker for the CIs generated by perspective takers, $t(400) = 4.24$, $p < .001$, $d = 0.21$, $CI_{95\%}$ [0.11, 0.31], than for the CIs generated by control participants, $t(400) =$

[11] Because of a programming error, 3 CIs were not rated (1 in the eyes-occluded condition, 2 in the eyes-and-mouth-occluded condition).

9.45, $p < .001$, $d = 0.47$, $CI_{95\%}$ [0.37, 0.58]. Approaching the interaction differently, perspective takers' outgroup CIs ($M = 0.35$, $SD = 0.26$) were judged as more trustworthy than were control participants' outgroup CIs ($M = 0.32$, $SD = 0.25$), $t(400) = -4.78$, $p < .001$, $d = -0.24$, $CI_{95\%}$ [-0.34, -0.14], whereas perspective takers' ingroup CIs ($M = 0.38$, $SD = 0.27$) were judged as less trustworthy than were control participants' ingroup CIs ($M = 0.40$, $SD = 0.27$), $t(400) = 2.69$, $p = .007$, $d = 0.13$, $CI_{95\%}$ [0.04, 0.23].

The Occlusion × Group Membership interaction was also significant, $F(2, 398) = 22.98$, $p < .001$, $\eta_p^2 = 0.10$, $CI_{90\%}$ [0.06, 0.15]. Simple-effects tests indicated that the ingroup CIs were judged as more trustworthy than were the outgroup CIs in the eyes-occluded condition, $t(275) = 9.25$, $p < .001$, $d = 0.56$, $CI_{95\%}$ [0.43, 0.68], and the mouth-occluded condition, $t(263) = 5.33$, $p < .001$, $d = 0.33$, $CI_{95\%}$ [0.20, 0.45]; however, this pattern of intergroup bias was not significant in the eyes-and-mouth-occluded condition, $t(261) = 1.31$, $p = .190$, $d = 0.08$.

Finally, and of greatest importance, the Occlusion × Group Membership × Perspective Taking interaction was significant, $F(2, 398) = 17.30$, $p < .001$, $\eta_p^2 = 0.08$, $CI_{90\%}$ [0.04, 0.12]. To further specify this interaction, we conducted separate 2 (Group Membership) × 2 (Perspective Taking) ANOVAs in each occlusion condition (see Fig. 8). When just the eyes were occluded, there were main effects of Group Membership, $F(1, 137) = 69.92$, $p < .001$, $\eta_p^2 = 0.34$, $CI_{90\%}$ [0.23, 0.43], and Perspective Taking, $F(1, 137) = 5.05$, $p = .026$, $\eta_p^2 = 0.04$, $CI_{90\%}$ [0.002, 0.10], and a Group Membership × Perspective Taking interaction, $F(1, 137) = 53.57$, $p < .001$, $\eta_p^2 = 0.28$, $CI_{90\%}$ [0.18, 0.38]. The pattern of responses underlying this interaction mirrored that found in the previous image-assessment experiments.

When just the mouth was occluded, the Group Membership main effect was significant, $F(1, 131) = 25.68$, $p < .001$, $\eta_p^2 = 0.16$, $CI_{90\%}$ [0.08, 0.26], but it was not moderated by Perspective Taking, $F(1, 131) = 1.69$, $p = .197$, $\eta_p^2 = 0.01$. When both the eyes and mouth were occluded, however, neither main effect nor the interaction was significant. Together, these results suggest that perspective taking appears to have uniquely altered the trustworthiness with which the mouths of ingroup and outgroup members were visually represented.

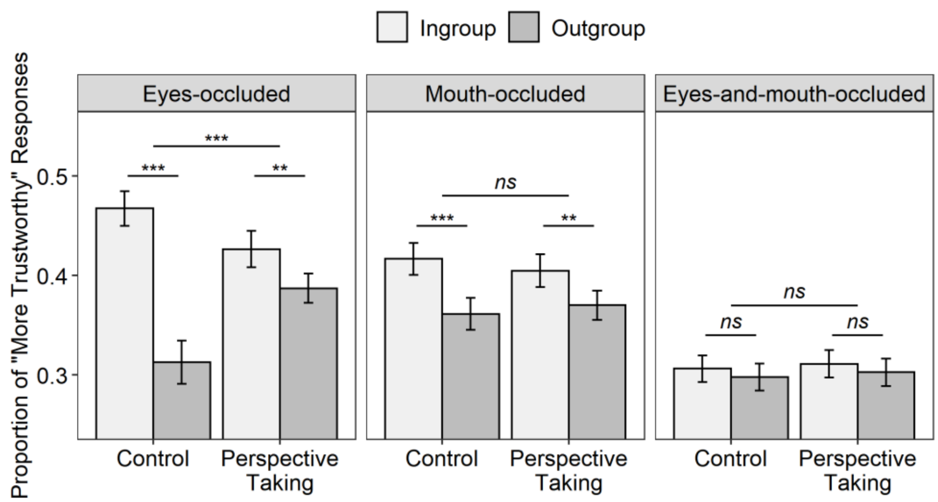

Figure 8. Proportion of "more trustworthy" responses explicitly ascribed to individual classification images by occlusion, group membership, and perspective taking in Experiment 5. Error bars depict 95% confidence intervals. $^{***}p < .001$, $^{**}p < .01$, *ns* = non-significant.

Discussion

These results help clarify findings from the previous image-assessment experiments by identifying a particular facial region—the mouth—wherein perspective taking alters intergroup visual representations. Specifically, when just the eyes were occluded, the interactive effects of group membership and perspective taking on the trustworthiness of the visual representations continued to emerge. In conditions where just the mouth was occluded or where both the mouth and the eyes were occluded, however, the positive effects of perspective taking on intergroup trustworthiness disappeared.

Experiment 6

The previous image-assessment experiment implicated the mouth as the critical region altered by perspective taking, but it remains unclear how the mouths differed in the visual representations generated by perspective takers and control participants. Impressions of facial trustworthiness are closely connected to subtle expressions of happiness and anger (Oosterhof & Todorov, 2008, 2009); thus, perspective taking may have altered the extent to which image generators visualized smiling or frowning/scowling on the group members' faces. In our final image-assessment experiment, we examined the interactive impact of group membership and perspective taking on judgments of smiling and frowning/scowling in a different set of subgroup CIs than what we used in Experiment 2A. We predicted that the ingroup CIs would be judged as smiling more than the outgroup CIs, but that this group difference would be tempered in perspective takers' CIs. In contrast, the outgroup CIs should be judged as scowling/frowning more than the ingroup CIs, but again this difference should be weaker in perspective takers' CIs. We further explored whether smiling or frowning/scowling expressions mediate the effect of perspective taking on the trustworthiness of the ingroup and outgroup CIs.

Method

Participants. MTurk workers (N = 151) participated for modest pay ($0.75). Data were excluded from participants who reported not taking the experiment seriously (n = 6) and who did not finish the entire experiment (n = 1). The final sample comprised 144 participants (67 women, 73 men, 4 non-binary; 112 White, 12 Black, 9 Asian, 8 Latinx, 1 Native American, 2 reporting other or multiple races/ethnicities; M_{age}= 39.00, SD_{age}= 12.25), which afforded >80% power to detect a two-way interaction with a small-to-medium effect size (η_p^2 = 0 .053; Faul et al., 2007).

Procedure. We first created a new set of subgroup CIs from new random subsets of individual CIs using the same procedure as before. Because there were more female than male participants in the image-generation experiment, we generated more female (n = 24) than male subgroup CIs (n = 16). Each subgroup CI comprised approximately 6 individual CIs.

Participants rated the subgroup CIs, which appeared individually in a randomized order, on smiling, frowning/scowling, and trustworthiness. Male participants rated the male subgroup CIs, whereas female and gender non-binary participants rated the female subgroup CIs. Participants were encouraged to rely on their first impressions and to "go with their gut" when rating the faces. First, participants rated (in counterbalanced order) the extent to which each face appeared to be smiling and frowning/scowling (1 = *not at all*, 7 = *very much so*). Next, they rated each face on trustworthiness (1 = *very untrustworthy*, 7 = *very trustworthy*).

Results

Smiling. A 2 (Group Membership) × 2 (Perspective Taking) repeated-measures ANOVA on the smiling ratings revealed a significant Group Membership main effect, $F(1, 143) = 429.63$, $p < .001$, η_p^2 = 0.75, $CI_{90\%}$ [0.70, 0.79], whereby the ingroup CIs (M = 4.12, SD = 1.06) were rated as smiling more than were the outgroup CIs (M = 2.49, SD = 1.05). A significant Perspective Taking main effect, $F(1, 143)$ = 12.98, $p < .001$, η_p^2 = 0.08, $CI_{90\%}$ [0.03, 0.16], indicated that the CIs generated by perspective takers (M =

3.37, SD = 1.24) were rated as smiling more than were the CIs generated by control participants (M = 3.24, SD = 1.42). The predicted Group Membership × Perspective Taking interaction was also significant, $F(1, 143) = 93.20$, $p < .001$, $\eta_p^2 = 0.39$, $CI_{90\%}$ [0.30, 0.48] (see Fig. 9, left panel).

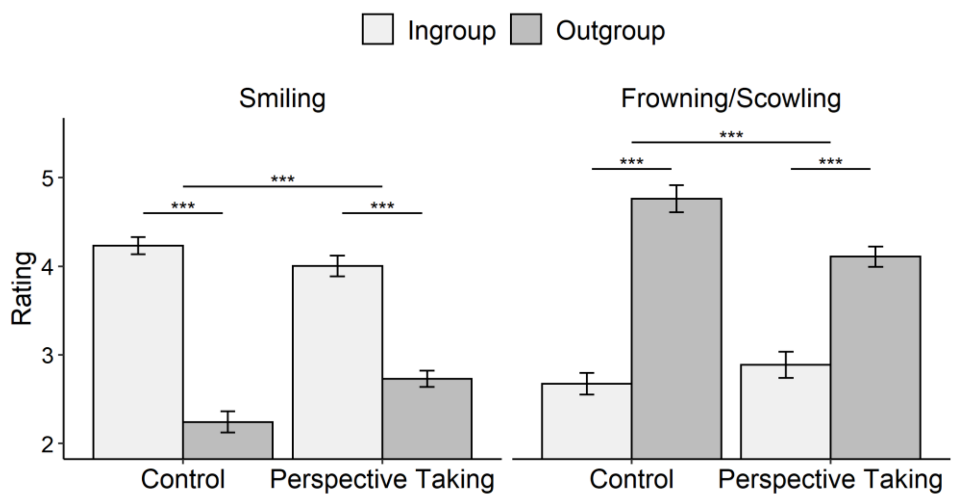

Figure 9. Smiling and frowning/scowling ratings of subgroup classification images by group membership and perspective taking in Experiment 6. Error bars reflect 95% confidence intervals. ***$p < .001$.

Simple-effects tests indicated that, although there was robust intergroup bias in smiling ratings in both conditions, it was smaller for the CIs generated by perspective takers, $t(143) = 14.90$, $p < .001$, d = 1.24, $CI_{95\%}$ [1.03, 1.46], than for the CIs generated by control participants, $t(143) = 22.43$, $p < .001$, d = 1.87, $CI_{95\%}$ [1.60, 2.15]. Exploring this interaction differently, perspective takers' outgroup CIs (M = 2.73, SD = 0.99) were rated as smiling more than were control participants' outgroup CIs (M = 2.24, SD = 1.05), $t(143) = -9.74$, $p < .001$, d = -0.81, $CI_{95\%}$ [-1.00, -0.62], whereas perspective takers' ingroup CIs (M = 4.00, SD = 1.14) were rated as smiling less than were control participants' ingroup CIs (M = 4.23, SD = 0.96), $t(143) = 4.40$, $p < .001$, d = 0.37, $CI_{95\%}$ [0.20, 0.54].

Frowning/Scowling. An identical 2 × 2 ANOVA on the frowning/scowling ratings revealed a significant Group Membership main effect, $F(1, 143) = 268.43$, $p < .001$, $\eta_p^2 = 0.65$, $CI_{90\%}$ [0.58, 0.71], whereby the outgroup CIs (M = 4.43, SD = 1.11) were rated as frowning/scowling more than were the ingroup CIs (M = 2.78, SD = 1.20). A significant Perspective Taking main effect, $F(1, 143) = 25.99$, $p <$ .001, $\eta_p^2 = 0.15$, $CI_{90\%}$ [0.07, 0.24], indicated that the CIs generated by perspective takers (M = 3.50, SD = 1.30) were rated as frowning/scowling less than were the CIs generated by control participants (M = 3.72, *SD* = 1.53). The Group Membership × Perspective Taking interaction was also significant, $F(1, 143) = 98.92$, $p < .001$, $\eta_p^2 = 0.41$, $CI_{90\%}$ [0.31, 0.49] (see Fig. 9, right panel).

Simple-effects tests indicated that, although intergroup bias in frowning/scowling ratings emerged in both conditions, it was smaller for the CIs generated by perspective takers, $t(143) = -11.40$, $p < .001$, $d = -0.95$, $CI_{95\%}$ [-1.15, -0.75], than for the CIs generated by control participants, $t(143) = -18.53$, $p < .001$, $d = -1.54$, $CI_{95\%}$ [-1.79, -1.31]. Examining the interaction differently, perspective takers' outgroup CIs (*M* = 4.11, *SD* = 1.05) were rated as frowning/scowling less than were control participants' outgroup CIs (*M* = 4.76, *SD* = 1.07), $t(143) = 10.66$, $p < .001$, $d = 0.89$, $CI_{95\%}$ [0.70, 1.08], whereas perspective takers' ingroup CIs (*M* = 2.89, *SD* = 1.23) were rated as frowning/scowling more than were control participants' ingroup CIs (*M* = 2.67, *SD* = 1.17), $t(143) = -3.45$, $p < .001$, $d = -0.29$, $CI_{95\%}$ [-0.46, -0.12].

Trustworthiness. An identical 2 × 2 ANOVA on the trustworthiness impressions yielded a significant Group Membership main effect, $F(1, 143) = 245.38$, $p < .001$, $\eta_p^2 = 0.63$, $CI_{90\%}$ [0.56, 0.69], whereby the ingroup CIs (*M* = 4.61, *SD* = 0.86) were rated as more trustworthy than were the outgroup CIs (*M* = 3.26, *SD* = 1.06). A significant Perspective Taking main effect, $F(1, 143) = 33.56$, $p < .001$, $\eta_p^2 = 0.19$, $CI_{90\%}$ [0.10, 0.28], indicated that perspective takers' CIs (*M* = 4.04, *SD* = 1.07) were rated as more trustworthy than were control participants' CIs (*M* = 3.83, *SD* = 1.27). Once again, the Group Membership × Perspective Taking interaction was also significant, $F(1, 143) = 68.98$, $p < .001$, $\eta_p^2 = 0.33$, $CI_{90\%}$ [0.23, 0.42] (see Fig. 10).

Simple-effects tests indicated that, although intergroup bias was evident in both conditions, it was smaller for the CIs generated by perspective takers, $t(143) = 11.48$, $p < .001$, $d = 0.96$, $CI_{95\%}$ [0.76, 1.16], than for the CIs generated by control participants, $t(143) = 16.99$, $p < .001$, $d = 1.42$, $CI_{95\%}$ [1.19, 1.65]. Examining the interaction differently, perspective takers' outgroup CIs (*M* = 3.52, *SD* = 0.97) were rated as more trustworthy than were control participants' outgroup CIs (*M* = 3.00, *SD* = 1.08), $t(143) = -10.71$, $p < .001$, $d = -0.89$, $CI_{95\%}$ [-1.09, -0.70]. Trustworthiness ratings for the ingroup CIs generated by perspective takers (*M* = 4.56, *SD* = 0.89) and control participants (*M* = 4.67, *SD* = 0.84), by contrast, did not significantly differ, $t(143) = 1.92$, $p = .056$, $d = 0.16$.

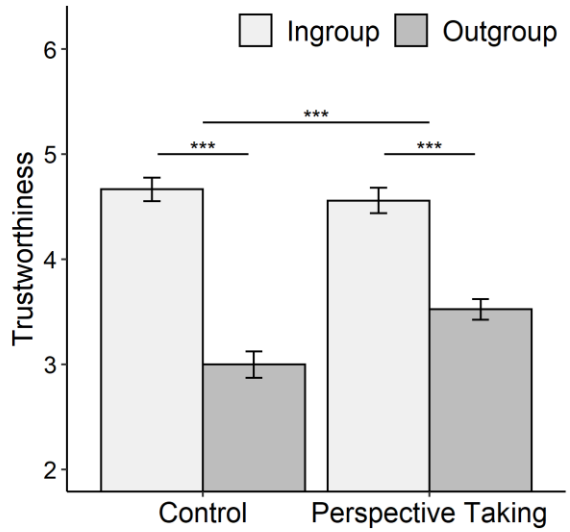

Figure 10. Trustworthiness explicitly ascribed to the subgroup classification images by group membership and perspective taking in Experiment 6. Error bars reflect 95% confidence intervals. $^{***}p < .001$.

Moderated Mediation with Smiling. Next, we tested a moderated mediation model wherein smiling ratings mediated the interactive effect of group membership and perspective taking on the trustworthiness impressions of the CIs. The primary predictor was Group Membership (ingroup = .5, outgroup = -.5), the moderator was Perspective Taking (control = .5, perspective taking = -.5), and the mediator was smiling ratings. Using the *JSmediate* 0.1.1 package, we followed Yzerbyt et al.'s (2018) component (i.e., joint-significance testing) approach, which requires the joint significance of individual parameter estimates of indirect effects to establish the presence of mediation.[12]

The Group Membership × Perspective Taking interaction predicted the mediator (i.e., smiling), a_{mod} = 0.72, $t(572) = 4.16$, $p < .001$, and smiling significantly predicted trustworthiness impressions, b = 0.44, $t(570) = 13.11$, $p < .001$, suggesting the presence of first-stage moderated mediation. The Group Membership × Perspective Taking interaction did not significantly predict trustworthiness impressions when controlling for smiling, c'_{mod} = 0.23, $t(570) = 1.30$, $p = .193$. Further supporting this conclusion, the Monte Carlo confidence interval (derived from 10,000 iterations) for the indirect effect did not contain zero, ab = 0.32, $CI_{95\%}$ [0.17, 0.48].

[12] One advantage of the component approach over the frequently used “index” approach (Hayes, 2013) is better protection against Type I error inflation.

Moderated Mediation with Frowning/Scowling. Finally, we tested a moderated mediation model wherein frowning/scowling ratings mediated the interactive effect of group membership and perspective taking on the trustworthiness impressions of the CIs. We again used the component (i.e., joint-significance testing) approach (Yzerbyt et al., 2018).

The Group Membership × Perspective Taking interaction predicted the mediator (i.e., frowning/scowling), $a_{mod} = -0.87$, $t(572) = 4.59$, $p < .001$, but frowning/scowling did not significantly predict trustworthiness impressions, $b = -0.03$, $t(570) = 0.78$, $p = .438$, suggesting an absence of first-stage moderated mediation. Further supporting this conclusion, the Monte Carlo confidence interval for the indirect effect contained zero, $ab = 0.02$, $CI_{95\%}$ [-0.04, 0.08].

Discussion

These results indicate that perspective taking changed the extent to which group members' faces were visualized as smiling and frowning/scowling. As predicted, ingroup faces were represented as smiling more than were outgroup faces, but this group difference was weaker for faces generated by perspective takers. Similarly, outgroup faces were represented as frowning/scowling more than were ingroup faces, but this difference was also weaker for faces generated by perspective takers. Adopting an outgroup member's perspective led participants to visualize that person's face as smiling more and frowning/scowling less, whereas adopting an ingroup member's perspective led participants to visualize that person's face as smiling less and scowling/frowning more, relative to the control condition. Furthermore, evidence of moderated mediation emerged, whereby changes in smiling, but not frowning/scowling, helped explain the interactive effect of group membership and perspective taking on trustworthiness impressions.

General Discussion

The current research investigated whether and how perspective taking alters intergroup biases in visual representations of group members' faces. Using a combination of group, subgroup, and individual classification images (CIs) generated in a reverse-correlation paradigm, and measures of both deliberate and spontaneous impressions and cooperative behavior, we found that faces of minimal ingroup members were visually represented more favorably than were faces of minimal outgroup members (Ratner et al., 2014). More important for the current work, perspective taking reduced, but did not eliminate, these patterns of intergroup bias. Relative to participants in a control condition, perspective takers generated more likable and trustworthy outgroup visual representations, which aligns with findings from numerous other studies indicating that perspective taking can increase the favorability of conceptual representations of outgroup members (Ku et al., 2015; Todd & Galinsky, 2014). With several exceptions, furthermore, the ingroup visual representations generated by perspective takers were less favorable than were those generated by control participants. This latter finding converges with mounting evidence that perspective taking does *not* produce uniformly positive effects on judgment and behavior (Sassenrath et al., 2016; Vorauer, 2013).

Our final two image-assessment experiments revealed that perspective taking altered visualizations of group members' mouths, specifically the extent to which the target person appeared to be smiling and frowning/scowling. When the CIs' eyes were occluded (but their mouths were visible), perspective taking reduced intergroup bias in trustworthiness impressions. When the CIs' mouths were occluded (but their eyes were visible), by contrast, there was robust intergroup bias in trustworthiness impressions, which aligns with prior work on the importance of the eyes in intergroup cognition (Kawakami et al., 2017; Xiao et al., 2016). Importantly, however, perspective taking did not temper intergroup bias in this condition. Finally, when the CIs' eyes and mouths were both occluded, no effects involving group membership or perspective taking emerged. The mouth, therefore, appears to be a critical facial region underlying the interactive effect of group membership and perspective taking on target impressions. Perspective takers also visualized outgroup (ingroup) faces as smiling more (less) and frowning/scowling less (more) than did control participants, with changes in the extent of smiling (but not in the extent of frowning/scowling) mediating the interactive impact of group membership and perspective taking on trustworthiness impressions. Future research using eye-tracking could help further corroborate the importance of the mouth by exploring whether differential gaze durations on this facial region predict trustworthiness impressions.

Together, these findings extend the perspective-taking literature into important new theoretical territory: Not only does perspective taking affect conceptual representations of what others are like; it also seems to shape visual representations of what others *look* like. Insofar as these visual representations guide early aspects of intergroup cognition (Brinkman et al., 2017), perspective taking may have important top-down influences on the visual processing of ingroup and outgroup faces. Findings from the final two image-assessment experiments, in particular, suggest the intriguing possibility that perspective taking might attenuate the tendency to more readily detect positive emotions on ingroup members' faces and negative emotions on outgroup members' faces (Hugenberg, 2005; Hugenberg & Bodenhausen, 2003; Lazerus et al., 2016). Future research should explore this possibility.

Insofar as engaging in perspective taking instills a general orientation whereby *any* target person is represented and processed as more "self-like" than they would be in the absence of perspective taking (cf. Ames et al., 2008; Davis et al., 1996), one might reasonably expect that adopting an ingroup member's perspective would increase ingroup positivity. Contrary to this expectation, however, we found just the opposite: Perspective takers' ingroup visual representations were generally, though not always, less favorable than were those of control participants. This pattern of results aligns more closely with evidence that engaging in perspective taking prompts a deviation from default modes of information processing. On one account (Todd et al., 2016), contemplating another person's mental states initiates a comparison process that entails considering how the person is both similar to *and* different from oneself (cf. Mussweiler, 2003). For targets initially assumed to be different from oneself (e.g., outgroup members), this comparison process is posited to shine light on previously overlooked self–other similarities, resulting in more projection of (typically positive) self-attributes to these targets than otherwise. For targets initially assumed to be similar to oneself (e.g., ingroup members), however, the comparison process triggered by perspective taking is posited to highlight self–other differences, resulting in less projection of self-attributes to these targets than otherwise.

Although our findings can be accommodated by Todd et al.'s (2016) account, we did not specifically identify the differential projection of self-attributes to ingroup versus outgroup targets as a mechanism underlying the current results. Prior reverse-correlation research, however, indicates that visual representations of ingroup members resemble self-representations more than visual representations of outgroup members do (Imhoff & Dotsch, 2013). Integrating this prior work on visual representations with Todd et al.'s (2016) account of perspective-taking effects on conceptual representations leads to the prediction that perspective taking should increase visual self-projection for outgroup targets but should decrease visual self-projection for ingroup targets, relative to control. Future research should test this prediction.

Limitations and Additional Directions for Future Research

We acknowledge several limitations of the current work, which suggest fruitful avenues for future research. First, although our image-assessment experiments consistently revealed that perspective taking altered visualizations of ingroup and outgroup faces, it is important to note that our conclusions rest on a single image-generation experiment. Indeed, this limitation applies to most published research using two-phase reverse-correlation paradigms (but see Hong & Ratner, 2021; Petsko et al., 2020, for exceptions). Because the results of our image-assessment experiments necessarily depend on the psychological processes occurring during the image-generation experiment, future research that replicates the image-generation experiment with a new sample of participants promises to provide more clarity on the robustness of our findings.

Second, we used a new procedure for creating subgroup CIs from random subsets of individual CIs (see also Brown-Iannuzzi et al., 2021). This decision was motivated by concerns about the inflation of Type I error in group CIs and by preliminary evidence indicating that using subgroup CIs adequately controls Type I error (Cone et al., 2020; Lei et al., 2020). As is the case with any novel approach, there has been little validation research on the use of subgroup CIs. For instance, subgroup CIs might decrease Type II error by averaging across random noise while boosting the signal in classification images, as has been claimed for group CIs. To our knowledge, however, no research has addressed this question as of yet.

Perhaps more central to the validity of the procedure, we note that no recommendations yet exist for the optimal number of subgroup CIs to use. Nevertheless, results from our two image-assessment experiments with subgroup CIs were closely aligned with those from our image-assessment experiments with group CI and individual CIs, thereby increasing confidence in the validity of the conclusions drawn from evidence based on this procedure. In addition, we created two different randomly generated sets of subgroup CIs (one set for Experiment 2A, another set for Experiment 6) to ensure that our results were not an artifact of the particular randomization of individual CIs used for creating a given set of subgroup CIs. Subgroup CIs offer a unique solution for increasing the number of trials included in CIs while retaining some of the variability in the individual data. Future work will be needed to provide clearer guidelines for the optimal use of subgroup CIs in reverse-correlation research.

Another limitation of the current work is that we only examined visual representations of *minimal* ingroup and outgroup members. Because participants had no prior experience with these groups, they necessarily had to construct visual representations of group members' faces on the spot. Although such arbitrarily determined groups are psychologically meaningful (see Dunham, 2018, for a review), different effects might have emerged had we used more established, familiar groups (e.g., political groups, racial/ethnic groups) for which stored visual representations are presumably available for retrieval. Insofar as people have multiple visual representations of the same social group in memory (Hinzman & Maddox, 2017), engaging in perspective taking might change which of several representations is retrieved. Future research could address this issue by testing whether perspective taking has comparable effects on the generation of visual representations of familiar and unfamiliar groups.

Moreover, perspective taking may not always have a positive impact on visual representations of outgroup members. Evidence from several studies suggests that perspective taking may fail to positively alter conceptual representations of highly disliked or extremely negatively portrayed outgroup targets (Paluck, 2010; Skorinko & Sinclair, 2013) and can even negatively affect conceptual representations in circumstances of longstanding intergroup conflict (Bruneau & Saxe, 2012). Thus, when initial representations of an outgroup are strongly negative, perspective taking may be ineffectual or may even backfire (Sassenrath et al., 2016; Vorauer, 2013). In such cases, it seems plausible that being instructed to engage in perspective taking might elicit reactance in some perceivers (Berndsen et al., 2018). Future research that examines the impact of perspective taking on visual representations of strongly disliked groups should be helpful in identifying potential boundary conditions of the effects reported here.

The current findings also raise questions about specifics of the visualization process initiated by perspective taking. Essay task instructions in the perspective-taking condition directed participants to consider the target person's mental states (i.e., thoughts, feelings, intentions, and goals), whereas instructions in the control condition directed participants to imagine the person from an outside perspective (i.e., to ignore these mental states). Given that internal states are only available to the self, perhaps being instructed to imagine another person's internal states encourages the adoption of a first-person perspective, whereas the control instructions encourage the adoption of a third-person perspective (cf. Nigro & Neisser, 1983). Because opposing perspective-taking effects emerged for ingroup and outgroup visualizations, however, it seems unlikely that the distinct impressions elicited by the different task instructions are entirely attributable to adopting a first-person versus a third-person perspective.

A related issue is that essay instructions in the control condition might have inhibited any natural inclinations to engage in perspective taking while writing about the target person. Thus, our findings might say as much (and potentially more) about the absence of perspective taking in the control condition as its presence in the perspective-taking condition. A recent meta-analysis reported evidence that echoes this point: Encountering a person in distress leads people to engage in perspective taking without being prompted to do so, whereas “remain objective” instructions appear to inhibit this unprompted perspective-taking tendency (McAuliffe et al., 2020).

Although actively engaging in perspective taking may be the default response to a target person in distress, we contend that it is unlikely to be the default mode of processing a target person who is *not* in distress, as was the case in the current work. In line with this view, previous research examining the effect of perspective taking on the projection of self-attributes to ingroup and outgroup members used a “pure” control condition and reported results consistent with those observed here: Relative to participants in a condition wherein perspective taking was neither explicitly encouraged nor explicitly discouraged, perspective takers were more likely to use self-knowledge to make inferences about an outgroup member and less likely to use self-knowledge to make inferences about an ingroup member (Todd et al., 2016). Ultimately, we concur with Davis et al. (1996, 2004) that the issue of greatest theoretical importance here is that our two sets of essay task instructions elicited *differential* perspective-taking activity.

A final limitation concerns our reliance on participant samples of convenience. Although participants in the image-generation experiment (university students in the midwestern United States) and participants in the image-assessment experiments (MTurk workers and university students in the western United States) were drawn from different regions in the United States, some evidence suggests that the effect of perspective taking on intergroup bias may depend on the relational mobility of a culture. For example, Wang et al. (2018) found that perspective taking decreased racial stereotyping among participants from relationally mobile cultures (e.g., United States), but it had negligible effects on racial stereotyping among participants from relationally stable cultures (e.g., Singapore). Importantly, this research differs from the current work in focusing on conceptual representations of a group (e.g., Black people) rather than visual representations of an individual (minimal) group member. Nevertheless, this documentation of cultural differences should motivate future research to continue exploring whether perspective taking has distinct effects on intergroup bias in different cultural settings.

Conclusion

In sum, the current research offers new insights into the effects of perspective taking on intergroup bias. Extending previous findings indicating that perspective taking can reduce various forms of intergroup bias in how group members are conceptually represented (Todd & Galinsky, 2014), our results provide some initial evidence that actively considering group members' mental states may also alter their imagined appearance, resulting in less biased visual representations of their faces.

References

Ahumada, A., & Lovell, J. (1971). Stimulus features in signal detection. *The Journal of the Acoustical Society of America*, *49*(6B), 1751–1756. https://doi.org/10.1121/1.1912577

Ames, D. L., Jenkins, A. C., Banaji, M. R., & Mitchell, J. P. (2008). Taking another person's perspective increases self-referential neural processing. *Psychological Science, 19*, 642–644. https://doi.org/10.1111/j.1467-9280.2008.02135.x

Ames, D. R. (2004a). Inside the mind reader's tool kit: Projection and stereotyping in mental state inference. *Journal of Personality and Social Psychology, 87*, 340–353. https://doi.org/10.1037/0022-3514.87.3.340

Ames, D. R. (2004b). Strategies for social inference: A similarity contingency model of projection and stereotyping in attribute prevalence estimates. *Journal of Personality and Social Psychology, 87*, 573–585. https://doi.org/10.1037/0022-3514.87.5.573

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*(46), 17948–17953. https://doi.org/10.1073/pnas.0705435104

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., … Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology*, *72*(1), 105–118. https://doi.org/10.1037/0022-3514.72.1.105

Ben-Shachar, M., Makowski, D., & Lüdecke, D. (2020). Compute and interpret indices of effect size. *CRAN*. R package, https://github.com/easystats/effectsize.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, *10*(1), 122–142. https://doi.org/10.1006/game.1995.1027

Berndsen, M., Thomas, E. F., & Pedersen, A. (2018). Resisting perspective-taking: Glorification of the national group elicits non-compliance with perspective-taking instructions. *Journal of Experimental Social Psychology*, *79*, 126–137. https://doi.org/10.1016/j.jesp.2018.07.007

Bernstein, M., Young, S., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18*, 709–712. https://doi.org/10.1111/j.1467-9280.2007.01964.x

Bjornsdottir, R. T., Yeretsian, S., West, G. L., & Rule, N. O. (2019). When words become borders: Ingroup favoritism in perceptions and mental representations of Anglo-Canadian and Franco-Canadian faces. *Group Processes and Intergroup Relations*, *22*, 1021–1038. https://doi.org/10.1177/1368430218789495

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

Broockman, D., & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science, 352*, 220-224. https://doi.org/10.1126/science.aad9713

Brown-Iannuzzi, J. L., Cooley, E., Marshburn, C. K., McKee, S. E., & Lei, R. F. (2021). Investigating the interplay between race, work ethic stereotypes, and attitudes toward welfare recipients and policies. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550620983051

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

Bruneau, E. G., & Saxe, R. (2012). The power of being heard: The benefits of 'perspective-giving' in the context of intergroup conflict. *Journal of Experimental Social Psychology, 48*, 855–866. https://doi.org/10.1016/j.jesp.2012.02.017

Chang, L. J., Doll, B. B., van 't Wout, M., Frank, M. J., & Sanfey, A. G. (2010). Seeing is believing: Trustworthiness as a dynamic belief. *Cognitive Psychology*, *61*(2), 87–105. https://doi.org/10.1016/j.cogpsych.2010.03.001

Cikara, M., & Van Bavel, J. J. (2014). The neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science, 9*, 245–274. https://doi.org/10.1177/1745691614527464

Cone, J., Brown-Iannuzzi, J., Lei, R., & Dotsch, R. (2020). Type I error is inflated in the two-phase reverse correlation procedure. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550620938616

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology*, *70*(4), 713–726. https://doi.org/10.1037/0022-3514.70.4.713

Davis, M. H., Soderlund, T., Cole, J., Gadol, E., Kute, M., Myers, M., & Weihing, J. (2004). Cognitions associated with attempts to empathize: How do we imagine the perspective of another? *Personality and Social Psychology Bulletin*, *30*(12), 1625–1635. https://doi.org/10.1177/0146167204271183

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

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science*, *3*(5), 562–571. https://doi.org/10.1177/1948550611430272

Dotsch, R., Wigboldus, D. H. J., 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

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

Evans, A. M., & Krueger, J. I. (2009). The Psychology (and Economics) of Trust. *Social and Personality Psychology Compass*, *3*(6), 1003–1017. https://doi.org/10.1111/j.1751-9004.2009.00232.x

Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology*, *79*, 200–216. https://doi.org/10.1016/j.jesp.2018.07.004

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/bf03193146

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism. *Journal of Personality and Social Psychology*, *78*(4), 708–724. https://doi.org/10.1037/0022-3514.78.4.708

Gawronski, B., Bodenhausen, G. V., & Banse, R. (2005). We are, therefore they aren't: Ingroup construal as a standard of comparison for outgroup judgments. *Journal of Experimental Social*

Psychology, 41, 515–526. https://doi.org/10.1016/j.jesp.2004.11.002

Gomila, R. (2021). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General, 150,* 700–709. https://doi.org/10.1037/xge0000920

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*(6), 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York, NY: Guilford Press.

Hehman, E., Stolier, R. M., Freeman, J. B., Flake, J. K., & Xie, S. Y. (2019). Toward a comprehensive model of impression formation: What we know, what we do not, and paths forward. *Social and Personality Psychology Compass*, e12431. https://doi.org/10.1111/spc3.12431

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Hinzman, L., & Maddox, K. B. (2017). Conceptual and visual representations of racial categories: Distinguishing subtypes from subgroups. *Journal of Experimental Social Psychology*, *70*, 95–109. https://doi.org/10.1016/j.jesp.2016.12.012

Hong, Y., & Ratner, K. G. (2021). Minimal but not meaningless: Seemingly arbitrary category labels can imply more than group membership. *Journal of Personality and Social Psychology, 120*, 576–600. https://doi.org/10.1037/pspa0000255

Hugenberg, K. (2005). Social categorization and the perception of facial affect: Target race moderates the response latency advantage for happy faces. *Emotion, 5*, 267–276. https://doi.org/10.1037/1528-3542.5.3.267

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science, 14*, 640–643. https://doi.org/10.1046/j.0956-7976.2003.psci_1478.x

Imhoff, R., & Dotsch, R. (2013). Do we look like me or like us? Visual projection as self- or ingroup-projection. *Social Cognition*, *31*(6), 806–816. https://doi.org/10.1521/soco.2013.31.6.806

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and consequences of social categorization. *Advances in Experimental Social Psychology*, *55*, 1–80. https://doi.org/10.1016/bs.aesp.2016.10.001

Klapper, A., Dotsch, R., van Rooij, I., & Wigboldus, D. H. J. (2016). Do we spontaneously form stable trustworthiness impressions from facial appearance? *Journal of Personality and Social Psychology*, *111*(5), 655–664. https://doi.org/10.1037/pspa0000062

Kleiman, T., Hassin, R. R., & Trope, Y. (2014). The control-freak mind: Stereotypical biases are eliminated following conflict-activated cognitive control. *Journal of Experimental Psychology: General*, *143*(2), 498–503. https://doi.org/10.1037/a0033047

Krieglmeyer, R., & Sherman, J. W. (2012). Disentangling stereotype activation and stereotype application in the stereotype misperception task. *Journal of Personality and Social Psychology*, *103*(2), 205–224. https://doi.org/10.1037/a0028764

Krosch, A. R., & Amodio, D. M. (2014). Economic scarcity alters the perception of race. *Proceedings of* *the National Academy of Sciences*, *111*(25), 9079–9084. https://doi.org/10.1073/pnas.1404448111

Ku, G., Wang, C. S., & Galinsky, A. D. (2015). The promise and perversity of perspective-taking in organizations. *Research in Organizational Behavior, 35*, 79–102. https://doi.org/10.1016/j.riob.2015.07.003

Lazerus, T., Ingbretsen, Z., Stolier, R., Freeman, J. B., & Cikara, M. (2016). Positivity bias in judgments of in-group members' emotional expressions. *Emotion, 16*, 1117–1125. https://doi.org/10.1037/emo0000227

Liberman, Z., Woodward, A., & Kinzler, K. D. (2017). The origins of social categorization. *Trends in Cognitive Sciences, 21*, 556–586. https://doi.org/10.1016/j.tics.2017.04.004

Lei, R. F., & Bodenhausen, G. V. (2017). Racial assumptions color the mental representation of social class. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00519

Lenth, R. V. (2020). *emmeans: Estimated Marginal Means, aka Least-Squares Means.* R package version 1.5.3. https://CRAN.R-project.org/package=emmeans

Lord, C. G., Lepper, M. R., & Preston, E. (1984). Considering the opposite: A corrective strategy for social judgment. *Journal of Personality and Social Psychology*, *47*(6), 1231–1243. https://doi.org/10.1037/0022-3514.47.6.1231

Luke S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, *49*, 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

Lundqvist, D., & Litton, J. E. (1998). The averaged Karolinska directed emotional faces. *Stockholm: Karolinska Institute, Department of Clinical Neuroscience, Section Psychology*.

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, *51*(1), 93–120. https://doi.org/10.1146/annurev.psych.51.1.93

Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science*, *28*(2), 209–226. https://doi.org/10.1016/j.cogsci.2003.11.004

McAuliffe, W. H., Carter, E. C., Berhane, J., Snihur, A. C., & McCullough, M. E. (2020). Is empathy the default response to suffering? A meta-analytic evaluation of perspective taking's effect on empathic concern. *Personality and Social Psychology Review, 24*(2), 141–162. https://doi.org/10.1177/1088868319887599

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*, 472–489. https://doi.org/10.1037/0033-295X.110.3.472

Nigro, G., & Neisser, U. (1983). Point of view in personal memories. *Cognitive Psychology*, *15*(4), 467–482. https://doi.org/10.1016/0010-0285(83)90016-6

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*(1), 128–133. https://doi.org/10.1037/a0014520

Otten, S. (2003). "Me and us" or "us and them"? The self as a heuristic for defining minimal ingroups. *European Review of Social Psychology*, *13*(1), 1–33. https://doi.org/10.1080/10463280240000028

Paluck, E. L. (2010). Is it better not to talk? Group polarization, extended contact, and perspective-taking in eastern Democratic Republic of Congo. *Personality and Social Psychology Bulletin, 36*, 1170–1185. https://doi.org/10.1177/0146167210379868

Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Review of Psychology*, *60*(1), 339–367. https://doi.org/10.1146/annurev.psych.60.110707.163607

Paulus, A., Rohr, M., Dotsch, R., & Wentura, D. (2016). Positive feeling, negative meaning: Visualizing the mental representations of in-group and out-group smiles. *PLoS ONE*, *11*(3), e0151230. https://doi.org/10.1371/journal.pone.0151230

Petsko, C. D., Lei, R. F., Kunst, J. R., Bruneau, E., & Kteily, N. (2020). Blatant dehumanization in the mind's eye: Prevalent even among those who explicitly reject it? *Journal of Experimental Psychology: General.* Advance online publication. https://doi.org/10.1037/xge0000961

Ratner, K. G., Dotsch, R., Wigboldus, D. H. J., van Knippenberg, A., & Amodio, D. M. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal of Personality and Social Psychology*, *106*(6), 897–911. https://doi.org/10.1037/a0036498

R Core Team. (2019). *R: A language and environment for statistical computing (Version 3.6.0).* Vienna, Austria: R Foundation for Statistical Computing.

Rees, H. R., Rivers, A. M., & Sherman, J. W. (2019). Implementation intentions reduce implicit stereotype activation and application. *Personality and Social Psychology Bulletin, 45(1)*, 37–53. https://doi.org/10.1177/0146167218775695

Rivers, A. M., Sherman, J. W., Rees, H. R., Reichardt, R., & Klauer, K. C. (2020). On the roles of stereotype activation and application in diminishing implicit bias. *Personality and Social Psychology Bulletin*, *46*(3), 349–364. https://doi.org/10.1177/0146167219853842

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science*, *19*(2), 109–111. https://doi.org/10.1111/j.1467-9280.2008.02054.x

Rule, N. O., & Ambady, N. (2009). She's got the look: Inferences from female chief executive officers' faces predict their success. *Sex Roles*, *61*(9–10), 644–652. https://doi.org/10.1007/s11199-009-9658-9

Rule, N. O., Ambady, N., Adams, R. B., & Macrae, C. N. (2008). Accuracy and awareness in the perception and categorization of male sexual orientation. *Journal of Personality and Social Psychology*, *95*(5), 1019–1028. https://doi.org/10.1037/a0013194

Sassenrath, C., Hodges, S. D., & Pfattheicher, S. (2016). It's all about the self: When perspective taking backfires. *Current Directions in Psychological Science*, *25*(6), 405–410. https://doi.org/10.1177/0963721416659253

Shih, M., Wang, E., Trahan, A., & Stotzer, B. (2009). Perspective taking: Reducing prejudice towards general outgroups and specific individuals. *Group Processes and Intergroup Relations, 12*, 565–577. https://doi.org/10.1177/1368430209337463

Simpson, A. J., & Todd, A. R. (2017). Intergroup visual perspective-taking: Shared group membership impairs self-perspective inhibition but may facilitate perspective calculation. *Cognition*, *166*, 371–381. https://doi.org/10.1016/j.cognition.2017.06.003

Skorinko, J. L., & Sinclair, S. A. (2013). Perspective-taking can increase stereotyping: The role of apparent stereotype confirmation. *Journal of Experimental Social Psychology, 49*, 10–18.

https://doi.org/10.1016/j.jesp.2012.07.009

Steiger, J. H. (2004). Beyond the *F* test: Effect size confidence intervals and tests of close fit in the analysis of variance and contrast analysis. *Psychological Methods, 9*, 164–182. https://doi.org/10.1037/1082-989X.9.2.164

Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition, 127*, 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

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

Todd, A. R., Bodenhausen, G. V., Richeson, J. A., & Galinsky, A. D. (2011). Perspective taking combats automatic expressions of racial bias. *Journal of Personality and Social Psychology*, *100*(6), 1027–1042. https://doi.org/10.1037/a0022308

Todd, A. R., & Burgmer, P. (2013). Perspective taking and automatic intergroup evaluation change: Testing an associative self-anchoring account. *Journal of Personality and Social Psychology*, *104*(5), 786–802. https://doi.org/http://dx.doi.org/10.1037/a0031999

Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass*, *8*(7), 374–387. https://doi.org/10.1111/spc3.12116

Todd, A. R., Galinsky, A. D., & Bodenhausen, G. V. (2012). Perspective taking undermines stereotype maintenance processes: Evidence from social memory, behavior explanation, and information solicitation. *Social Cognition*, *30*(1), 94–108. https://doi.org/10.1521/soco.2012.30.1.94

Todd, A. R., Hanko, K., Galinsky, A. D., & Mussweiler, T. (2011). When focusing on differences leads to similar perspectives. *Psychological Science, 22*, 134–141. https://doi.org/10.1177/0956797610392929

Todd, A. R., Simpson, A. J., & Tamir, D. I. (2016). Active perspective taking induces flexible use of self-knowledge during social inference. *Journal of Experimental Psychology: General*, *145*(12), 1583–1588. https://doi.org/http://dx.doi.org/10.1037/xge0000237

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623–1626. https://doi.org/10.1126/science.1110589

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Uleman, J. S. (1999). Spontaneous versus intentional inferences in impression formation. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 141–160). New York, NY: Guilford Press.

van 't Wout, M., & Sanfey, A. G. (2008). Friend or foe: The effect of implicit trustworthiness judgments in social decision-making. *Cognition, 108*, 796–803. https://doi.org/10.1016/j.cognition.2008.07.002

Vorauer, J. (2013). The case for and against perspective-taking. *Advances in Experimental Social Psychology*, *48*, 59–115. https://doi.org/10.1016/B978-0-12-407188-9.00002-8

Wang, C. S., Lee, M., Ku, G., & Leung, A. K. y. (2018). The cultural boundaries of perspective-taking: When and why perspective-taking reduces stereotyping. *Personality and Social Psychology Bulletin*, *44*(6), 928–943. https://doi.org/10.1177/0146167218757453

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science*, *17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

Wilson, J. P., & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science*, *26*(8), 1325–1331. https://doi.org/10.1177/0956797615590992

Wilson, J. P., & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: The role of facial trustworthiness. *Social Psychological and Personality Science*, *7*(4), 331–338. https://doi.org/10.1177/1948550615624142

Xiao, Y. J., Coppin, G., & Van Bavel, J. J. (2016). Perceiving the world through group-colored glasses: A perceptual model of intergroup relations. *Psychological Inquiry, 27*, 255–274. https://doi.org/10.1080/1047840X.2016.1199221

Young, A. I., Ratner, K. G., & Fazio, R. H. (2014). Political attitudes bias the mental representation of a presidential candidate's face. *Psychological Science*, *25*(2), 503–510. https://doi.org/10.1177/0956797613510717

Yzerbyt, V., Muller, D., Batailler, C., & Judd, C. M. (2018). New recommendations for testing indirect effects in mediational models: The need to report and test component paths. *Journal of Personality and Social Psychology, 115*, 929–943. https://doi.org/10.1037/pspa0000132

Appendix

ANOVA Results at the Level of the Individual CI

In the image-assessment experiments reported in the main text, we conducted all analyses at the level of the participant (i.e., image rater) by collapsing across individual CIs. Image-assessment experiments in other reverse-correlation research have conducted analyses at the level of the individual CI by collapsing across image raters. In addition, recent simulation work examining Type I error inflation in two-phase reverse-correlation procedures (Cone et al., 2020) conducted their analysis at the level of the individual CIs. Below, we report the results summarizing at the level of the individual CIs for Experiments 2B–5.

Experiment 2B

A 2 (Group Membership) × 2 (Perspective Taking) ANOVA on the likability of the individual CIs revealed that the ingroup CIs (M = 3.91, SD = 0.58) were rated as more likable than were the outgroup CIs (M = 3.29, SD = 0.65), $F(1, 250) = 64.44$ $p < .001$, $\eta_p^2 = 0.21$, $CI_{90\%}$ [0.14, 0.28]. Although the Group Membership × Perspective Taking interaction was not significant, $F(1, 250) = 1.69$, $p = .195$, $\eta_p^2 = 0.01$, we examined the underlying pattern of simple effects.

Simple-effects tests indicated that the ingroup CIs were rated as more likable than were the outgroup CIs in both conditions, but this group difference was descriptively (but not significantly) smaller for the CIs created by perspective takers, $t(123.91) = 4.72$, $p < .001$, $d = 0.84$, $CI_{95\%}$ [0.47, 1.20], than for the CIs created by control participants, $t(124.81) = 6.66$, $p < .001$, $d = 1.18$, $CI_{95\%}$ [0.80, 1.56]. When approaching this interaction differently, no significant differences emerged between the ingroup CIs generated by control participants and those generated by perspective takers, $t(122.98) = 0.13$, $p = .900$, $d = 0.02$, nor did significant differences emerge between the outgroup CIs generated by control participants and those generated by perspective takers, $t(126.77) = –1.65$, $p = .101$, $d = –0.29$.

Experiment 3A

A 2 (Group Membership) × 2 (Perspective Taking) ANOVA on the proportion of “more likable” responses revealed that the ingroup CI primes (M = 0.61, SD = 0.05) elicited more likability than did the outgroup CI primes (M = 0.57, SD = 0.06), $F(1, 250) = 27.96$, $p < .001$, $\eta_p^2 = 0.10$, $CI_{90\%}$ [0.05, 0.16]. The

Group Membership × Perspective Taking interaction was also significant, $F(1, 250) = 7.53$, $p = .007$, $\eta_p^2 = 0.03$, $CI_{90\%}$ [0.005, 0.07].

Simple-effects tests indicated that the ingroup CIs created by control participants elicited more likability than did the outgroup CIs created by control participants, $t(121.70) = 6.03$, $p < .001$, $d = 1.06$, $CI_{95\%}$ [0.69, 1.43]. In contrast, the difference between the ingroup and outgroup CIs created by perspective takers was not significant, $t(124.38) = 1.71$, $p = .089$, $d = 0.30$. Approaching the interaction differently, perspective takers' ingroup CIs elicited more likability than did control participants' ingroup CIs, $t(116.78) = 2.10$, $p = .038$, $d = 0.38$, $CI_{95\%}$ [0.02, 0.73], whereas the difference between the outgroup CIs created by control participants and those created by perspective takers was not significant, $t(127.00) = –1.80$, $p = .075$, $d = –0.32$.

Experiment 3B

A 2 (Group Membership) × 2 (Perspective Taking) ANOVA on the proportion of "more trustworthy" responses revealed that the ingroup CIs ($M = 0.59$, $SD = 0.06$) elicited more trustworthiness than did the outgroup CIs ($M = 0.55$, $SD = 0.06$), $F(1, 250) = 20.73$, $p < .001$, $\eta_p^2 = .08$, $CI_{90\%}$ [0.03, 0.13]. Although the predicted Group Membership × Perspective Taking interaction was not significant, $F(1, 250) = 2.22$, $p = .137$, $\eta_p^2 = 0.01$, we examined the underlying pattern of simple effects.

Simple-effects analyses indicated that the ingroup CIs elicited more trustworthiness than did the outgroup CIs in both conditions, but this group difference in trustworthiness was directionally (but not significantly) smaller for the CIs created by perspective takers, $t(124.89) = 1.99$, $p = .048$, $d = 0.35$, $CI_{95\%}$ [0, 0.70], than for the CIs created by control participants, $t(124.00) = 4.71$, $p < .001$, $d = 0.84$, $CI_{95\%}$ [0.47, 1.20]. Approaching this interaction differently, no significant differences emerged between the ingroup CIs generated by control participants and those generated by perspective takers, $t(120.33) = 1.17$, $p = .245$, $d = 0.21$, nor did significant differences emerge between the outgroup CIs generated by control participants and those generated by perspective takers, $t(121.77) = –0.94$, $p = .351$, $d = –0.17$.

Experiment 4

A 2 (Group Membership) × 2 (Perspective Taking) ANOVA on the average amount of money transferred revealed that more money was entrusted to interaction partners represented by ingroup CIs ($M = 20.20$ cents, $SD = 1.05$) than to partners represented by outgroup CIs ($M = 19.70$ cents, $SD = 0.99$),

$F(1, 250) = 17.82$, $p < .001$, $\eta_p^2 = 0.07$, $CI_{90\%}$ [0.03, 0.12]. The predicted Group Membership × Perspective Taking interaction was significant, $F(1, 250) = 4.09$, $p = .044$, $\eta_p^2 = 0.02$, $CI_{90\%}$ [0.0002, 0.05].

Simple-effects tests indicated that participants entrusted more money to the ingroup CIs than to the outgroup CIs created by control participants, $t(124.41) = 4.38$, $p < .001$, $d = 0.78$, $CI_{95\%}$ [0.42, 1.14]. In contrast, the difference in money entrusted to the ingroup and outgroup CIs created by perspective takers was not significant, $t(122.83) = 1.56$, $p = .120$, $d = 0.28$. Approaching this interaction differently, no significant differences in trust decisions emerged between the ingroup CIs created by control participants and those created by perspective takers, $t(122.97) = 0.93$, $p = .354$, $d = 0.17$, nor did significant differences emerge between the outgroup CIs created by control participants and those created by perspective takers, $t(126.70) = -1.97$, $p = .051$, $d = -0.35$.

Experiment 5

A 3 (Occlusion) × 2 (Group Membership) × 2 (Perspective Taking) ANOVA on the proportion of individual CIs in each condition that elicited a "more trustworthy" response revealed main effects of Group Membership, $F(1, 247) = 32.21$, $p < .001$, $\eta_p^2 = 0.12$, $CI_{90\%}$ [0.06, 0.18], and Occlusion, $F(1.61, 396.91) = 50.12$, $p < .001$, $\eta_p^2 = 0.17$, $CI_{90\%}$ [0.12, 0.22]. The ingroup CIs ($M = 0.39$, $SD = 0.13$) were judged as more trustworthy than were the outgroup CIs ($M = 0.33$, $SD = 0.12$). Additionally, the CIs with both the eyes and mouth occluded ($M = 0.31$, $SD = 0.07$) were judged as less trustworthy than were the CIs with just the eyes occluded ($M = 0.40$, $SD = 0.18$), $t(332.45) = -7.51$, $p < .001$, $d = -0.67$, $CI_{95\%}$ [–0.85, –0.49], and the CIs with just the mouth occluded ($M = 0.39$, $SD = 0.10$), $t(451.30) = -9.93$, $p < .001$, $d = -0.88$, $CI_{95\%}$ [–1.06, –0.70]. Trustworthiness impressions for the eyes-occluded CIs and the mouth-occluded CIs did not significantly differ, $t(405.67) = 0.88$, $p = .380$, $d = 0.08$.

Although the predicted Group Membership × Perspective Taking interaction was not significant, $F(1, 247) = 3.67$, $p = .056$, $\eta_p^2 = 0.01$, we examined the underlying pattern of simple effects. The ingroup CIs were judged as more trustworthy than were the outgroup CIs, but this pattern of intergroup bias was directionally (but not significantly) weaker for the CIs generated by perspective takers, $t(376.09) = 2.75$, $p = .006$, $d = 0.28$, $CI_{95\%}$ [0.08, 0.48], than for the CIs generated by control participants, $t(367.79) = 5.79$, $p < .001$, $d = 0.60$, $CI_{95\%}$ [0.39, 0.80]. Approaching this interaction differently, perspective takers' outgroup CIs were judged as more trustworthy than were control participants' outgroup CIs, $t(381.35) = -2.56$, $p =$

.011, d = –0.26, $CI_{95\%}$ [–0.46, –0.06]. However, the difference in trustworthiness judgments for the ingroup CIs generated by control participants and those generated by perspective takers was not significant, $t(370.98) = 0.57$, $p = .567$, $d = 0.06$.

The Occlusion × Group Membership interaction was also significant, $F(1.61, 396.91) = 12.76$, $p < .001$, $\eta_p^2 = 0.05$, $CI_{90\%}$ [0.02, 0.08]. Simple-effects tests indicated that the ingroup CIs were judged as more trustworthy than were the outgroup CIs in both the eyes-occluded condition, $t(249.95) = 5.23$, $p < .001$, $d = 0.66$, $CI_{95\%}$ [0.40, 0.91], and the mouth-occluded condition, $t(251.94) = 3.74$, $p < .001$, $d = 0.47$, $CI_{95\%}$ [0.22, 0.72]; however, this pattern of intergroup bias was not significant in the eyes-and-mouth-occluded condition, $t(243.03) = 1.29$, $p = .200$, $d = 0.16$.

Finally, and of greatest importance, the Occlusion × Group Membership × Perspective Taking interaction was significant, $F(1.61, 396.91) = 5.69$, $p = .007$, $\eta_p^2 = 0.02$, $CI_{90\%}$ [0.004, 0.05]. To further specify this interaction, we conducted separate 2 (Group Membership) × Perspective Taking ANOVAs in each occlusion condition. When just the eyes were occluded, the Group Membership main effect, $F(1, 248) = 27.84$, $p < .001$, $\eta_p^2 = 0.10$, $CI_{90\%}$ [0.05, 0.16], and the Group Membership × Perspective Taking interaction, $F(1, 248) = 7.50$, $p = .007$, $\eta_p^2 = 0.03$, $CI_{90\%}$ [0.005, 0.07], were both significant. Simple-effects tests in the eyes-occluded condition indicated that the ingroup CIs created by control participants were judged as more trustworthy than were the outgroup CIs created by control participants, $t(123.42) = 5.81$, $p < .001$, $d = 1.04$, $CI_{95\%}$ [0.66, 1.41]; however, no significant differences emerged between the ingroup and outgroup CIs created by perspective takers, $t(123.63) = 1.75$, $p = .082$, $d = 0.31$. When just the mouth was occluded, the Group Membership main effect was significant, $F(1, 250) = 13.88$, $p < .001$, $\eta_p^2 = 0.05$, $CI_{90\%}$ [0.02, 0.10], but it was not moderated by Perspective Taking, $F(1, 250) = 0.13$, $p = .717$, $\eta_p^2 = 0.0005$. When both the eyes and mouth were occluded, however, neither main effect nor the interaction was significant.

ANOVA Results involving Gender in Experiment 6

In the main text, we collapsed across participant/CI gender in all analyses. For each experiment, we conducted the same analyses reported in the main text while retaining participant/CI gender as a factor. Below, we report cases (both in Experiment 6) in which Gender significantly moderated the critical Group Membership × Perspective Taking interaction.

Frowning/Scowling Ratings

A significant Gender × Group Membership × Perspective Taking interaction emerged on the frowning/scowling ratings, $F(1, 142) = 12.63$, $p = .001$, $\eta_p^2 = 0.08$, $CI_{90\%}$ [0.02, 0.16]. To decompose this three-way interaction, we conducted separate 2 (Group Membership) × 2 (Perspective Taking) ANOVAs for the female CIs and the male CIs.

For the female CIs rated by women, there were significant main effects of Group Membership, $F(1, 70) = 181.79$, $p < .001$, $\eta_p^2 = 0.72$, $CI_{90\%}$ [0.63, 0.78], and Perspective Taking, $F(1, 70) = 44.22$, $p < .001$, $\eta_p^2 = 0.39$, $CI_{90\%}$ [0.24, 0.51], and a significant Group Membership × Perspective Taking interaction, $F(1, 70) = 32.22$, $p < .001$, $\eta_p^2 = 0.32$, $CI_{90\%}$ [0.17, 0.44]. Female outgroup CIs generated by perspective takers ($M = 4.23$, $SD = 1.01$) were rated as frowning/scowling less than were female outgroup CIs generated by control participants ($M = 4.90$, $SD = 1.10$), $t(70) = 7.80$, $p < .001$, $d = 0.93$, $CI_{95\%}$ [0.65, 1.21], whereas female ingroup CIs generated by perspective takers and control participants did not significantly differ, $t(70) = 1.60$, $p = .114$, $d = 0.19$.

For the male CIs rated by men, there was a significant Group Membership main effect, $F(1, 72) = 145.71$, $p < .001$, $\eta_p^2 = 0.67$, $CI_{90\%}$ [0.57, 0.74], and a significant Group Membership × Perspective Taking interaction, $F(1, 72) = 74.97$, $p < .001$, $\eta_p^2 = 0.51$, $CI_{90\%}$ [0.38, 0.61]. The male outgroup CIs generated by perspective takers ($M = 3.99$, $SD = 1.08$) were rated as frowning/scowling less than were the male outgroup CIs generated by control participants ($M = 4.63$, $SD = 1.02$), $t(72) = 7.24$, $p < .001$, $d = 0.85$, $CI_{95\%}$ [0.58, 1.12], whereas the male ingroup CIs generated by perspective takers ($M = 3.39$, $SD = 1.11$) were rated as frowning/scowling more than were the male ingroup CIs generated by control participants ($M = 2.87$, $SD = 1.25$), $t(72) = –5.83$, $p < .001$, $d = –0.68$, $CI_{95\%}$ [–0.94, –0.43].

Trustworthiness Impressions

A significant Gender × Group Membership × Perspective Taking interaction also emerged on the trustworthiness impressions, $F(1, 142) = 15.20$, $p < .001$, $\eta_p^2 = 0.10$, $CI_{90\%}$ [0.03, 0.18]. To decompose this three-way interaction, we conducted separate 2 (Group Membership) × 2 (Perspective Taking) ANOVAs for the female CIs and the male CIs.

For female CIs rated by women, there were significant main effects of Group Membership $F(1, 70) = 173.27$, $p < .001$, $\eta_p^2 = 0.71$, $CI_{90\%}$ [0.62, 0.78], and Perspective Taking, $F(1, 70) = 45.45$, $p < .001$,

$\eta_p^2 = 0.39$, $CI_{90\%}$ [0.25, 0.51], and a significant Group Membership × Perspective Taking interaction, $F(1, 70) = 23.66$, $p < .001$, $\eta_p^2 = 0.25$, $CI_{90\%}$ [0.12, 0.39]. The female outgroup CIs generated by perspective takers ($M = 3.34$, $SD = 0.91$) were rated as more trustworthy than were the female outgroup CIs generated by control participants ($M = 2.84$, $SD = 1.06$), $t(70) = –8.16$, $p < .001$, $d = –0.97$, $CI_{95\%}$ [–1.26, –0.69]. Similarly, the female ingroup CIs generated by perspective takers ($M = 4.91$, $SD = 0.75$) were rated as more trustworthy than were the female ingroup CIs generated by control participants ($M = 4.76$, $SD = 0.68$), $t(70) = –2.60$, $p = .011$, $d = –0.31$, $CI_{95\%}$ [–0.55, –0.07].

For male CIs rated by men, there was a significant main effect of Group Membership, $F(1, 72) = 111.43$, $p < .001$, $\eta_p^2 = 0.61$, $CI_{90\%}$ [0.49, 0.69], and a significant Group Membership × Perspective Taking interaction, $F(1, 72) = 52.76$, $p < .001$, $\eta_p^2 = 0.42$, $CI_{90\%}$ [0.28, 0.54]. The male outgroup CIs generated by perspective takers ($M = 3.71$, $SD = 1.00$) were rated as more trustworthy than were the male outgroup CIs generated by control participants ($M = 3.15$, $SD = 1.07$), $t(72) = –7.19$, $p < .001$, $d = –0.84$, $CI_{95\%}$ [–1.11, –0.58], whereas the male ingroup CIs generated by perspective takers ($M = 4.22$, $SD = 0.89$) were rated as less trustworthy than were the male ingroup CIs generated by control participants ($M = 4.58$, $SD = 0.96$), $t(72) = 4.28$, $p < .001$, $d = 0.50$, $CI_{95\%}$ [0.26, 0.75].

Mixed-Effects Models

In image-assessment Experiments 2B–5, we used the individual CIs, which afford the use of more conservative mixed-effects models that simultaneously capture participant (i.e., image rater) and stimulus (i.e., image generator CIs) variance. In the results reported below, we fit mixed-effects models that fully crossed image raters and CIs, treating both as random factors.

Experiment 2B

The linear mixed-effects regression included fixed effects for Group Membership (outgroup = -1, ingroup = 1) and Perspective Taking (control = -1, perspective taking = 1). For the outcome variable, higher scores reflect more likability explicitly ascribed to the CI (0 = “less likable” response, 1 = “more likable” response).[13] The reported model contains the maximal random-effects structure, which includes random intercepts for image raters and CIs and by-rater random slopes for Group Membership,

[13] We fit linear mixed-effects models on our binary data in response to recent work highlighting the equivalence of linear and logistic mixed-effects models (Gomila, 2021).

Perspective Taking, and the Group Membership × Perspective Taking interaction. Because participants in the image-generation experiment generated a CI for only one of the four between-subjects conditions, the maximal model contained only a random intercept for CIs.

This analysis revealed a significant main effect of Group Membership, $b = 0.31$, $SE = 0.04$, $F(1, 261.47) = 63.27$, $p < .001$,[14] indicating that the ingroup CIs were rated as more likable than were the outgroup CIs. Although the predicted Group Membership × Perspective Taking interaction was not significant, $b = 0.05$, $SE = 0.04$, $F(1, 251.33) = 1.63$, $p = .202$, we examined the underlying pattern of simple effects.

Pairwise comparisons[15] indicated that, although a robust pattern of intergroup bias emerged in both conditions, the difference in likability impressions for the ingroup and outgroup CIs was descriptively (but not significantly) smaller for the CIs generated by perspective takers, $b = 0.52$, $SE = 0.11$, $t(256) = 4.75$, $p < .001$, than for the CIs generated by control participants, $b = 0.72$, $SE = 0.11$, $t(256) = 6.55$, $p < .001$.[16] Approaching this interaction differently, pairwise comparisons revealed no significant difference between the outgroup CIs, $b = 0.18$, $SE = 0.10$, $t(251) = 1.68$, $p = .094$, nor between the ingroup CIs, $b = –0.02$, $SE = 0.11$, $t(251) = –0.14$, $p = .888$, generated by perspective takers and control participants.

Experiment 3A

As in Experiment 2B, the linear mixed-effects model in Experiment 3A included fixed effects for Group Membership and Perspective Taking. The outcome of interest was the probability that a CI evoked a "more likable" response, with higher scores reflecting more likability spontaneously elicited by the CI. The initial logistic mixed-effects model had the maximal random-effects structure (i.e., random intercepts by-rater and by-CI, as well as by-rater random slopes for the two fixed effects and their interaction); however, this model had singular fit. We removed the random slopes for Perspective Taking and the Group Membership × Perspective Taking interaction because these two parameters accounted for near zero variance. The reported random-effects structure contained by-rater and by-CI random intercepts and a by-rater random slope for Group Membership.

---

[14] We used the Kenward-Roger approximation for degrees-of-freedom to generate *F*-statistics and *p*-values. This approach confers adequate control of Type I error rates (Luke, 2017).
[15] We computed pairwise comparisons using the *pairs* function in the *emmeans* package (Lenth, 2020).
[16] Degrees-of-freedom for post-hoc analyses were computed using Satterthwaite's method.

The linear mixed-effects regression revealed a significant main effect of Group Membership, b = 0.02, SE = 0.004, $F(1, 203.78) = 23.00$, $p < .001$, indicating that the ingroup CIs were more likely to elicit a "more likable" response than were outgroup CIs. The Group Membership × Perspective Taking interaction was also significant, b = 0.01, SE = 0.003, $F(1, 254.31) = 8.23$, $p = .004$.

Pairwise comparisons indicated that, for control participants, the ingroup CIs elicited more likability than did the outgroup CIs, b = 0.06, SE = 0.01, $t(213) = 5.50$, $p < .001$. In contrast, for perspective takers, the difference in likability elicited by the ingroup and outgroup CIs was not significant, b = 0.02, SE = 0.01, $t(214) = 1.73$, $p = .085$. Examining the interaction differently, the ingroup CIs generated by perspective takers elicited less likability than did the ingroup CIs generated by control participants, b = –0.02, SE = 0.01, $t(219) = –2.03$, $p = .044$, whereas the outgroup CIs generated by perspective takers elicited more likability than did the outgroup CIs generated by control participants, b = 0.02, SE = 0.01, $t(213) = 2.03$, $p = .043$.

Experiment 3B

The reported model had the maximal random-effects structure: a by-CI random intercept, a by-rater random intercept, and by-rater random slopes for all fixed-effects parameters.[17] As in the previous experiments, there was a significant main effect of Group Membership, b = 0.02, SE = 0.004, $F(1, 286.76) = 20.93$, $p < .001$, indicating that the ingroup CIs were more likely to elicit a "more trustworthy" response than were the outgroup CIs. Although the Group Membership × Perspective Taking interaction was not significant, b = 0.007, SE = 0.004, $F(1, 259.08) = 3.43$, $p = .065$, we examined the underlying pattern of simple effects.

Significant intergroup bias emerged in both conditions, but it was nominally (albeit non-significantly) weaker for the CIs generated by perspective takers, b = 0.02, SE = 0.01, $t(421) = 2.13$, $p = .034$, than for the CIs generated by control participants, b = 0.05, SE = 0.01, $t(420) = 4.62$, $p < .001$. The trustworthiness elicited by neither the outgroup CIs, b = 0.01, SE = 0.01, $t(420) = 1.16$, $p = .249$, nor the ingroup CIs, b = –0.02, SE = 0.01, $t(421) = –1.46$, $p = .144$, differed for perspective takers and control participants.

[17] This model had singular fit but removing random effects accounting for near-zero variance resulted in convergence errors. Parameter estimates were nearly identical across all the fitted models, and no model suggested a different interpretation of the data.

Experiment 4

The linear mixed-effects model contained the same fixed effects and contrast coding as in the previous experiments. The outcome of interest was the size of the transfer in cents. The initial model contained the maximal random-effects structure (i.e., by-rater and by-CI random intercepts and by-rater slopes for Group Membership, Perspective Taking, and their interaction); however, this model had singular fit. We removed the random slopes for the Group Membership × Perspective Taking interaction because this parameter accounted for near-zero variance. The final model included by-rater and by-CI random intercepts and random slopes for Group Membership and Perspective Taking.

A significant Group Membership main effect emerged, $b = 0.29$, $SE = 0.06$, $F(1, 251.77) = 23.25$, $p < .001$, indicating that partners represented by ingroup CIs were entrusted with more money than were partners represented by outgroup CIs. The Group Membership × Perspective taking interaction was also significant, $b = 0.12$, $SE = 0.05$, $F(1, 251.01) = 7.05$, $p = .008$.

Although there was a clear pattern of intergroup bias in both conditions, it was weaker for the CIs generated by perspective takers, $b = 0.33$, $SE = 0.15$, $t(310) = 2.21$, $p = .028$, than for the CIs generated by control participants, $b = 0.81$, $SE = 0.15$, $t(311) = 5.44$, $p < .001$. Examining this interaction differently, whereas the ingroup CIs generated by perspective takers received less money than did the ingroup CIs generated by control participants, $b = –0.27$, $SE = 0.13$, $t(240) = –2.03$, $p = .042$, the amount of money entrusted to outgroup CIs generated by perspective takers and those generated by control participants did not significantly differ, $b = 0.22$, $SE = 0.13$, $t(239) = 1.65$, $p = .100$.

Experiment 5

The linear mixed-effects model contained the same fixed effects as the previous experiments as well as a fixed effect for occlusion. Because the occlusion variable had three levels, it was coded across two variables (*Mask1*: eye-and-mouth occluded = 1, eyes occluded = 0, mouth occluded = -1; *Mask2*: eye-and-mouth occluded = 0, eyes occluded = 1, mouth occluded = -1). We first fit the maximal model (i.e., by-rater and by-CI random intercepts, by-CI random slope for Occlusion, and by-rater random slopes for Group Membership, Perspective Taking, and their interaction), but this model had singular fit. We simplified this model by removing the by-rater random slopes for Perspective Taking and the Group Membership × Perspective Taking interaction.

As in previous experiments, there was a significant Group Membership main effect, *b* = 0.03, *SE* = 0.01, $F(1, 271.56) = 29.82$, $p < .001$, indicating that the ingroup CIs were rated as more trustworthy than were the outgroup CIs. There was also a significant Occlusion main effect, $b_{Mask1} = –0.06$, *SE* = 0.02, $b_{Mask2} = 0.03$, *SE* = 0.02, $F(2, 452.96) = 5.49$, $p = .004$. The CIs with the eyes and mouth occluded were rated as less trustworthy than were the CIs with just the eyes occluded, *b* = –0.09, *SE* = 0.03, $t(432) = –3.01$, $p = .008$, and as less trustworthy than were the CIs with just the mouth occluded, *b* = –0.08, *SE* = 0.03, $t(448) = –2.62$, $p = .018$. The difference in the trustworthiness of the CIs with just the eyes occluded and those with just the mouth occluded was not significant, *b* = 0.01, *SE* = 0.03, $t(496) = 0.31$, $p = .754$.

The Group Membership × Perspective Taking interaction was also significant, *b* = 0.01, *SE* = 0.005, $F(1, 237.44) = 4.84$, $p = .029$. Although there was significant intergroup bias in both conditions, it was smaller for the CIs generated by perspective takers, *b* = 0.03, *SE* = 0.01, $t(258) = 2.43$, $p = .016$, than for the CIs generated by control participants, *b* = 0.08, *SE* = 0.01, $t(257) = 5.47$, $p < .001$. Examining the interaction differently, the outgroup CIs created by perspective takers were rated as more trustworthy than were outgroup CIs generated by control participants, *b* = 0.03, *SE* = 0.01, $t(238) = 2.30$, $p = .022$, whereas the ingroup CIs generated by perspective takers and control participants did not significantly differ, *b* = –0.01, *SE* = 0.01, $t(237) = –0.82$, $p = .412$.

In addition, there was a Group Membership × Occlusion interaction, $b_{Group \times Mask1} = –0.02$, *SE* = 0.005, $b_{Group \times Mask2} = 0.03$, *SE* = 0.007, $F(2, 331.02) = 15.33$, $p < .001$. Intergroup bias was evident for the CIs with just the eyes occluded, *b* = 0.11, *SE* = 0.02, $t(507) = 6.51$, $p < .001$, and for the CIs with just the mouth occluded, *b* = 0.05, *SE* = 0.02, $t(526) = 2.60$, $p = .010$, but not for the CIs with both the eyes and mouth occluded, *b* = 0.01, *SE* = 0.01, $t(496) = 0.93$, $p = .353$.

Finally, and of most importance, the Group Membership × Perspective Taking × Occlusion interaction was significant, $b_{Group \times PT \times Mask1} = –0.01$, *SE* = 0.004, $b_{Group \times PT \times Mask2} = 0.02$, *SE* = 0.006, $F(2, 366.67) = 6.36$, $p = .002$. To interpret this interaction, we fit separate models for each occlusion condition.[18] For the CIs with the eyes and mouth occluded, there were no significant effects. For the CIs with only the mouth occluded, there was a significant Group Membership main effect, *b* = 0.02, *SE* =

[18] These models had a simplified random effects structure (only random intercepts) due to convergence errors.

0.01, $F(1, 236.15) = 14.29$, $p < .001$. No other effects were significant. Finally, for the CIs with only the eyes occluded, the Group Membership main effect was significant, $b = 0.06$, $SE = 0.01$, $F(1, 241.43) = 28.31$, $p < .001$, as was the Group Membership × Perspective Taking interaction, $b = 0.03$, $SE = 0.01$, $F(1, 241.62) = 7.11$, $p = .008$.