Visualizing Minimal Ingroup and Outgroup Faces: Implications for Impressions, Attitudes, and Behavior

Kyle G. Ratner
The Ohio State University and New York University

Ron Dotsch, Daniel H. J. Wigboldus, and
Ad van Knippenberg
Radboud University Nijmegen

David M. Amodio
New York University

More than 40 years of research have shown that people favor members of their ingroup in their impressions, attitudes, and behaviors. Here, we propose that people also form different mental images of minimal ingroup and outgroup members, and we test the hypothesis that differences in these mental images contribute to the well-established biases that arise from minimal group categorization. In Study 1, participants were assigned to 1 of 2 groups using a classic minimal group paradigm. Next, a reverse correlation image classification procedure was used to create visual renderings of ingroup and outgroup face representations. Subsequently, a 2nd sample naive to the face generation stage rated these faces on a series of trait dimensions. The results indicated that the ingroup face was significantly more likely than the outgroup face to elicit favorable impressions (e.g., trusting, caring, intelligent, attractive). Extending this finding, Study 2 revealed that ingroup face representations elicited more favorable implicitly measured attitudes than did outgroup representations, and Study 3 showed that ingroup faces were trusted more than outgroup faces during an economic game. Finally, Study 4 demonstrated that facial physiognomy associated with trustworthiness more closely resembled the facial structure of the average ingroup than outgroup face representation. Together, these studies suggest that minimal group distinctions can elicit different mental representations, and that this visual bias is sufficient to elicit ingroup favoritism in impressions, attitudes and behaviors.

Keywords: ingroup favoritism, minimal group paradigm, reverse correlation, face processing, mental representation

Intergroup biases are not always the result of long-standing rivalries. They may arise even in conditions where people do not know each other and have no conflicts of interest or any preexisting animosity toward each other. Indeed, mere identification with one of two distinct groups is sufficient to elicit a preference for one group over another (Tajfel, 1970; Tajfel, Billig, Bundy, & Flament, 1971; Tajfel & Turner, 1986). Once formed, a *minimal group* identity can influence a wide range of responses, including impressions, attitudes, and behaviors that have consequences for members of the ingroup as well as the outgroup (e.g., Ashburn-Nardo, Voils, & Monteith, 2001; Brewer & Silver, 1978; Locksley,

Ortiz, & Hepburn, 1980; Otten & Wentura, 1999; Tajfel et al., 1971; Van Bavel & Cunningham, 2009).

In the last decade, researchers have begun to examine minimal group effects on visual perception and memory (e.g., Bernstein, Young, & Hugenberg, 2007; Hugenberg & Corneille, 2009; Ratner & Amodio, 2013; Van Bavel, Packer, & Cunningham, 2008, 2011). The recent focus on this topic reflects a broader recognition in social psychology that studying how people make sense of faces and other visual information provides insight into the nature of social interactions (Hassin & Trope, 2000; Macrae, Quinn, Mason, & Quadflieg, 2005; Todorov, Said, Engell, & Oosterhof, 2008;

Zebrowitz & Montepare, 2008). However, unlike other minimal group research that has focused on idealized representations of ingroup members as pleasant and trustworthy (e.g., Brewer & Silver, 1978; Tajfel et al., 1971), the work in visual processing has largely focused on how group membership may affect the degree of attention and perceptual resources allocated to the processing of ingroup as opposed to outgroup faces (e.g., Bernstein et al., 2007; Ratner & Amodio, 2013; Van Bavel et al., 2008, 2011; Young & Hugenberg, 2010). As a result, the literature has not addressed the question of how one's actual mental representation of an ingroup or outgroup member's face may be distorted by the top-down influence of mere group categorization or how such distortion might contribute to intergroup bias.

The Contribution of Facial Representations to Intergroup Responses

Facial representations are key to social interactions because they convey information about a person's intentions, reactions, and dispositions. Not surprisingly, people are highly attuned to facial information. Among both children and adults, faces receive preferential attention and processing over other objects (Birmingham, Bischof, & Kingstone, 2009a, 2009b; Goren, Sarty, & Wu, 1975; Johnson & Morton, 1991), and this preference is supported by the rapid encoding of faces in neural regions involved in high-level vision (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Haxby, Hoffman, & Gobbini, 2000; Kanwisher, McDermott, & Chun, 1997). Neuroimaging findings suggest that mental images of faces are processed in much the same way as the direct perception of a face. Indeed, the same neural regions involved in direct face perception are also activated in response to an imagined face, suggesting that, as with perceiving actual faces, mental imagery of a face may profoundly shape the way information about a person is processed (Ishai, Haxby, & Ungerleider, 2002; Mechelli, Price, Friston, & Ishai, 2004; O'Craven & Kanwisher, 2000).

Given the weight afforded to facial cues during information processing, it is not surprising that information about appearance figures prominently in group-based knowledge structures. Kessler and McKenna (1978) reported that when people are asked to make gender inferences, they rely on physical characteristics to a larger degree than traits and behaviors. Furthermore, the inclusion of information about physical attributes increases the vividness and specificity of trait-based stereotypes (Andersen & Klatzky, 1987) and the likelihood that trait stereotypes will be applied during impression formation (Deaux & Lewis, 1984). Thus, representations of physical appearance promote the use of group-based stereotypes when making trait inferences.

The visual component of social group representations not only triggers trait inferences but also contributes to attitudes and behaviors. For instance, in the United States, Afrocentric features have been shown to activate negative attitudes toward Black individuals (Livingston & Brewer, 2002). In other work, White Americans who simulated the role of a patrolling police officer during a video game scenario were more likely to shoot unarmed Black men whose appearance was more prototypical of the Black stereotype (Ma & Correll, 2011). Furthermore, it has been shown that Black men accused of a crime receive harsher sentences when their appearance is closer to this stereotypical prototype (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, &

Johnson, 2006). Thus, visual representations constitute a powerful mechanism for conveying group-based prejudices.

Group identity is also related to the mental images people form of others, and these images have implications for social judgments and behavior. For example, Beaupré and Hess (2003) found that people were more likely to assume that targets portrayed in neutrally written vignettes were smiling if they shared the same ethnic group as the perceiver. Additionally, Dotsch, Wigboldus, Langner, and van Knippenberg (2008) observed an association between their Dutch participants' implicitly measured attitudes and the extent to which these participants formed mental images of Moroccan faces as appearing criminal and untrustworthy. In other work, Blair, Ma, and Lenton (2001) found that the strength of people's stereotypic associations could be weakened by asking them to engage in counterstereotypical mental imagery. Similarly, imagined contact with outgroups has been shown to reduce intergroup anxiety and improve attitudes toward several groups, including the elderly, homosexuals, and Muslims (Crisp & Turner, 2009; Turner & Crisp, 2010; Turner, Crisp, & Lambert, 2007). Together, the existing literature suggests that mental representations are important for guiding intergroup responses. However, the present question—of whether the simple act of categorizing others into the same or different social group as our own contributes to these mental representations and subsequent reactions—has not yet been directly addressed.

Clues That Mere Group Categorization Influences Visual Representations

Although most research on visual processing has focused on ingroup processing advantages (e.g., promoting perceptual accuracy), there are clues in the literature that ingroup favoritism under minimal group conditions can arise from distortions in facial representations, as we hypothesize. In the first wave of minimal group studies, Doise and colleagues (1972) examined whether anticipating a cooperative or a competitive context influences images of minimally generated groups and whether this effect on imagery could justify one's intended behavior toward the outgroup. In their study, they included evaluations of the groups on four physical trait dimensions: blond–dark, tall–short, fat–thin, colorful–quiet. Of importance for the present purposes, they also included a control group that consisted of minimal ingroups and outgroups with no anticipated interaction. Doise and colleagues did not conduct statistical analyses on differences between the ingroup and outgroup representations for the control group, but the data suggest that the ingroup was viewed to have physical traits deemed more favorable than the outgroup. A later study by Johnson (1981) investigated whether people have more favorable mental images of people belonging to their political party (a group with no clear perceptual basis). He found that when participants were presented with photographs of people who could be members either of their political party or of an opposing party, they selected more attractive people as belonging to their party. Most recently, Dunham (2011) reported that rapidly presented happy faces were more likely to be remembered as belonging to ingroup members. This link between happy expressions and the ingroup is consistent with the existence of a top-down expectation that ingroup faces comprise more favorable physiognomic information than outgroup faces.

On the basis of past work, we propose that mere membership in a social group has implications for the way faces of ingroup and outgroup members are represented in the mind and that differences in these mental representations can influence trait impressions, attitudes, and behaviors toward ingroup and outgroup members. In this way, mental representations of ingroup and outgroup members' faces may provide a mechanism through which mere group categorization contributes to stereotyping, prejudice, and discrimination. We tested this hypothesis across four studies.

Study 1

The objective of Study 1 was to test for differences in participants' mental representations of ingroup and outgroup faces and to assess whether these differences contribute to trait impressions formed by perceivers. This study was conducted in two parts, in which we (a) obtained visual renderings of participants' mental representations of minimally defined ingroup or outgroup members and (b) collected trait ratings of these renderings from a separate group of participants naive to the origin of the face images. This approach is described in more detail in what follows.

In Part 1, individuals were assigned to minimal groups and then categorized faces as belonging to either their ingroup or an outgroup. To capture the representational bias that we hypothesize can give rise to ingroup favoritism, we used a technique called *reverse correlation image classification.* Reverse correlation image classification belongs to a class of techniques that follow *reverse correlation* logic. Reverse correlation methods examine responses to many different stimuli and infer patterns in the stimuli that may have caused the responses. These patterns can be visualized and provide an approximation of the mental representations upon which participants based their responses. These techniques were developed over 40 years ago and have been applied to study a diverse range of topics, including auditory cognition (Ahumada & Lovell, 1971), neurophysiology (Ringach & Shapley, 2004; Victor, 2005), and low-level vision (Ahumada, 2002; Solomon, 2002). Recently, they have proved useful for investigating representations underlying face categorization (Dotsch & Todorov, 2012; Gosselin & Schyns, 2003; Mangini & Biederman, 2004; Martin-Malivel, Mangini, Fagot, & Biederman, 2006) and social cognitive biases in face processing (Dotsch et al., 2008; Dotsch, Wigboldus, & van Knippenberg, 2011; Imhoff, Dotsch, Bianchi, Banse, & Wigboldus, 2011; Jack, Caldara, & Schyns, 2012; Karremans, Dotsch, & Corneille, 2011; Young, Ratner, & Fazio, 2014). Thus, reverse correlation image classification provides a purely data-driven method for creating visual renderings of people's mental representations of faces that are resistant to experimenter bias (Mangini & Biederman, 2004; Todorov, Dotsch, Wigboldus, & Said, 2011).

In Part 2, we assessed whether these visual renderings could reveal biases that contribute to differential impressions of ingroup and outgroup members. To this end, we averaged together the visual renderings (also called classification images) generated for each participant to create grand-averaged ingroup and outgroup classification images. Then, an independent sample of participants was used to assess the impressions elicited by these grand-averaged ingroup and outgroup face representations. In order to determine the types of impressions people might discern from facial representations, we focused on trait dimensions that Oost-

erhof and Todorov (2008) have shown to be the most common types of impressions that people naturally draw from faces.

Study 1 was designed to test the role of facial representations as a mechanism involved in the effect of minimal group assignment on ingroup favoritism, in a manner analogous to a mediation pattern. Because much previous research has established the "direct path" between minimal group assignment and ingroup favoritism (e.g., Brewer & Silver, 1978), our design adopted the "experimental mediation" approach to examine a mechanism that contributes to this well-established effect (Spencer, Zanna, & Fong, 2005). We predicted, in line with the previously reviewed research and theory, that the ingroup facial representation would be rated more highly on traits signaling prosociality (e.g., trustworthiness, caring, sociality) and overall group fitness (e.g., intelligence, attractiveness, confidence) compared with the outgroup face representation.

Method

Part 1: Generating visual renderings of group-based facial representations

Participants. One hundred and seventy-six New York University students were recruited to participate in exchange for course credit. Up to four participants were run simultaneously in separate cubicle rooms.

Procedure. Upon arrival at the study session, participants were asked to complete a consent form and were told that they would perform several tasks on a computer. Next, a classic "dot estimation" minimal group procedure was used to assign participants to arbitrary, but believable, groups (Brown, Collins, & Schmidt, 1988; Gerard & Hoyt, 1974; Mussweiler, Gabriel, & Bodenhausen, 2000; Tajfel et al., 1971).

Numerical Estimation Style Test (NEST). Participants were led to believe that people vary in numerical estimation style, which was defined as the tendency to overestimate or underestimate the number of objects they encounter. To underscore the arbitrary nature of numerical estimation style, we told participants that approximately half the population are overestimators and half are underestimators, and that research has not related numerical estimation style to any other cognitive tendency or personality trait.

Participants were told they would categorize photographs of students from a previous semester whose numerical estimation style had been determined with a well-established task called the *Numerical Estimation Style Test.* They were then informed that past research had shown that people are able to reliably detect numerical estimation style from faces and that the purpose of the current study was to test whether people can determine numerical estimation style even when face images appear blurry. The instructions explained that the study was important because often people need to make judgments about others from far distances or at night, when perceptual information is not completely clear. It was crucial that participants accepted this aspect of the cover story, because the reverse correlation procedure requires that visually noisy images are used during the face categorization task.

Next, participants were told that in order to gain a concrete understanding of numerical estimation style and how it is measured, they should first complete the NEST themselves. In brief, the task consisted of estimating the number of dots in 10

rapidly presented dot patterns. At the end of the test, the computer program provided predetermined feedback (counterbalanced across participants), indicating that each participant was either an overestimator or an underestimator. The NEST was not actually used to assess any perceptual tendency; it simply provided a rationale for the manipulation of group assignment.

Several procedures were used to ensure that the novel group category was salient in participants' minds throughout the entire face categorization task. First, participants were prompted to report their numerical estimation style to the experimenter, which served as a public act of commitment to the group. The experimenter then wrote each participant's identification number and numerical estimation style in large letters on the cover page of a post-task questionnaire packet and placed the packet in the participants' line of sight to remind them of their group membership during the categorization task. Participants also typed their numerical estimation style group into the computer, in another behavioral act of association. These procedures served to keep participants' minimal group membership salient during the critical categorization task.

Face categorization. Following the NEST, participants completed a forced-choice face categorization task for 450 trials. On each trial, participants viewed two adjacent face images, each consisting of 211×270 grayscale pixels. Participants were told that on each trial one of the faces was an overestimator and the other was an underestimator, and that all the people whose faces they were viewing had completed the numerical estimation style test during a previous semester. According to a counterbalanced schedule, some participants were asked on every trial to choose which of the two faces was an overestimator and the other participants were asked on every trial to select the underestimator face. Thus, targets were ingroup members when their numerical estimation style was shared with the participant. Conversely, targets were outgroup members when their numerical estimation style was different from that of the participant.

Each face was actually derived from the same *base face* image: the grayscale neutral male average face of the Averaged Karolinska Directed Emotional Faces Database (Lundqvist & Litton, 1998). Noise patterns were layered on the images, to make each face look unique, distorting the various facial features and overall facial structure. The noise pattern added to each image consisted of 4,092 superimposed truncated sinusoid patches spanning two cycles in six orientations (0°, 30°, 60°, 90°, 120°, and 150°) $\times$ five spatial frequencies (1, 2, 4, 8, and 16 patches per image) $\times$ two phases (0, $\pi/2$), with random contrasts (amplitudes) as parameters.

On each trial, a random noise pattern was generated. This noise pattern was applied to one of the images in the pair, and the inverse of the noise pattern was added to the other image (see Figure 1). The image with the inverse noise was equally presented on the left and right sides of the screen in a random order. The same noise patterns were used for all participants.

It is important to note that because a response was required on each trial and participants were tasked with selecting on the basis of only one group (either overestimators or underestimators) for all the trials, an equal number of trials were used to generate visualizations of ingroup and outgroup representations
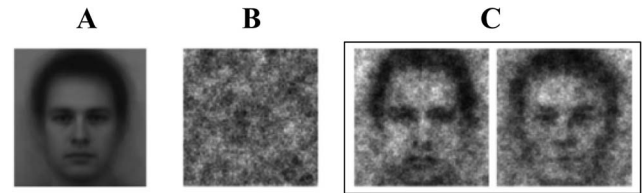
Figure 1. Example stimuli used in the Study 1 face categorization task. (A) Base face, (B) random noise pattern, (C, left) noise pattern superimposed on the face, and (C, right) inverse of the noise pattern superimposed on the base image.

(i.e., effects were not biased by the well-established tendency to exclude ambiguous individuals from the ingroup; Castano, Yzerbyt, Bourguignon, & Seron, 2002; Krosch, Berntsen, Amodio, Jost, & Van Bavel, 2013; Leyens & Yzerbyt, 1992; Quanty, Keats, & Harkins, 1975).

Face representation data processing. The logic of the reverse correlation method is that, on each trial, participants solve the task by comparing the two faces presented on the screen with their mental representation of ingroup or outgroup members, as defined by numerical estimation style. They presumably select whichever face best matches this mental representation. In order to create clear visualizations of social category representations, several hundred of these forced choice categorizations are averaged to form a classification image (Dotsch & Todorov, 2012; Dotsch et al., 2008, 2011; Imhoff et al., 2011). It is notable that if a participant categorized faces randomly, an average of the participant's responses would produce the base face image, because the counterbalanced visual noise would be canceled out (assuming enough trials are included in the average). If participants respond on some systematic basis (e.g., group membership), then systematic patterns in the pixel intensities will emerge that will reveal the mental representation. It is important to note that the visualization of the mental representation is dependent on the base face and type of noise that is used. Additionally, because there are limitations to the number of trials that can be presented, not all the noise will cancel out. However, averaging across participants provides adequate noise reduction for visualization.

Participant-level classification images. Custom MATLAB scripts were used to conduct the reverse correlation analysis. First, for each participant, a classification image was created by averaging together all of the parameters of the 450 noise patterns that he or she selected and superimposing the normalized average on the original base image. The images reflected participants' mental representation of faces belonging to the ingroup or outgroup, as a function of the counterbalanced numerical estimation style factor, the base face, group-specific features, and noise (i.e., error variance).

Group-level classification images. Because our theoretical interest concerned the effects of minimal group membership, and not numerical estimation style, ingroup ($n = 86$) and outgroup ($n = 90$) classification images were created by averaging the appropriate participant-level mean parameters and superimposing the normalized average on the original base

image (collapsing across the numerical estimation style dimension; see Figure 2).[1]

Part 2: Assessing impressions of the ingroup and outgroup face representations. Having implemented the minimal group manipulation in Part 1, which produced visual images representing minimal ingroup and outgroup members, we tested in Part 2 whether ingroup face representations elicited more favorable impressions than outgroup representations, as hypothesized. To this end, an independent sample was recruited to measure trait judgments linked to the ingroup and outgroup face representations estimated during Part 1.

Participants. One hundred and nine participants were recruited through the Amazon Mechanical Turk website (www.mturk.com) to complete an online survey administered through Qualtrics (www.qualtrics.com). Participants were given 30 minutes to complete the study and were remunerated with $0.20. On average, participants finished the survey in approximately six minutes. Data from three participants were removed because, based on their Internet Protocol addresses, it appeared they had already completed the survey. Data from five other participants were excluded because their surveys were incomplete. The final data set included 101 participants. Past research suggests that the quality and reliability of data collected on Mechanical Turk are comparable to those of data collected in laboratory settings (Buhrmester, Kwang, & Gosling, 2011; Paolacci, Chandler, & Ipeirotis, 2010). All participants were naive to how the faces were generated—no mention was made of numerical estimation or any aspect of group membership.

Procedure. After providing informed consent, participants viewed the group-level ingroup and outgroup classification images and rated each on 13 trait dimensions (*trustworthy, attractive, dominant, caring, sociable, confident, emotionally stable, responsible, intelligent, aggressive, mean, weird,* and *unhappy*; Oosterhof & Todorov, 2008). The two group-level face images were presented adjacent to each other in the upper half of the screen, in counterbalanced positions, in order to draw attention to the comparison of the two images. Ratings were made on a 7-point scale. The order of the trait ratings was randomly determined by the Qualtrics stimulus presentation software.

Results

For each trait dimension, we conducted paired *t* tests comparing the ratings of the ingroup and outgroup classification images. The

means, *t* values, and *p* values for each comparison are presented in Table 1. These tests revealed a striking pattern of results, such that the ingroup face was rated substantially higher than the outgroup face on all the traits considered to be desirable of an ingroup member (*attractiveness, intelligence, responsibility, confidence, trustworthiness, caring, emotional stability,* and *sociality*). By contrast, the outgroup face representation was judged as significantly more *weird* than the ingroup face. Ingroup and outgroup face ratings did not differ statistically on *dominance, aggressiveness, meanness,* and *unhappiness*.

Discussion

In Study 1, we investigated whether mere group membership can give rise to contrasting representations of ingroup and outgroup faces, and whether these can then lead to different social impressions. Our results indicate that sharing a minimal group membership with another person can indeed bias facial representations, in a manner that evokes more favorable impressions on a range of traits. Minimally defined ingroup and outgroup faces differed primarily on traits dimensions that signal whether a target should be approached or avoided, but they did not differ on traits found to signify dominance (Oosterhof & Todorov, 2008). This pattern comports with the broader literature on minimal group effects, whereby mere categorization evokes an ingroup preference but not necessarily competition or dominance hierarchies (Mummendey & Otten, 1998). Although past research has shown that people rely on appearance-related stereotypes when forming impressions (Deaux & Lewis, 1984; Kessler & McKenna, 1978; McArthur & Baron, 1983) and that information about a group member's behavior can trigger assumptions about appearance (Dotsch, Wigboldus, & van Knippenberg, 2013), this study is the first to suggest that even seemingly inconsequential group memberships can evoke facial representations that facilitate more favorable impressions of ingroup than outgroup members.

The use of a reverse correlation task was a particularly notable feature of the study because mental representations of faces are difficult to assess with traditional behavioral techniques. Because the reverse correlation approach, used in Part 1, allows all the features of the representation to vary freely and thus does not make strong assumptions a priori about the informational content of the face, it is able to generate estimates of participants' facial representations unbiased by the experimenters' preconceptions. Thus, reverse correlation methodology allowed us to create purely data-driven visualizations of ingroup and outgroup face representations. In Part 2, we then objectively ascertained the impressions elicited by ingroup and outgroup facial representations by asking a naive sample to rate the faces on traits that are highly relevant to face-based evaluations (Oosterhof & Todorov, 2008). It was important that we separated the image generation process from the impression formation stage, because without doing so it would have been difficult to mechanistically show that the facial representations can directly contribute to trait inferences.

Ingroup   **Outgroup**

Figure 2. Group-level ingroup and outgroup classification images from Study 1.

---

[1] The ingroup image was averaged across 44 overestimator and 42 underestimator participant-level classification images. The outgroup image was averaged across 42 overestimator and 48 underestimator images.
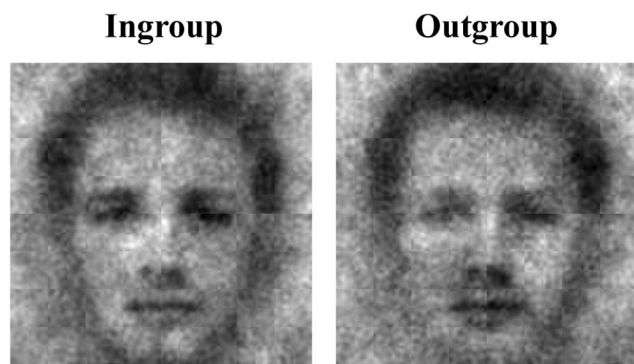
Table 1

Means and Paired t-Test Results for the Study 1 Trait Ratings

Trait	Ingroup mean [95% CI]	Outgroup mean [95% CI]	t	p
Attractive	**4.85 [4.57, 5.13]**	**3.73 [3.42, 4.05]**	**6.17**	**<.001**
Intelligent	**4.96 [4.70, 5.18]**	**4.19 [3.93, 4.46]**	**5.17**	**<.001**
Responsible	**4.69 [4.42, 4.96]**	**3.92 [3.67, 4.19]**	**4.65**	**<.001**
Confident	**4.78 [4.54, 5.03]**	**4.09 [3.81, 4.37]**	**4.17**	**<.001**
Trustworthy	**4.73 [4.48, 4.99]**	**3.96 [3.68, 4.24]**	**3.89**	**<.001**
Caring	**4.50 [4.23, 4.76]**	**3.91 [3.64, 4.18]**	**3.44**	**.001**
Emotionally stable	**4.58 [4.33, 4.84]**	**4.12 [3.87, 4.37]**	**2.59**	**.01**
Sociable	**4.53 [4.28, 4.79]**	**4.15 [3.88, 4.42]**	**2.16**	**.03**
Mean	4.32 [4.02, 4.61]	4.06 [3.77, 4.36]	1.35	.18
Dominant	4.26 [3.97, 4.55]	4.08 [3.78, 4.38]	0.95	.34
Aggressive	3.87 [3.55, 4.15]	4.03 [3.72, 4.34]	−0.80	.43
Unhappy	4.00 [3.67, 4.33]	4.29 [3.96, 4.60]	−1.40	.17
Weird	**3.86 [3.31, 3.86]**	**3.96 [3.68, 4.24]**	**−2.33**	**.02**

Note. Items are ordered from the largest positive *t* value to the most negative. Negative *t* values indicate that the means were greater for the outgroup than the ingroup. Bold type indicates items with significant effects. CI = confidence interval.

Although the reverse correlation approach provided a powerful method for probing the role of facial representations in the minimal group effect, some limitations of this approach should be considered. First, it is unclear whether this representational process occurs spontaneously and in all situations. In all likelihood, there are many situations in which visual representations are not generated or in which group-based preferences are not driven by such representations. Thus, the visual biases identified in Study 1 may represent one of multiple processes that may contribute to minimal group effects. Additionally, it is possible that participants in Part 1 of Study 1 spontaneously generated trait inferences and attitudes that influenced their mental representations of the minimal groups, suggesting a more dynamic interplay of impressions and visual representations. Nevertheless, our experiments showed that once these facial representations are formed, the differences in these representations are themselves sufficient to evoke intergroup bias in trait impressions.

Another potential limitation concerns the effect of the visual clarity of the images produced by the reverse correlation method. For example, it is possible that differences in the clarity of the ingroup and outgroup images contributed to the observed effects, beyond structural differences between these facial images. However, such an effect is difficult to discern from our data because the images used in this study were aggregated across participants. The clarity of the aggregated images could result from different across-group variability for ingroup and outgroup faces. It is additionally possible that image normalization of the noise influenced the pixel intensities across participants in a nonlinear manner complicating the interpretation of clarity differences in the images. Nonetheless, Study 1 served as an initial proof-of-principle that mere group categorization can distort mental representations of faces in a manner that leads to more favorable impressions of ingroup than outgroup members.

Study 2

Study 1 provided preliminary evidence that mental representations of faces elicited by minimally defined ingroup and outgroup

members can vary and lead to trait impressions that favor the ingroup. The ingroup favoritism evident in these trait ratings suggests that differences in mental representations, based on minimal group membership, could also contribute to prejudiced attitudes (i.e., evaluative positions or associations that are not tied to a particular trait attribute; Amodio & Devine, 2006; Allport, 1954). Therefore, in Study 2, we tested whether the face representations computed in Study 1 on the basis of mere group assignment also elicit ingroup favoritism in attitudes. The strategy of Study 2 was similar to that of Study 1, except that instead of asking an independent sample to explicitly rate the face images on trait dimensions, we recruited an independent sample to complete an implicit measure of attitudes toward the ingroup and outgroup images generated in Part 1 of Study 1.

Additionally, instead of presenting participants with the aggregate ingroup and outgroup classification images, as in Study 1, we used individual participant-level classification images as stimuli in Study 2. Use of the participant-level images allowed for repeated measurements of attitudes associated with the ingroup and outgroup representations, which increased the likelihood that the attitude measure would provide stable effects. The use of participant-level images also precluded potential effects of face variability, which could affect clarity in group-level images, thereby addressing a potential limitation of Study 1.

Finally, given that social attitudes may be expressed without conscious deliberation (Devine, 1989; Dovidio, Evans, & Tyler, 1986; Fazio, Jackson, Dunton, & Williams, 1995), use of a measure that implicitly assesses attitudes provided an opportunity to test whether the minimal group effects on face representations can elicit pro-ingroup attitudes at an implicit level of processing.

Method

Participants. One hundred and one undergraduates from New York University participated in exchange for course credit and were run simultaneously in individual cubicles in groups of up to four participants.

Procedure and materials. After providing consent, participants completed an *Affect Misattribution Procedure* (AMP; Payne, Cheng, Govorun, & Stewart, 2005), an implicit measure of attitudes. Past work has validated the AMP's ability to implicitly assess evaluative associations, including those related to racial prejudice, addiction cravings, political ideology, and evocative images (Payne et al., 2005; Payne, McClernon, & Dobbins, 2007). The logic of the AMP is that if people are asked to judge an ambiguous target, their evaluation of this target will be biased by the attitudinal information associated with an image that directly preceded that target.

Before they began the AMP, participants were told that the purpose of the study was to measure their responses under distracting conditions and that, on each trial, two images would appear in sequence. The first image would be a grayscale picture and the second image would be a Chinese character. Participants were instructed to do nothing in response to the first image but to respond to the Chinese character by guessing how pleasant or unpleasant the symbol appeared on a scale of 1 to 4 (with 4 being very pleasant). Participants were also told that they should not let the first picture influence their judgment of the symbol. They then completed three practice trials, followed by 176 experimental

trials, which used each ingroup and outgroup face image generated by the participants from Part 1 of Study 1 as primes. Trial order was determined randomly according to the DirectRT stimulus presentation software. Additionally, the Chinese characters were randomly chosen on each trial from a set of 200 characters that was downloaded from http://www.unc.edu/~bkpayne/materials.html.

Each trial began with a face (ingroup or outgroup) presented for 75 ms, followed by a blank screen for 125 ms, and then a Chinese character for 100 ms. The character was replaced with a mask consisting of grayscale visual noise, which remained onscreen until a response was registered (following Payne et al., 2005). At the conclusion of the AMP, participants were carefully debriefed. As part of this debriefing, participants indicated whether they could read Chinese characters. Eight participants reported being able to read these characters; their data were excluded from analysis, leaving data from 93 participants.

Results

We predicted that if faces representing minimal ingroup members trigger more favorable attitudes than those representing outgroup members, Chinese characters should be rated as more pleasant when preceded by an ingroup than an outgroup classification image. To test our prediction, we computed the average pleasantness rating associated with each participant-level classification image. We then conducted a paired t test to examine differences in the pleasantness ratings associated with the ingroup and outgroup images. Consistent with our hypothesis, Chinese characters that followed the ingroup faces ($M = 2.59$, $SD = 0.32$) were rated more positively than those that followed the outgroup faces ($M = 2.56$, $SD = 0.32$), $t(92) = 2.33$, $p = .02$, $d = .09$, 95% CI of difference [0.004, 0.056].[2]

Discussion

In Study 2, we tested whether differences in mental representations of mere ingroup and outgroup members are sufficient to create pro-ingroup attitudes. The results revealed that the visual renderings of ingroup faces did indeed evoke a more positive attitude than those of outgroup faces, suggesting that subtle differences in the way people visualize ingroup and outgroup members may be sufficient to produce prejudices that favor the ingroup. Although the effect of minimal group assignment on attitudes has been shown in much previous research, our findings show that this effect could, at least in part, be mediated through mental representations of faces.

Our findings also revealed that attitudes elicited by ingroup and outgroup face representations operate implicitly, complementing the observation of explicit impression effects in Study 1. This result is in line with previous research demonstrating that positive ingroup attitudes following mere group categorization can occur implicitly (Ashburn-Nardo et al., 2001; Otten & Wentura, 1999; Van Bavel & Cunningham, 2009). Furthermore, this finding is consistent with evidence that brief, even nonconscious, presentations of faces can elicit neural responses indicative of affect (de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999; Morris, de Gelder, Weiskrantz, & Dolan, 2001; Whalen et al., 1998; see also Amodio, Harmon-Jones, & Devine, 2003).

From a methodological standpoint, it is notable that the results of Study 2 reflected responses to the participant-level images, as opposed to the two group-level images used in Study 1. This finding addresses the concern regarding variability effects on group-level aggregate images and provides converging evidence for the validity of the procedure (as in Dotsch et al., 2008). Furthermore, it is notable that the standard deviations of the AMP responses to ingroup and outgroup participant-level images were close to equivalent. Therefore, it is unlikely that differential variability between the ingroup and outgroup images could explain the observed difference in attitudes.

Study 3

Having demonstrated that mere group membership alters mental representations of ingroup and outgroup faces in a way that elicits differential trait impressions and attitudes, we aimed in Study 3 to assess the degree to which mental representations of ingroup and outgroup members can influence behavior. Here, we focused on trust behavior, because trust is a central dimension for evaluating faces (Oosterhof & Todorov, 2008) and a critical facilitator of harmonious interactions among ingroup members (Foddy, Platow, & Yamagishi, 2009; Kramer, 1999; Tanis & Postmes, 2005). To assess the degree to which trust is evoked by ingroup facial representations, in comparison with outgroup representations, we conducted a study in which participants completed an economic trust game with interaction partners represented by the ingroup and outgroup classification images from Study 1 (Berg, Dickhaut, & McCabe, 1995). As in Study 2, participants responded to the participant-level classification images.

Method

Participants. Eighty-one undergraduates from New York University completed the study in exchange for course credit. Up to four participants were run simultaneously in separate cubicles.

Procedure.

Trust game. After providing informed consent, participants completed a hypothetical trust game with people depicted by the ingroup and outgroup face classification images. Participants were instructed to imagine that they were given $10 on each trial, and that they could choose either to keep this money or to take the chance to increase their share by engaging in a hypothetical economic interaction with individuals who purportedly provided responses in an earlier session. On each interaction trial, participants were given the choice to share a portion of $10 (i.e., $0, $2, $4, $6, $8, or $10). Any money they shared would be quadrupled and given to the interaction partner. The partner would then have the option to return half of the sum to the participant who had

---

[2] We reported the results for a paired t test collapsing across the numerical estimation style dimension because we were interested in group differences. However, because our analyses for this study used the participant-level images, and we thus had information about the group and numerical estimation style for each image, we were also able to conduct a 2 (group: ingroup vs. outgroup) × 2 (numerical estimation style: overestimator vs. underestimator) repeated measures analysis of variance (ANOVA). Consistent with the results of our paired t test, this analysis produced a main effect of group, such that, irrespective of numerical estimation style, Chinese characters that followed the ingroup faces ($M = 2.59$, $SD = 0.32$) were rated more positively than those that followed the outgroup faces ($M = 2.56$, $SD = 0.32$), $F(1, 92) = 5.33$, $p = .02$, $\eta_p^2 = .06$. There were no numerical estimation style or interaction effects ($ps < .31$).

shared the money. In this way, it would be possible for the participant to make more money than if he or she had not shared. Hence, the amount of money shared indicates the degree to which the participants trusted the interaction partners.

Participants played the trust game with 176 different partners, represented only by a photograph. In actuality, these photographs were the participant-level classification images of the ingroup and outgroup faces, presented in a random order across trials. Participants were told that each partner's face was blurred to protect his or her identity.

Results

We predicted that participants would entrust more money to individuals depicted by the ingroup classification images than to those depicted by the outgroup classification images. As a test of this, the average amount entrusted to ingroup faces and outgroup faces was computed for each participant and submitted to a paired t test. This analysis indicated that ingroup face depictions ($M =$ 3.01, $SD = 1.62$) were trusted significantly more than outgroup face depictions ($M = 2.77, $SD = 1.62$), $t(80) = 6.64$, $p < .001$, $d = .15$, 95% CI of difference [0.17, 0.30].[3]

Discussion

In Study 3, we tested the hypothesis that representations of ingroup and outgroup faces formed on the basis of a minimal group distinction could elicit different patterns of behavior, in the form of trust decisions in an economic game. We found that participants acted in a more trusting manner toward faces reflecting mental representations of ingroup members, as compared with outgroup members, evidencing a pattern of ingroup favoritism in behavior. As in Study 2, the variance in responses to participant-level ingroup and outgroup faces did not differ, further suggesting that the effects were not driven by differential variability in representations of ingroup and outgroup faces. Thus, this study suggests that subtle biases in the way an individual mentally represents the face of an ingroup or outgroup member could result in a difference in trust decisions.

The results of Study 3 add to previous research on trust behavior under minimal group conditions by extracting the effect of minimal group assignment on facial representations and then showing that this information alone can cause differences in trust behavior. In past research (Foddy et al., 2009; Platow, Foddy, Yamagishi, Lim, & Chow, 2012), minimal group effects were examined in the absence of group-specific facial information. Given that ingroup favoritism, especially in zero-sum situations, can result in disadvantaging outgroup members, this work suggests a novel pathway through which group-based facial representations can contribute to prejudice, discrimination, and intergroup conflict.

It is important to mention that our conclusions are drawn from trust behavior during hypothetical transactions, in the sense that participants could not actually earn money during the game. Differences between hypothetical and real scenarios tend to occur because people risk more money and exhibit self-presentation concerns more strongly during hypothetical situations (Ajzen, Brown, & Carvajal, 2004). Although we cannot conclusively rule out the possibility that participants' behavior would have differed if real money had been exchanged, we believe it is unlikely. First,

our procedure should have precluded any self-presentation concerns related to ingroup and outgroup membership, given that participants were unaware of the group-based origins of the face images. Second, although it is possible that participants would have risked less money in their decisions if the money had been real, our interest was in the relative difference in money entrusted to ingroup and outgroup representations rather than the mean amount.

Study 4

The aim in Study 4 was to examine, more directly, how facial cues signaling trust contribute to the observed effects of group membership. Because the reverse correlation procedure used to render the images in Study 1 allows for unconstrained representations of facial features, it is sensitive to the many possible ways in which group membership could be instantiated in the physiognomy of a face. For instance, it is possible that trust was conveyed primarily in the eyes or by the shape of the mouth (Schul, Mayo, & Burnstein, 2004; Zebrowitz, 1997). Thus, our goal was to objectively test the correspondence between ingroup (vs. outgroup) face representations and an independently produced representation of a trustworthy face. Insight into this physiognomic correspondence would clarify the mechanism underlying the results of Studies 1–3.

A second, related goal was to examine the extent to which the resemblance between trust and ingroup representations is distributed across the face. This analysis would shed light on whether the ingroup favoritism effects observed thus far are due to specific or gestalt differences between the ingroup and outgroup face representations.

To address these goals in Study 4, we recruited a new sample of participants to complete a face categorization task. Although similar to that used in Study 1, it pertained to whether a face appears trustworthy. The reverse correlation analysis method was then used to produce visual renderings of participants' mental images of a trustworthy face. Because people have a relatively strong notion of what a trustworthy face looks like (Dotsch & Todorov, 2012), a smaller sample was needed to generate a clear estimate of

---

[3] As with Study 2, the analyses for Study 3 used the participant-level images. Therefore, we were also able to conduct a 2 (group: ingroup vs. outgroup) × 2 (numerical estimation style: overestimator vs. underestimator) repeated measures ANOVA. In line with the results from the paired t test, this analysis produced a significant main effect for group, such that ingroup face depictions ($M = 3.01$, $SD = 1.62$) were trusted significantly more than outgroup face depictions ($M = 2.77$, $SD = 1.62$), $F(1, 80) =$ 44.02, $p < .001$, $\eta_p^2 = .36$. Of interest, and unexpectedly, a significant main effect also emerged for numerical estimation style, indicating that overestimators ($M = 3.01$, $SD = 1.61$) were generally trusted significantly more than underestimators ($M = 2.76$, $SD = 1.63$), $F(1, 80) =$ 37.28, $p < .001$, $\eta_p^2 = .32$. These main effects were qualified by a significant interaction, $F(1, 80) = 10.46$, $p < .01$, $\eta_p^2 = .12$, which revealed that the group effect was larger in response to faces representing underestimators ($M_{(ingroup - outgroup)} = .33$, $SD_{(ingroup - outgroup)} = $ $.40$) than to faces representing overestimators ($M_{(ingroup-outgroup)} = .11$, $SD_{(ingroup-outgroup)} = .46$), $t(80) = 3.24$, $p = .002$, $d = .51$. It is notable, however, that the ingroup was trusted significantly more than the outgroup for both the overestimators, $t(80) = 2.14$, $p = .04$, $d = .07$, and the underestimators, $t(80) = 7.56$, $p < .001$, $d = .20$. Thus, our hypothesis about the effect of group membership was supported independently of the effect observed for the numerical estimation style factor.

physiognomic information associated with trust. We then compared the resulting classification image of a *trustworthy* face (collapsed across subjects) to the group-level classification images from Study 1 that were used to estimate ingroup and outgroup face representations.

Method

Participants. Fourteen undergraduates from New York University participated in exchange for extra course credit and completed the study in private cubicles, in groups of one to four participants at a time.

Procedure. The procedure and stimuli were identical to those in Study 1, except that there was no group assignment and, thus, no minimal group paradigm was used. Instead, during the face categorization task, participants were instructed to choose the face from each trial pair *that looks the most trustworthy*. As was predicted, a visual inspection of the aggregate trust classification image (see Figure 3A) indicated that a sample size of 14 participants was sufficient to generate a clear estimate of participants' mental image of a trustworthy face. This image closely resembled the trustworthy face classification image generated by Dotsch and Todorov (2012).

Data reduction. Participant-level classification images were first created by averaging the 450 faces selected as appearing trustworthy for each participant. A composite mental representation of a trustworthy face was then created by aggregating these participant-level classification images. The classification image depicting the trust representation is presented alongside the Study 1 ingroup and outgroup group-level images in Figure 3A.

To quantitatively test similarities between the trust- and group-based representations, we correlated the patterns of pixel intensities that composed these classification images. Before calculating the correlations, we removed the base face from all the images, leaving only pixel patterns representing the variation due to participants' mental images, as inclusion of the base face would have artificially inflated observed correlations between the images. Regions of the image outside the face, including the hair, were masked and excluded from analysis (see Figure 3B to see the masked noise patterns). Finally, for each image, the intensity

Figure 3. (A) Ingroup, outgroup, and trustworthy classification images. (B) Masked classification images with the base face removed used during the Study 4 analyses.

values of the remaining 89,177 pixels were converted into a single vector and correlated with the other two vectors.

Results

Correlations between the trust and ingroup/outgroup images. We hypothesized that the pixel intensities of the trust classification image would be more highly correlated with the ingroup than the outgroup classification image, indicating a greater correspondence between the facial representations of trustworthiness and the ingroup face, as compared with the outgroup face. Supporting our prediction, the vectors representing pixel intensities for the ingroup and trust classification images were significantly correlated ($r = .46$, $p < .001$). The outgroup and trust classification images were also significantly correlated ($r = .23$, $p < .001$), but the magnitude of this correlation was significantly lower than that of the trust–ingroup correlation ($z = 55.56$, $p < .001$). This test confirmed that the resemblance of the ingroup and trust classification images was significantly greater than that of the outgroup and trust classifications.

Given the large sample size (i.e., number of pixels), the significance of the p value is not as notable as the fact that the ingroup classification image accounted for over four times as much variance in the trust classification image as did the outgroup classification image ($r^2 = .21$, 95% CI [0.205, 0.215] vs. $r^2 = .05$, 95% CI [0.047, 0.052]). Moreover, if partial correlations are considered to adjust for shared variance between the ingroup and outgroup images, the variance accounted for in the trust image is nine times as large for the ingroup image ($r^2 = .18$) as it is for the outgroup image ($r^2 = .02$). These results objectively demonstrate that the facial information associated with people's representation of an ingroup member overlaps more with a mental representation of a trust face than does their representation of an outgroup member.

Region of interest analyses. To assess whether the differences in the correspondences between the ingroup and outgroup faces with the trustworthy face representation were localized or distributed across the face, we conducted the same partial correlation analyses on pixels representing noise patterns within regions of interest (ROIs) covering the eye region, nose, and mouth (see Figure 4). The results from these analyses are summarized in Table 2. Of importance, for each of the regions of interest, the similarity between the trust classification image and the ingroup classification image was significantly larger than the trust classification image and the outgroup classification image. Consistent with research demonstrating that the eyes are especially important for trust judgments (Schul et al., 2004; Zebrowitz, 1997), the differences in correlations for the ingroup and outgroup were largest for the eye region, followed by the nose, and then the mouth.

Discussion

Study 4 examined the hypothesis that differences in facial representations of minimal ingroup and outgroup members could be attributed, at least in part, to the ascription of features associated with trustworthiness to the ingroup face. By comparing patterns in the pixel intensities of the ingroup and outgroup classification images to those of the trust classification image produced by an independent sample, this study provided an especially stringent and objective test of this hypothesis. These results support our
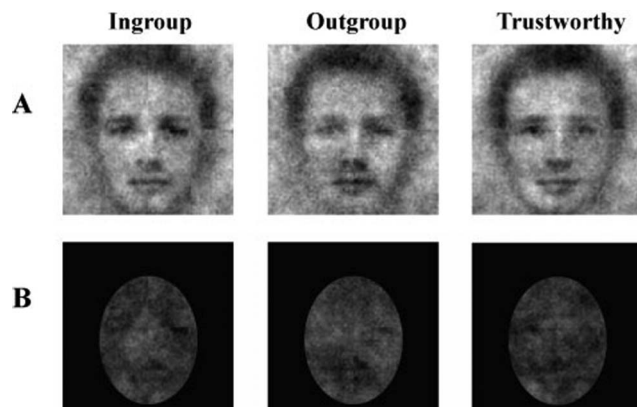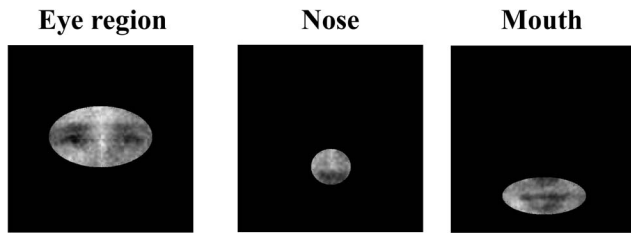
Eye region **Nose** **Mouth**

Figure 4. Masks used to conduct the Study 4 region of interest analyses, shown on the trust classification image. The number of pixels in each region of interest were 36,994 (eye region), 8,269 (nose), and 18,251 (mouth).

theoretical proposal that shared group membership can bias mental representations of faces toward a facial geometry that communicates affiliative signals, which would elicit more favorable impressions, attitudes, and behaviors from perceivers.

It is also noteworthy that the ingroup face representation was positively correlated with the trustworthy face presentation, but the outgroup face representation was not negatively correlated with the trustworthy representation. This is consistent with the existing minimal group literature, which has shown that merely separating people into groups increases ingroup favoritism but does not elicit outgroup derogation (Mummendey & Otten, 1998).

Finally, in Study 4, we also examined the ingroup favoritism effect in three regions of the face (the eyes, nose, and mouth). In all three regions, the trust image was correlated more highly with the ingroup image than the outgroup image. This finding suggests that the ingroup bias in mental representation is distributed across multiple regions of the face or, at the very least, is represented in different regions by different participants. Although trust information was communicated in multiple regions of the ingroup face, the effects were strongest for the eyes. This result is consistent with a growing body of research that suggests that the eyes are attended to when making decisions related to trust, possibly because they are assumed to act as a window into the unobservable mind (Schul et al., 2004; Zebrowitz, 1997).

General Discussion

Over four decades of research has demonstrated that mere group categorization alone is sufficient to cause discriminatory behavior. Simply assigning a person to an arbitrarily defined group produces behavioral and evaluative preferences for the ingroup over the outgroup (e.g., Ashburn-Nardo et al., 2001; Brewer & Silver, 1978; Locksley et al., 1980; Otten & Wentura, 1999; Tajfel et al., 1971; Van Bavel & Cunningham, 2009). In the past several years, numerous studies have investigated how sharing a group identity with others can influence the processing of and memory for their faces. Although this work has revealed that ingroup faces are processed more deeply and accurately than outgroup faces (Bernstein et al., 2007; Hugenberg & Corneille, 2009; Ratner & Amodio, 2013; Van Bavel et al., 2008, 2011; Young & Hugenberg, 2010), the present research investigated a different mechanism through which mere group categorization can influence face processing. That is, we proposed that mere group assignment can distort mental representations of faces in a way that leads to more favorable responses to ingroup than to outgroup members.

This hypothesis was examined across four studies. In Study 1, we showed that membership in a minimally defined group was sufficient to produce differences in the facial representations of ingroup and outgroup members and that, when shown to naive participants, a representation of an ingroup face elicited more favorable trait impressions than a representation of an outgroup face. The ingroup facial representation was judged to be more trustworthy, attractive, caring, sociable, emotionally stable, intelligent, responsible, and confident than the outgroup face representation: a pattern of ingroup traits signifying group fitness and affiliative tendencies. Furthermore, we found that ingroup face representations elicited more positive implicitly measured attitudes (Study 2) and trusting behavior (Study 3) than did outgroup face representations. Finally, Study 4 demonstrated that participants' representation of facial trustworthiness more closely resembled the information contained within an ingroup face representation than an outgroup face representation. Together, these results suggest that people form subtly different mental images of ingroup and outgroup members that can contribute to significant differences in trait impressions, attitudes, and behaviors and that the difference in mental images can be driven, in part, by the ascription of trust-related features to the ingroup face.

Inferring a Causal Influence of Facial Representations on Intergroup Responses

The overarching objective of our work was to study the effects of mere group categorization on facial representations as a means to understand how group memberships can lead to intergroup bias. Although several mechanisms likely contribute to the effect of minimal group assignment on intergroup bias, we examined internal representations of facial information as one causal pathway (cf. Bullock, Green, & Ha, 2010; Green, Ha, & Bullock, 2010). In order to provide inferential traction when multiple mechanisms might exist, methodologists have encouraged experimentally testing each step in a predicted causal path (Spencer et al., 2005). In line with this recommendation, we first manipulated the effect of mere group categorization on facial representations (Study 1, Part 1) and then independently assessed the effect of the resultant facial representations on trait impressions (Study 1, Part 2), attitudes (Study 2), and behavior (Study 3). This approach was built upon the already well-established "direct path" between minimal group induction and ingroup favoritism (e.g., Ashburn-Nardo et al., 2001; Brewer & Silver, 1978; Van Bavel & Cunningham, 2009). The use of this design suggests that facial representations can

Table 2

Study 4 Partial Correlations Between Trust and Ingroup/Outgroup Classification Images as a Function of Face Region

Group	Trust		
	Eyes	Nose	Mouth
Ingroup	0.48	0.55	0.47
Outgroup	−0.09	0.15	0.31

Note. All the correlation coefficients were significant ($p < .001$). Within each region, the two correlation coefficients were different from each other ($p < .001$).

contribute to minimal group effects on attitudes, inferences, and trust decisions.

To be clear, our results suggest that biased facial representations can operate as a mechanism leading to intergroup bias, but, like other mediation-based approaches, they do not rule out a role for other mechanisms at any stage of the process. For instance, our work does not preclude the possibility that participants in Part 1 of Study 1 explicitly generated trait inferences as a means to develop mental representations of the minimal groups. It is also possible that clarity differences in the mental images of ingroup and outgroup members contributed to downstream favoritism effects, although the results of Studies 2 and 3 suggest that this is unlikely to be the case. Some people might also use their self-image or prototypical features of other groups they belong to as a basis for discerning their image of a minimally defined ingroup member (see Gramzow, Gaertner, & Sedikides, 2001; Imhoff & Dotsch, 2013, for research consistent with these possibilities). Further research will be required to fully illuminate the range of facial representation mechanisms that contribute to the effects of minimal group distinctions.

Implications for Mental Simulations Involving Ingroup and Outgroup Members

Our research suggests that facial representations associated with ingroup favoritism may influence social relations. One way this may occur is by guiding mental simulation. Mental simulation refers to people's ability to transcend the here and now by role-playing past and future experiences (Schacter, Addis, & Buckner, 2007; Tulving, 1983). Functionally, mental simulation allows people to plan for and feel emotions associated with scenarios that they are not currently experiencing (Marks, 1999; Taylor, Pham, Rivkin, & Armor, 1998).

An implication for intergroup interactions is that mental representations of an ingroup face could result in self-fulfilling prophesies (Chen & Bargh, 1997; Darley & Fazio, 1980; Merton, 1948). For instance, when anticipating an interaction with an ingroup member, expectations that the target person will look trustworthy, responsible, and sociable could contribute to approach-related attitudes and cooperative behaviors during the interaction, which in turn could increase the likelihood that the ingroup member develops warm attitudes toward the perceiver and reciprocates affiliative behavior. Similar mental simulation effects could also emerge when people are interacting with an ingroup member over e-mail or in other situations that do not provide access to appearance information about the target individual. These self-reinforcing processes could perpetuate good will that promotes positive ingroup relations. Given that outgroup representations elicit less favorable impressions than ingroup representations, outgroup members may be less likely to receive the benefit of the doubt than ingroup members. This would put outgroup members at a further disadvantage.

Implications for Intergroup Face Perception and Self-Regulation

In most research on minimal group effects on face processing, the guiding assumption is that the ingroup is more motivationally relevant, which leads to deeper and more accurate processing of ingroup faces and relatively superficial processing of outgroup faces (Hugenberg & Corneille, 2009; Ratner & Amodio, 2013; Van Bavel et al., 2008, 2011; Young & Hugenberg, 2010). This work has been particularly useful for demonstrating the generality of the *other race effect* (i.e., the finding that same-race faces are remembered better than other-race faces; Malpass & Kravitz, 1969). However, as we have mentioned, in minimal group work outside of the domain of face processing, researchers have suggested that ingroup biases are driven by enhanced positivity toward the ingroup (e.g., Ashburn-Nardo et al., 2001; Brewer & Silver, 1978). Motivated in part by this apparent disconnect in literatures, the current work examined whether people see ingroup members through "rose-colored glasses" and whether these distortions in facial representations contribute to biased intergroup impressions, evaluations, and behaviors. In this way, the perceptual biases promoted by these mental representations provide a mechanism for facilitating one's goal to favor ingroup members (i.e., an example of how motivated perception serves self-regulation; Amodio, 2010).

Our findings also raise the possibility that internally generated representations of ingroup and outgroup faces influence bottom-up processes through which faces are attended to and perceived. Predictive coding theories state that expectations can guide the acquisition of visual information (Friston, 2005; Mumford, 1992; Summerfield et al., 2006). From this perspective, it is possible that if people expect an ingroup member to have certain facial characteristics but the person's appearance diverges from this prediction, this discordant information would receive more processing. As such, formal models could be developed that use visual renderings of ingroup and outgroup face representations as Bayesian priors for predicting the duration and location of endogenous attention and eye saccades allocated to processing attributes of a particular face. This type of an analysis could advance theories of how mere group knowledge biases the perceptual information that reaches the retina and is subsequently encoded in the visual cortex. Moreover, such work may lead to the development of strategies that protect against this type of bias by changing the predictive visual codes that people use to disambiguate facial information of group members.

Conclusion

Although the role of appearance cues in group-based processes had been emphasized in seminal theoretical articles (e.g., Brewer, 1988; Carlston, 1992; McArthur & Baron, 1983), until recently, the bulk of social cognition research overlooked the contributions of facial information (Macrae & Bodenhausen, 2006; Zebrowitz, 2006). Our studies contribute to the recent movement to reintegrate research on visual face processing into the social psychological understanding of intergroup responses (e.g., Dotsch et al., 2008; Hugenberg & Corneille, 2009; Ito & Urland, 2005; Kaul, Ratner, & Van Bavel, 2014; Ofan, Rubin, & Amodio, 2011, in press; Ratner & Amodio, 2013; Ratner, Kaul, & Van Bavel, 2013; Van Bavel et al., 2011).

In this spirit, our work provides evidence that mere group categorization can lead to ingroup favoritism through effects on mental representations of faces. Whereas past research had shown that simply separating people into novel groups triggers biases that favor ingroup members, it had not been previously established that

representations of faces could contribute to this phenomenon. Given that facial representations are utilized to make sense of other people, this research provides insight into a previously underexplored route through which group-based biases can influence society.

References

Ahumada, A. J. (2002). Classification image weights and internal noise level estimation. *Journal of Vision, 2,* 121–131. doi:10.1167/2.1.8

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

Ajzen, I., Brown, T. C., & Carvajal, F. (2004). Explaining the discrepancy between intentions and actions: The case of hypothetical bias in contingent valuation. *Personality and Social Psychology Bulletin, 30,* 1108–1121. doi:10.1177/0146167204264079

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.

Amodio, D. M. (2010). Coordinated roles of motivation and perception in the regulation of intergroup responses: Frontal cortical asymmetry effects on the P2 event-related potential and behavior. *Journal of Cognitive Neuroscience, 22,* 2609–2617. doi:10.1162/jocn.2009.21395

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology, 91,* 652–661. doi:10.1037/0022-3514.91.4.652

Amodio, D. M., Harmon-Jones, E., & Devine, P. G. (2003). Individual differences in the activation and control of affective race bias as assessed by startle eyeblink responses and self-report. *Journal of Personality and Social Psychology, 84,* 738–753. doi:10.1037/0022-3514.84.4.738

Andersen, S. M., & Klatzky, R. L. (1987). Traits and social stereotypes: Levels of categorization in person perception. *Journal of Personality and Social Psychology, 53,* 235–246. doi:10.1037/0022-3514.53.2.235

Ashburn-Nardo, L., Voils, C. I., & Monteith, M. J. (2001). Implicit associations as the seeds of intergroup bias: How easily do they take root? *Journal of Personality and Social Psychology, 81,* 789–799. doi:10.1037/0022-3514.81.5.789

Beaupré, M. G., & Hess, U. (2003). In my mind, my friend smiles: A case of in-group favoritism. *Journal of Experimental Social Psychology, 39,* 371–377. doi:10.1016/S0022-1031(03)00012-X

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8,* 551–565. doi:10.1162/jocn.1996.8.6.551

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

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychological Science, 18,* 706–712. doi:10.1111/j.1467-9280.2007.01964.x

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009a). Get real! Resolving the debate about equivalent social stimuli. *Visual Cognition, 17,* 904–924. doi:10.1080/13506280902758044

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009b). Saliency does not account for fixations to eyes within social scenes. *Vision Research, 49,* 2992–3000. doi:10.1016/j.visres.2009.09.014

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15,* 674–679. doi:10.1111/j.0956-7976.2004.00739.x

Blair, I. V., Ma, J. E., & Lenton, A. P. (2001). Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery. *Journal of Personality and Social Psychology, 81,* 828–841. doi:10.1037/0022-3514.81.5.828

Bodenhausen, G. V., & Macrae, C. N. (2006). Putting a face on person perception. *Social Cognition, 24,* 511–515.

Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull & R. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55,* 429–444. doi:10.1111/0022-4537.00126

Brewer, M. B., & Silver, M. (1978). Ingroup bias as a function of task characteristics. *European Journal of Social Psychology, 8,* 393–400. doi:10.1002/ejsp.2420080312

Brown, J. D., Collins, R. L., & Schmidt, G. W. (1988). Self-esteem and direct versus indirect forms of self-enhancement. *Journal of Personality and Social Psychology, 55,* 445–453. doi:10.1037/0022-3514.55.3.445

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3–5. doi:10.1177/1745691610393980

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology, 98,* 550–558. doi:10.1037/a0018933

Carlston, D. E. (1992). Impression formation and the modular mind: The associated systems theory. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 301–341). Hillsdale, NJ: Erlbaum.

Castano, E., Yzerbyt, V., Bourguignon, D., & Seron, E. (2002). Who may enter? The impact of in-group identification on in-group/out-group categorization. *Journal of Experimental Social Psychology, 38,* 315–322. doi:10.1006/jesp.2001.1512

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self-fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541–560. doi:10.1006/jesp.1997.1329

Crisp, R. J., & Turner, R. N. (2009). Can imagined interactions produce positive perceptions? Reducing prejudice through simulated social contact. *American Psychologist, 64,* 231–240. doi:10.1037/a0014718

Darley, J. M., & Fazio, R. H. (1980). Expectancy confirmation processes arising in the social interaction sequence. *American Psychologist, 35,* 867–881. doi:10.1037/0003-066X.35.10.867

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology, 46,* 991–1004. doi:10.1037/0022-3514.46.5.991

de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Nonconscious recognition of affect in the absence of striate cortex. *NeuroReport, 10,* 3759–3763. doi:10.1097/00001756-199912160-00007

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18. doi:10.1037/0022-3514.56.1.5

Doise, W., Csepeli, G., Dann, H. D., Gouge, C., Larsen, K., & Ostell, A. (1972). An experimental investigation into the formation of intergroup representations. *European Journal of Social Psychology, 2,* 202–204. doi:10.1002/ejsp.2420020208

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological & Personality Science, 3,* 562–571. doi: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,* 978–980. doi:10.1111/j.1467-9280.2008.02186.x

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

Dotsch, R., Wigboldus, D. H. J., & van Knippenberg, A. (2013). Behavioral information biases the expected facial appearance of members of

novel groups. *European Journal of Social Psychology, 43,* 116–125. doi:10.1002/ejsp.1928

Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology, 22,* 22–37. doi:10.1016/0022-1031(86)90039-9

Dunham, Y. (2011). An angry = outgroup effect. *Journal of Experimental Social Psychology, 47,* 668–671. doi:10.1016/j.jesp.2011.01.003

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17,* 383–386. doi:10.1111/j.1467-9280.2006.01716.x

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobstrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69,* 1013–1027. doi:10.1037/0022-3514.69.6.1013

Foddy, M., Platow, M. J., & Yamagishi, T. (2009). Group-based trust in strangers. *Psychological Science, 20,* 419–422. doi:10.1111/j.1467-9280.2009.02312.x

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360,* 815–836. doi:10.1098/rstb.2005.1622

Gerard, H. B., & Hoyt, M. F. (1974). Distinctiveness of social categorization and attitude toward ingroup members. *Journal of Personality and Social Psychology, 29,* 836–842. doi:10.1037/h0036204

Goren, C. C., Sarty, M., & Wu, P. Y. K. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics, 56,* 544–549.

Gosselin, F., & Schyns, P. G. (2003). Superstitious perceptions reveal properties of internal representations. *Psychological Science, 14,* 505–509. doi:10.1111/1467-9280.03452

Gramzow, R. H., Gaertner, L., & Sedikides, C. (2001). Memory for ingroup and outgroup information in a minimal group context: The self as an informational base. *Journal of Personality and Social Psychology, 80,* 188–205. doi:10.1037/0022-3514.80.2.188

Green, D. P., Ha, S. E., & Bullock, J. G. (2010). Enough already about "black box" experiments: Studying mediation is more difficult than most scholars suppose. *Annals of the American Academy of Political and Social Science, 628,* 200–208. doi:10.1177/0002716209351526

Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 78,* 837–852. doi:10.1037/0022-3514.78.5.837

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233. doi:10.1016/S1364-6613(00)01482-0

Hugenberg, K., & Corneille, O. (2009). Holistic processing is tuned for in-group faces. *Cognitive Science, 33,* 1173–1181. doi:10.1111/j.15551-6709.2009.01048.x

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

Imhoff, R., Dotsch, R., Bianchi, M., Banse, R., & Wigboldus, D. H. (2011). Facing Europe: Visualizing spontaneous in-group projection. *Psychological Science, 22,* 1583–1590. doi:10.1177/0956797611419675

Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: Effects of memory and attention revealed by fMRI. *NeuroImage, 17,* 1729–1741. doi:10.1006/nimg.2002.1330

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, and Behavioral Neuroscience, 5,* 21–36.

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General, 141,* 19–25. doi:10.1037/a0023463

Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development: The case of face recognition.* Oxford, England: Blackwell.

Johnson, R. W. (1981). Perceived physical attractiveness of supporters of Canada's political parties: Stereotype or in-group bias? *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 13,* 320–325. doi:10.1037/h0081203

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17,* 4302–4311.

Karremans, J. C., Dotsch, R., & Corneille, O. (2011). Romantic relationship status biases memory of faces of attractive opposite-sex others: Evidence from a reverse-correlation paradigm. *Cognition, 121,* 422–426. doi:10.1016/j.cognition.2011.07.008

Kaul, C., Ratner, K. G., & Van Bavel, J. J. (2014). Dynamic representations of race: Processing goals shape race encoding in the fusiform gyri. *Social Cognitive and Affective Neuroscience, 9,* 326–332. doi:10.1093/scan/nss138

Kessler, S. J., & McKenna, W. (1978). *Gender: An ethnomethodological approach.* New York, NY: Wiley.

Kramer, R. M. (1999). Trust and distrust in organizations: Emerging perspectives, enduring questions. *Annual Review of Psychology, 50,* 569–598. doi:10.1146/annurev.psych.50.1.569

Krosch, A. R., Berntsen, L., Amodio, D. M., Jost, J. T., & Van Bavel, J. J. (2013). On the ideology of hypodescent: Political conservatism predicts categorization of racially ambiguous faces as Black. *Journal of Experimental Social Psychology, 49,* 1196–1203. doi:10.1016/j.jesp.2013.05.009

Leyens, J., & Yzerbyt, V. Y. (1992). The ingroup overexclusion effect: Impact of valence and confirmation on stereotypical information search. *European Journal of Social Psychology, 22,* 549–569. doi:10.1002/ejsp.2420220604

Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology, 82,* 5–18. doi:10.1037/0022-3514.82.1.5

Locksley, A., Ortiz, V., & Hepburn, C. (1980). Social categorization and discriminatory behavior: Extinguishing the minimal intergroup discrimination effect. *Journal of Personality and Social Psychology, 39,* 773–783. doi:10.1037/0022-3514.39.5.773

Lundqvist, D., & Litton, J. E. (1998). *The averaged Karolinska directed emotional faces—AKDEF* [CD-ROM]. Stockholm, Sweden: Psychology Section, Karolinska Institutet.

Ma, D. S., & Correll, J. (2011). Target prototypicality moderates racial bias in the decision to shoot. *Journal of Experimental Social Psychology, 47,* 391–396. doi:10.1016/j.jesp.2010.11.002

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

Macrae, C. N., Quinn, K. A., Mason, M. F., & Quadflieg, S. (2005). Understanding others: The face and person construal. *Journal of Personality and Social Psychology, 89,* 686–695. doi:10.1037/0022-3514.89.5.686

Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology, 13,* 330–334. doi:10.1037/h0028434

Mangini, M. C., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classifications. *Cognitive Science, 28,* 209–226. doi:10.1207/s15516709cog2802_4

Marks, D. F. (1999). Consciousness, mental imagery and action. *British Journal of Psychology, 90,* 567–585. doi:10.1348/000712699161639

Martin-Malivel, J., Mangini, M. C., Fagot, J., & Biederman, I. (2006). Do humans and baboons use the same information when categorizing human and baboon faces? *Psychological Science, 17,* 599–607. doi:10.1111/j.1467-9280.2006.01751.x

McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review, 90,* 215–238. doi:10.1037/0033-295X.90.3.215

Mechelli, A., Price, C. J., Friston, K. J., & Ishai, A. (2004). Where bottom-up meets top-down: Neuronal interactions during perception and imagery. *Cerebral Cortex, 14,* 1256–1265. doi:10.1093/cercor/bhh087

Merton, R. K. (1948). The self-fulfilling prophecy. *Antioch Review, 8,* 193–210. doi:10.2307/4609267

Morris, J. S., de Gelder, B., Weiskrantz, L., & Dolan, R. J. (2001). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain, 124,* 1241–1252. doi:10.1093/brain/124.6.1241

Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics, 66,* 241–251. doi:10.1007/BF00198477

Mummendey, A., & Otten, S. (1998). Positive–negative asymmetry in social discrimination. *European Review of Social Psychology, 9,* 107–143. doi:10.1080/14792779843000063

Mussweiler, T., Gabriel, S., & Bodenhausen, G. V. (2000). Shifting social identities as a strategy for deflecting threatening social comparisons. *Journal of Personality and Social Psychology, 79,* 398–409. doi:10.1037/0022-3514.79.3.398

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12,* 1013–1023. doi:10.1162/08989290051137549

Ofan, R. H., Rubin, N., & Amodio, D. M. (2011). Seeing race: N170 responses to race and their relation to automatic racial attitudes and controlled processing. *Journal of Cognitive Neuroscience, 23,* 3153–3161. doi:10.1162/jocn_a_00014

Ofan, R. H., Rubin, N., & Amodio, D. M. (in press). Situation-based social anxiety enhances the neural processing of faces: Evidence from an intergroup context. *Social Cognitive and Affective Neuroscience.* doi:10.1093/scan/nst087

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

Otten, S., & Wentura, D. (1999). About the impact of automaticity in the minimal group paradigm: Evidence from affective priming tasks. *European Journal of Social Psychology, 29,* 1049–1071. doi:10.1002/(SICI)1099-0992(199912)29:8<1049::AID-EJSP985>3.0.CO;2-Q

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5,* 411–419.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89,* 277–293. doi:10.1037/0022-3514.89.3.277

Payne, B. K., McClernon, F. J., & Dobbins, I. G. (2007). Automatic affective responses to smoking cues. *Experimental and Clinical Psychopharmacology, 15,* 400–409. doi:10.1037/1064-1297.15.4.400

Platow, M. J., Foddy, M., Yamagishi, T., Lim, L., & Chow, A. (2012). Two experimental tests of trust in in-group strangers: The moderating role of common knowledge of group membership. *European Journal of Social Psychology, 42,* 30–35. doi:10.1002/ejsp.852

Quanty, M. B., Keats, J. A., & Harkins, S. G. (1975). Prejudice and criteria for identification of ethnic photographs. *Journal of Personality and Social Psychology, 32,* 449–454. doi:10.1037/h0077093

Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology, 49,* 298–301. doi:10.1016/j.jesp.2012.10.017

Ratner, K. G., Kaul, C., & Van Bavel, J. J. (2013). Is race erased? Decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. *Social Cognitive and Affective Neuroscience, 8,* 750–755. doi:10.1093/scan/nss063

Ringach, D., & Shapley, R. (2004). Reverse correlation in neurophysiology. *Cognitive Science, 28,* 147–166. doi:10.1207/s15516709cog2802_2

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience, 8,* 657–661. doi:10.1038/nrn2213

Schul, Y., Mayo, R., & Burnstein, E. (2004). Encoding under trust and distrust: The spontaneous activation of incongruent cognitions. *Journal of Personality and Social Psychology, 86,* 668–679. doi:10.1037/0022-3514.86.5.668

Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision, 2,* 105–120. doi:10.1167/2.1.7

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology, 89,* 845–851. doi:10.1037/0022-3514.89.6.845

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006, November 24). Predictive codes for forthcoming perception in the frontal cortex. *Science, 314,* 1311–1314. doi:10.1126/science.1132028

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223,* 96–102. doi:10.1038/scientificamerican1170-96

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

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

Tanis, M., & Postmes, T. (2005). A social identity approach to trust: Interpersonal perception, group membership and trusting behaviour. *European Journal of Social Psychology, 35,* 413–424. doi:10.1002/ejsp.256

Taylor, S. E., Pham, L. B., Rivkin, I. D., & Armor, D. A. (1998). Harnessing the imagination: Mental simulation, self-regulation, and coping. *American Psychologist, 53,* 429–439. doi:10.1037/0003-066X.53.4.429

Todorov, A., Dotsch, R., Wigboldus, D. H. J., & Said, C. P. (2011). Data-driven methods for modeling social perception. *Social and Personality Psychology Compass, 5,* 775–791. doi:10.1111/j.1751-9004.2011.00389.x

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

Tulving, E. (1983). *Elements of episodic memory*. New York, NY: Oxford University Press.

Turner, R. N., & Crisp, R. J. (2010). Imagining intergroup contact reduces implicit prejudice. *British Journal of Social Psychology, 49,* 129–142. doi:10.1348/014466609X419901

Turner, R. N., Crisp, R. J., & Lambert, E. (2007). Imagining intergroup contact can improve intergroup attitudes. *Group Processes & Intergroup Relations, 10,* 427–441. doi:10.1177/1368430207081533

Van Bavel, J. J., & Cunningham, W. A. (2009). Self-categorization with a novel mixed-race group moderates automatic social and racial biases. *Personality and Social Psychology Bulletin, 35,* 321–335. doi:10.1177/0146167208327743

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological Science, 19,* 1131–1139. doi:10.1111/j.1467-9280.2008.02214.x

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the fusiform face area following minimal exposure to motivationally relevant faces: Evidence of in-group enhancement (not out-group disregard). *Journal of Cognitive Neuroscience, 23,* 3343–3354. doi:10.1162/jocn_a_00016

Victor, J. D. (2005). Analyzing receptive fields, classification images and functional images: Challenges with opportunities for synergy. *Nature Neuroscience, 8,* 1651–1656. doi:10.1038/nn1607

Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience, 18,* 411–418.

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,* 503–510. doi:10.1177/0956797613510717

Young, S. G., & Hugenberg, K. (2010). Mere social categorization modulates identification of facial expressions of emotion. *Journal of Personality and Social Psychology, 99,* 964–977. doi:10.1037/a0020400

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Boulder, CO: Westview Press.

Zebrowitz, L. A. (2006). Finally, faces find favor. *Social Cognition, 24,* 657–701. doi:10.1521/soco.2006.24.5.657

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass, 2,* 1497–1517. doi:10.1111/j.1751-9004.2008.00109.x