Group Processes & Intergroup Relations

Effects of Minimal Grouping On Implicit Prejudice, Infrahumanization, and Neural Processing Despite Orthogonal Social Categorizations

Journal:	Group Processes & Intergroup Relations
Manuscript ID	GPIR-18-184.R2
Manuscript Type:	Original Manuscript
Keywords:	minimal group, infrahumanization, bias, implicit
Abstract:	Racial prejudice is a pervasive and pernicious form of intergroup bias. However, a mounting number of studies show that re-categorization— even into minimal groups—can overcome the typical consequences of racial and other group classifications. We tested the effects of minimal grouping on implicit prejudice and infrahumanization using a paradigm in which race was orthogonal to group membership. This allowed us to examine whether knowledge of group membership overrides obvious category differences. We found that participants infrahumanized and showed implicit bias toward the minimal out-group, despite the cross- cutting presence of race, and in fact did not show any of the usual implicit racial bias. In addition, Event-Related Potentials (ERPs) showed an early race effect followed by distinct reactions on the basis of group as processing continued. This is evidence that arbitrary social classifications can engender in-group preference even in the presence of orthogonal, visually salient categorizations.



Running head: EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING

Racial bias remains a pervasive and powerful factor in American society, but complex modern social environments require people to build group affiliations that transcend racial categories. The long history of inequality and the physical dissimilarities between racial groups make race a psychologically salient classifier that would seem difficult to overcome. In contrast, minimal groups draw distinctions between sets of people with no historical, social, or physical differences. When minimal groupings are pitted against established social categories like race, however, they can override automatic racial biases shown in neural and implicit reactions (e.g., Van Bavel, Packer, & Cunningham, 2008). Investigating how and when novel, top-down social categorization affects social processing in the presence of more established and ostensibly more salient and stable categories is an important step toward understanding the complexity of groupcategorization in an ever-changing social environment.

The current research aimed to conceptually replicate and extend previous findings regarding the effects of minimal groups in a cross-categorization paradigm in which minimal group assignment is orthogonal to racial category. To understand multiple levels of bias, we used measures that tap into complex and reflective processes and into early perceptual and attentional processes, none of which had been previously used in this paradigm. These tests included an explicit measure of infrahumanization, a less extreme form of dehumanization, implicit dehumanization and group bias measured with Implicit Association Tests, and early neural processing indexed with event-related potentials (ERPs) and neural motor resonance, which indicate bias in early attention and processing over time. These contributions will further our knowledge not only of whether arbitrary yet salient categorizations predict behavior despite more important orthogonal classifications, but when these categories are processed thanks to the temporal specificity of electroencephalography (EEG). We seek to understand whether bias will

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING

follow arbitrarily salient categorizations (minimal groups) in the presence of societally-ingrained categorizations (racial groups), including looking for the effect of both groupings in neural processing.

Minimal Groups and Categorization

People instinctively categorize others, creating well-documented social boundaries that result in intergroup bias (Brewer, 1979; Hewstone, Rubin, & Willis, 2002; Mullen, Brown, & Smith, 1992). Categories like age, gender, and race are almost instantly perceived (Cikara & Van Bavel, 2014; Fiske & Neuberg, 1990; Freeman & Johnson, 2016; Ito & Bartholow, 2009), and in-groups, however minimal, can receive favoritism (Brewer, 1979; Pinter & Greenwald, 2011; Ratner, Dotsch, Wigboldus, Knippenberg, & Amodio, 2014). At the same time, top-down selfand social-categorization dynamically interacts with bottom-up perception to determine our categorization of and response to others (Freeman & Ambady, 2011, 2014; Freeman & Johnson, 2016; Macrae & Bodenhausen, 2000; Ratner et al., 2014). For instance, studies have linked the top-down processes of group categorization to automatic neurological responses (Cunningham, Van Bavel, Arbuckle, Packer, & Waggoner, 2012; Decety, 2011; Gutsell & Inzlicht, 2013; Hackel, Looser, & Van Bavel, 2014; Ito & Urland, 2005; Ratner & Amodio, 2013) and implicit biases (Ashburn-Nardo, Voils, & Monteith, 2001; Capozza, Andrighetto, Di Bernardo, & Falvo, 2012; DeSteno, Dasgupta, Bartlett, & Cajdric, 2004; McCaslin, 2011; Pinter & Greenwald, 2011; Van Bavel & Cunningham, 2010), emphasizing the effect of social categorization on our perception of those around us.

If categorization is automatic, what happens when targets have multiple relevant group memberships? Sometimes one category dominates the others, sometimes categories are hierarchically ordered, and sometimes there are additive and independent effects of the

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING

categories (Brewer, Ho, Lee, & Miller, 1987; Hewstone, Islam, & Judd, 1993; Klauer, Ehrenberg, & Wegener, 2003). One category can be more salient while another category is still meaningful, and there can, in addition, be emergent properties of their conjunction (Bodenhausen & Peery, 2009; Hehman et al., 2011). Crossing categories, moreover, reduces intergroup bias as common identities can be found (Crisp & Hewstone, 2007). Using multiply-categorized targets allows us to test the intractability of prejudices, crossing charged racial grouping with arbitrary minimal grouping to pit one of the most historically important category biases against a novel classification.

The minimal group paradigm (Taifel, Billig, Bundy, & Flament, 1971) allows us to create experimental social categories that participants have never met, cannot influence, and have no history with, allowing researchers to control for variation in individual experiences, opinions, and motives. Favoritism toward these minimal in-group members-despite participants' knowledge that these groups are new and arbitrary-is prime evidence for how fundamental ingroup-out-group distinctions are to social cognition. Such biases have been shown on a variety of behavioral and neural measures, including distribution of rewards (Demoulin et al., 2009; Tajfel et al., 1971), implicit prejudice (Ashburn-Nardo et al., 2001; Van Bavel & Cunningham, 2009; Capozza et al., 2012), facial attention and recognition (Bernstein, Young, & Hugenberg, 2007; Van Bavel & Cunningham, 2010, 2012), mental images (Ratner et al., 2014), neural face processing indexed by the amplitude of the N170 ERP component (Ratner & Amodio, 2013), and differential activation in a variety of brain regions involved in social perception and social cognition (e.g., Molenberghs & Morrison, 2014). In sum, participants in minimal group studies show many of the biases usually found for more salient and established group distinctions. These effects seem to be driven by favoritism for the minimal in-group, perhaps motivated by identity

enhancement concerns (Ashburn-Nardo et al., 2001; Pinter & Greenwald, 2011).

Minimal group research has found that multiple categorizations are simultaneously encoded but that behavior follows the most salient membership, including minimal groups over racial classifications (Ratner, Kaul, & Van Bavel, 2013; Van Bavel & Cunningham, 2009). Moreover, distinguishing between in- and out-group, even with multiple competing categories, produces in-group favoritism rather than out-group derogation (Brewer, 1979; Scroggins, Mackie, Allen, & Sherman, 2016; Van Bavel, Packer, & Cunningham, 2011). Interestingly, minimal group effects can be elicited despite using mixed-race minimal groups, suggesting that minimal group categorization can overturn race-based biases (Van Bavel & Cunningham, 2009, 2012). Such cross-categorization of race and minimal groups has been shown to override race bias in various domains including neural processing of faces (Van Bavel et al., 2008, 2011), attention to and memory of faces (Van Bavel & Cunningham, 2012), and implicit evaluation (Van Bavel & Cunningham, 2009, 2012). We will extend these results by using the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998), the best-known measure of implicit bias. While there is significant debate over the behavioral correlates of the IAT (e.g. Greenwald, Banaji, & Nosek, 2015; Oswald, Mitchell, Blanton, Jaccard, & Tetlock, 2013), it is a useful measure for our purposes because its popularity gives us a benchmark of typical bias. Insofar as the IAT indexes automatic associations (Nosek, Greenwald, & Banaji, 2007), we can test whether an arbitrary categorization can engender an automatic positivity bias despite the presence of a known stigmatized categorization.

We also evaluate orthogonal categorization's effect on dehumanization, an important and distinct form of bias (Kahn, Goff, & McMahon, 2015), which can also be measured implicitly.

Moreover, to better understand the time course of such biases, we look at neural signatures of reflexive and controlled processing in response to multiply-categorized targets.

Dehumanization

Researchers have tested the effects of multiple categorization on dehumanization by crossing ethnic or immigrant categorization with individuating descriptions including religion, age, nationality, and other classifications (Albarello & Rubini, 2012; Prati, Crisp, Meleady, & Rubini, 2016). But no one has used the strongest contrast of a minimal group versus a socially important classification. Dehumanization is distinct from prejudice (Kahn et al., 2015; Wilde, Martin, & Goff, 2014) as it does not require antipathy (Haslam & Loughnan, 2012). This makes it a valuable test of whether grouping affects in-group preference rather than out-group dislike as it assesses attitude differences without inherently measuring animosity. The present work therefore measures implicit dehumanization to expand our understanding of the extent of the effects of crossed categorizations. Can an arbitrary categorization change automatic associations with humanness, in addition to engendering a positivity bias?

Work using other measures of dehumanization suggests that minimal out-group members are perceived as less human (Capozza, Boccato, Andrighetto, & Falvo, 2009; Hackel et al., 2014) and that multiple categorization of out-group targets reduces dehumanization (Albarello & Rubini, 2012; Prati et al., 2016), but these approaches have not been combined. We investigated not whether multiple categorization overcomes typical biases, but if minimal groups can elicit dehumanization even in the presence of historically important orthogonal classifications. To do so we used two measures, infrahumanization and implicit dehumanization, asking participants to assess both groups' emotional capacities and testing their unconscious associations with each.

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING

Infrahumanization is the perception of out-group members as less human than in-group members (Leyens et al., 2001). Infrahumanization exists independent of out-group derogation and in-group favoritism, instead suggesting a literally impersonal form of bias (Haslam & Loughnan, 2014), operationalized to index intergroup denials of humanity (Haslam & Loughnan, 2014). The measure was developed to assess implicit attitudes (Leyens et al., 2000) and telling participants to purposefully bias their responses in either direction does not change its results, suggesting that infrahumanization is difficult to fake (Eyssel & Ribas, 2012). Implicit Association Tests (IATs) have been used multiple times to assess dehumanization implicitly (Goff, Eberhardt, Williams, & Jackson, 2008; Mekawi, Bresin, & Hunter, 2015; Rudman & Mescher, 2012), but the only minimal group study primarily manipulated disgust (Buckels & Trapnell, 2013). We seek to show implicit dehumanization and infrahumanization¹ due only to minimal grouping, which would show that the mere label of "other" is sufficient to cause differential assignment of humanity.

Neural Processing

Numerous studies have endeavored to find the neurological components of social processes, including grouping, face perception, empathy, and action perception (see Cikara & Van Bavel, 2014 for a review). This literature aims to illuminate the interplay of bottom-up perceptions and top-down knowledge to better understand how we interact with the world. Multiple group memberships are a perfect test, pairing a visual categorization with a memorized, arbitrary one. We used ERPs to test the effects and time-course of these complex social stimuli. We also pre-registered and measured neural resonance as the suppression of mu band oscillations in the sensorimotor cortex—which is thought to show simulation of observed actions within the

¹ Previous work investigating infrahumanization of minimal groups was underpowered with only 16 participants (Demoulin et al., 2009).

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING

observer's own sensorimotor system (Fox et al., 2015)—but results were inconclusive and can be found in the Supplementary Materials along with additional information on this measure.

Visual and attentional differences are reflected very quickly after a stimulus is presented. Facial race and gender are differentiated in the brain as early as 122 ms after stimulus onset (e.g., Ito & Urland, 2003), though face sensitivities prior to 170 ms have been shown to be due to lowlevel visual cues such as stimulus amplitude and color (Rossion & Caharel, 2011). The N170 is thought to be selective to face stimuli, in that it is typically larger in response to human faces compared to objects and nonhuman faces and is thus thought to be involved in face-processing (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Eimer, 2011; Eimer, Kiss, & Nicholas, 2010; Rossion & Jacques, 2009). It is also affected by race (e.g. Balas & Nelson, 2010; Gajewski, Schlegel, & Stoerig, 2008; Ito & Urland, 2005) and minimal in-group (Ratner & Amodio, 2013), although the direction of these effects seems to be context-specific (Ito & Senholzi, 2013; Senholzi & Ito, 2013) and the two have not been juxtaposed. The P200, thought to reflect the orientation of attention to relevant and salient features (Czigler & Gexcy; Eimer, 1997; Kenemans, Kok, & Smulders, 1993), is usually stronger in response to other-race and out-group as opposed to own-race and in-group faces (Hehman, Stanley, Gaertner, & Simons, 2011; Ito & Urland, 2003), while the N200 is usually larger in response to same-race faces as opposed to other-race faces (Ito & Urland, 2003).

Later processing does not respond solely to visual differences. The P3 is considered to indicate the allocation of processing resources to motivationally relevant stimuli and also seems to be affected by race (Ito & Urland, 2003). We do not know whether it responds more to visually apparent or memorized groupings; perhaps the longer time frame allows for more complex target representations or more influence of top-down self-categorization. Looking at the

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING

event-related neural response in this study allows us to test whether, in addition to affecting implicit attitudes, crossed-categorizations alter the neural processing of social categories.

Current Research

The current research aims to contribute to the literature on social re-categorization by making race orthogonal to minimal groups, which we induced using a group memorization paradigm (Van Bavel et al., 2008). We chose this paradigm because it engenders strong grouping effects while creating true minimal groups that do not interact, share history, or choose their affiliations (Hewstone et al., 2002; Pinter & Greenwald, 2011; Van Bavel, Packer, & Cunningham, 2008). We tested its effects on well-established measures of implicit and explicit bias, as well as neural processing. As outlined in our preregistration of this study (https://osf.io/73h6n/), we predicted that assigning participants to an arbitrary minimal in-group that included own-race and other-race individuals and having them memorize the faces of both the minimal in- and out-groups would produce infrahumanization of, implicit bias toward, and less motor resonance with the minimal out-group, with no effect of target race. The motor resonance results are reported in Supplementary Materials along with an exploratory moderation analysis testing whether, in accordance with Social Identity Theory, participants who identify more strongly with their minimal-group would show more implicit bias, infrahumanization, and implicit dehumanization towards the minimal out-group.

To gain a more fine-grained differentiation of the independent effects and interactions of race and minimal group dimensions on early stimulus processing over time, we performed an exploratory ERP analysis, going from early attentional processes that have previously been shown to reflect racial and gender biases (the N170, up to the N200; Hehman et al., 2011; Ito &

Urland, 2003) to more reflective, motivation-driven processing (e.g. the P3; see Willadsen-Jensen & Ito (2006) for a similar approach).

Methods

Participants

75 White university students participated in the study for either money or course credit. 13 participants were excluded due to computer (n = 4) or experimenter error (n = 3), or because they did not self-identify as White (n = 1) or did not meet the 70% correct threshold on the target group memorization task (n = 4). Another was excluded for not remembering the group assignment (n = 1). This left a final sample of 62 participants (32 female, mean age = 21.65, range: 17-56), just below our preregistered sample size of 65, chosen from power analysis of our infrahumanization pilot data, which suggested it to be sufficient for a power of .80. The subjects were distributed evenly into two minimal groups called the Asteroids (n = 32; EEG n = 19) and Comets (n = 30; EEG n = 20). Of these, a subset (n = 39, 20 female, 37 right-handed, mean age = 20.46, range: 17-30) completed the within-subject EEG portion of the study in addition to the behavioral measures. The Institutional Review Board approved the study.

Measures

Implicit Association Tests (IAT)

Participants took three IATs, testing their association of memorized photos of minimal in-group and out-group members with human and animal traits to assess implicit dehumanization, with positive and negative words to assess implicit prejudice, and with self and other concepts. The self-other IAT was included as a manipulation check, always came third, and showed the same general pattern of results as the other two (see Supplementary Materials). Scores on the positivity bias IAT were correlated with scores on the self-other IAT, r = .51, 95%

CI[.30, .68), p < .001, but dehumanization IAT scores were unrelated to either positivity bias, r = .01, 95% CI[-.26, .24), p = .94, or self-other scores, r = .05, 95% CI[-.20, .30), p = .68, suggesting that dehumanization is indeed a distinct form of bias.

Each IAT had the standard seven blocks (Nosek et al., 2007), including three sets of 20 trials introducing faces or concepts (Blocks 1, 2, and 5), two practice blocks of 20 trials (Blocks 3 and 6), and two test blocks of 40 trials (Blocks 4 and 7). The order of blocks in each IAT was counterbalanced across participants, and the IATs were scored using the *D*-algorithm (Greenwald, Nosek, & Banaji, 2003) to compare reaction times in associating in- and out-group faces with the concepts presented. On the positivity bias IAT we doubled the standard number of test trials in order to look at sub-*D* scores for face stimuli so that there were 80 rather than 40 trials in blocks 4 and 7.

Positivity Bias. The positivity bias IAT used the standard set of positive ("joy," "love," "peace," "wonderful," "pleasure," "glorious," "laughter," "happy") and negative words ("agony," "terrible," "horrible," "nasty," "evil," "awful," "failure," "hurt;" Nosek et al., 2007), testing participants' preferences for minimal group faces (see Figure 1). It has been shown to have satisfactory test-retest reliability (median r = .56; Nosek, et al., 2007). Participants categorized words as "Good" or "Bad" and faces as "Asteroid" or "Comet."

Dehumanization. The dehumanization IAT tested the association of in- and out-group photos with words related to animals (*"animals," "nature," "instinct," "physical, "bodies*") and words related to humans (*"culture," "society," "mind," "symbols," "monuments"*; Rudman & Meschner, 2012). We predicted that this would provide convergent evidence with the infrahumanization measure described below. Participants categorized words as "Animal" or "Human" and faces as "Asteroid" or "Comet."

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING 11

Emotional Attribution Task

Infrahumanization is operationalized as the differential attribution of secondary emotions to in-group rather than out-group members (Demoulin et al., 2004). Secondary emotions *sentiments* in French—like love and hope are uniquely human, while primary emotions *émotions*—like anger and fear are also attributed to non-human agents like animals (Leyens et al., 2000; Leyens et al., 2001).

During this task, participants read a list of 20 adjectives and were asked to pick every word they thought was characteristic of the minimal in-group or of the minimal out-group. They were then presented with the same list and asked to choose again for the other group. The order of groups was randomized between participants. The 20 adjectives included 14 filler words related to warmth and competence, three primary emotions (*"happiness," "surprise," "pleasure"*) and three secondary emotions (*"tenderness," love," "hope"*). The emotion words were chosen based on their similar valences and consensus among raters (Cortes, Demoulin, Rodriguez, Rodriguez, & Leyens, 2005; Demoulin et al., 2004). Only positive emotions were used to more closely replicate Demoulin et al.'s (2009) work.

Procedure

After signing the consent form, participants were told they would be learning about two groups, the Asteroids and Comets, and that they had been assigned to one of the two. They were then photographed in front of a blank wall at a standard distance and zoom intended to maximize the photographs' similarity to the stimuli. For EEG participants, recording was set up at this point, as detailed below. Participants completed two learning tasks on the computer in which they were shown 16 faces (plus their own to enhance their sense of affiliation) split into two groups labeled Asteroids and Comets. Both the Asteroids and the Comets were always made up

of four Black and four White faces, but the assignment of faces to groups was randomized such that all participants saw the same 16 faces with unique group assignments. The first learning task consisted of participants seeing this presentation of 17 faces—eight in their minimal in-group, plus their own, on one side of screen, eight in their out-group on the other side of the screen—and having three minutes to memorize them. The side of the screen on which each group was presented was counterbalanced across participants, and the position of each face within the two grids was random. The second, sorting task presented the faces one at a time in the center of the screen without group labels, and participants were asked to sort them into the proper minimal group by pressing the 'E' and 'I' keys on the keyboard. Incorrect responses produced an "X" on the screen and required participants to click the correct key to proceed. This served to confirm participants' knowledge of the groups; on average participants were correct on 90.85% of trials and participants who scored below 70% were excluded from analysis.

Participants then took the IATs, which comprised the same images of minimal group members' faces as the learning tasks, again without any indication of group membership. Because we were concerned about the efficacy of repeated IATs, the group-identification test the least theoretically important—was always presented last, and the order of the first two was randomized. Next, participants completed several questionnaires described below in "Additional Measures."

At this point, participants in the EEG version of the study watched four three-minute blocks of videos of right hands squeezing stress balls. The memorized faces were paired with hand videos to suggest that the hands belonged to the target minimal group members. ERPs were computed for each face presentation. The stress ball squeezes were used to calculate motor resonance.

Lastly, all participants filled out a brief demographic questionnaire, before finishing with the final manipulation check asking them to report which minimal group was theirs. Participants were then debriefed, compensated, and dismissed.

Stimuli

Target images were of males photographed in grey t-shirts on white backgrounds with only their heads and shoulders visible. These were obtained from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015), and were chosen by eliminating all targets who were not between 18 and 25 years old and labeled as White or Black by 100% of raters. We then standardized the attractiveness and threatening-ness ratings for the remaining faces and selected eight of each race that were within one standard deviation of the mean on both measures.

EEG task. The order of stimulus presentation was as follows: 2000 ms of white noise followed by a jittered fixation cross for between 500 and 800 ms, a 1500 ms presentation of a memorized face, another jittered 500 – 800 ms fixation cross, and then three repetitions of a 2000 ms video of a hand squeezing a stress ball, with jittered fixation crosses between each presentation (see Figure 2). The intertrial period was between 10 and 1000 ms. The white noise and videos covered 20 and 19 percent of the screen, respectively, while the images covered four percent. All stimuli were presented centered on 1920 x 1080 pixel monitors. Each of the 16 memorized faces was presented once per block and there were four blocks. These face presentations were used for ERP calculation.

Presentation order was randomized within block and each face was paired with a hand video of the same race. We made the videos in lab, recording people squeezing a stress ball at 1 Hz and cutting the videos into two-second clips (see Figure 2 for a still image). We used eight clips of Black hands and eight clips of White hands, and told participants there were multiple

videos for each face, allowing hands and faces to be matched randomly within race. We have used similar stimuli in other studies (Hager, Yang, & Gutsell, 2018).

EEG recording and processing. EEG was recorded from 33 active electrodes embedded in a stretch-lycra cap (ActiCap, BrainProducts GmbH, Munich, Germany) arranged according to the 10-20 system with impedances kept below 20 k Ω . The EEG was digitized at 500 Hz using BrainAmp amplifiers and BrainVision recorder software (BrainProducts GmbH, Munich, Germany) with an initial reference at FCz and re-referenced offline to the average of all EEG electrodes. The data was filtered through 30 Hz low-pass and 0.01 Hz high-pass zero-phase Butterworth filters (24 dB) and ocular artifacts were corrected using the VEOG channel² and an ICA-based procedure for isolating and removing ocular artifacts (Croft & Barry, 2000). Remaining artifacts exceeding \pm 100 μ V in amplitude, with a voltage step larger than 200 μ V between sample points, or a maximum voltage difference of less than 0.5 μ V within a 100 ms interval were rejected for individual channels in each trial.

For the ERP analysis, we created artifact-free epochs from 200 ms pre-picture onset to 1500 ms post-picture onset for each face presentation. The N170 was evaluated as the average activity in a 150 to 250 ms window at left posterior electrodes TP9, P7, and P09 and at right posterior electrodes TP10, P8, and P10; for P200 we computed mean amplitudes between 150 and 250 ms at Cz; the N200 was evaluated as the average activity between 200 and 350 ms at central electrode Cz; and the P3 was evaluated as the average activity between 350 and 700 ms at Pz.

Additional Measures

² Technical issues with the VEOG channel led to the use of the FP2 channel for five participants and FP1 for an additional three.

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING 15

Participants also completed a four-item Likert scale measure of their identification with both their racial and minimal groups (e.g., "I identify with being an Asteroid" from 1 = "*Strongly Disagree*" to 7 = "*Strongly Agree*"; Postmes, Haslam, & Jans, 2012); the Symbolic Racism 2000 scale (Henry & Sears, 2002), an eight-item measure of prejudice toward African Americans; and an exploratory visual measure of perceived conflict between groups. We also pre-registered a final manipulation check but a programming mistake misreported its data for most participants.

Results

There were no significant differences between participants assigned to be "Asteroids" and participants assigned to be "Comets" in age, gender, expressed commitment to minimal group, expressed commitment to racial group, symbolic racism, perception of conflict between the two groups, or IAT *D*-scores on any of the three IATs (all ps > .27). We also found no significant differences between male and female participants on any of our dependent variables, including resonance (all ps > .10). Therefore, neither participant gender nor specific minimal group membership was considered in the following analyses. Confirming our manipulation, participants implicitly associated themselves more strongly with their minimal in-group than with the minimal out-group, according to the self/other IAT, D = .42, 95% CI[.31, .52].

Implicit Bias

IATs were scored using the *D*-algorithm (Greenwald et al., 2003), which measures the effect size of the response time difference between stimulus pairings [(in-group and good + out-group and bad) versus (in-group and bad + out-group and good)]. In accordance with our pre-registration and the recommendation of Nosek, Greenwald, and Banaji (2007), we excluded trials over 10,000 ms as well as the participant for whom more than 10% response times were under 300 ms. For the dehumanization and group-identification IATs, practice and task blocks were

scored separately, then averaged, but due to the doubling of task trials on the positivity IAT we did not separate practice and task scores, instead simultaneously computing a single *D*-score for all trials. This is not standard practice, but it gave us more power to tease apart the effects of race and minimal group. We would expect that, if anything, the added trials would decrease bias scores due to practice effects.

To test whether a particular combination of race and minimal group membership was driving our effects, we ran a 2 (race: Black or White) x 2 (group: in- or out-) repeated measures ANOVA using mean response time difference scores (bad minus good; animal minus human; other minus self). The positivity bias IAT and group-identification IAT response times were inverse rooted and the dehumanization IAT response times were inverted to normalize the data, as these response times are positively skewed.³ Transformed values were used in statistical models; reported *D*-scores and means were calculated with untransformed data.

Participants had a mean *D*-score of .35, 95% CI[.26, .43] on the positivity bias IAT, indicating a preference for the minimal in-group relative to the minimal out-group. We also calculated *D*-scores using only trials in which participants categorized target faces. These trials, unlike the word stimuli trials, contain information about both race and minimal group—one memorized, one visually apparent—allowing us to directly compare response times. This unorthodox approach, aided by the doubled number of trials, showed no bias on the basis of race, D = .031, 95% CI [-.016, .078], and similar minimal group bias, D = .34, 95% CI[.25, .42], although it is important to note that we did not include a condition that required participants to sort faces according to race, so we cannot truly compare minimal group bias to racial bias.

³ Transformation methods were determined by a Box-Cox test, R MASS package (Venables & Ripley, 2002). Transformation did not alter any results.

The ANOVA showed a significant main effect of minimal group, F(1, 61) = 52.24, p < .001, $\eta^2 = .46$, 95% CI[.27, .59],⁴ and no main effect of race, F(1, 61) = 1.18, p = .28, $\eta^2 = .02$, 95% CI[0, .13], nor interaction of minimal group and race, F(1, 61) = 2.86, p = .096, $\eta^2 = .045$, 95% CI[0, .18] (Figure 3). Pre-planned simple comparisons found a significant difference between response times for White in-group (M = 229.60ms, SD = 379.53 ms)⁵ versus White outgroup targets (M = .188.79 ms, SD = 267.12 ms), F(1, 61) = 42.51, p < .001, $\eta^2 = .41$, 95% CI[.22, .55], and for Black in-group (M = 173.04 ms, SD = 259.23 ms) versus Black out-group targets (M = .174.59 ms, SD = 311.22 ms), F(1, 61) = 39.61, p < .001, $\eta^2 = .39$, 95% CI[20, .54]. The difference in response times between White in-group members and Black in-group members was outside the traditional threshold of significance, F(1, 61) = 3.79, p = .056, $\eta^2 = .059$, 95% CI[0, .20], and the racial difference for out-group members was also not significant, p = .56. **Dehumanization**

Infrahumanization.

To assess condition effects on infrahumanization we ran a 2 (group: in- vs. out-) x 2 (emotion: primary or secondary) repeated measures ANOVA, which revealed a main effect of minimal group, F(1, 61) = 4.81, p = .032, $\eta^2 = .073$, 95% CI[0, .22]. There was no effect of emotion, F(1, 61) = 1.00, p = .32, $\eta^2 = .016$, 95% CI[0, .12], nor was there an interaction, F(1, 61) = .14, p = .71, $\eta^2 = .002$, 95% CI[0, .08]. Pre-planned simple comparisons of emotion showed a significant difference between in-group (M = .77) and out-group attribution (M = .53)

⁴ Confidence intervals reported here are on η^2 , calculated with the 'apaTables' package in R (Stanley, 2018). They cannot be negative.

⁵ The means reported here are untransformed mean difference scores. The White in-group mean indicates that participants were 230 ms faster on average to categorize White in-group targets when they were paired with the positive words than with the negative. A negative mean indicates that the targets were more quickly associated with bad than good.

of secondary emotions, F(1, 61) = 4.67, p = .035, $\eta^2 = .071$, 95% CI[0, .21], confirming our preregistered hypothesis. The difference in primary emotion attribution was not significant (p = .13; see Figure 4). This suggests that in-group members were seen as having more complex emotions than out-group members: out-group members were infrahumanized. Other work has similarly reported infrahumanization results through planned comparison of secondary emotions in the absence of a significant interaction (Cehajic, Brown, & González, 2009), but it is possible that the outgroup was simply attributed fewer emotions of all kinds.

IAT

The dehumanization IAT showed the same pattern of results as the positivity bias IAT. Participants were quicker to associate the in-group rather than the out-group with human versus animal concepts, D = .22, 95% CI[.12, .33], and there was no race effect in the face stimuli trials, D = .031, 95% CI[-.099, .038].

The ANOVA showed the expected significant main effect of minimal group, F(1, 61) =10.51, p = .002, $\eta^2 = .15$, 95% CI[.02 .31], no effect of race, F(1, 61) = 3.13, p = .082, $\eta^2 = .049$, 95% CI[0, .18], and no interaction between group and race, F(1, 61) = 1.06, p = .31, $\eta^2 = .017$, 95% CI[0, .12] (see Figure 5). Although the difference was not significant, Black targets (M =37.28 ms, SD = 474.24) elicited less dehumanization bias than White targets (M = .18.95 ms, SD= 477.00). Simple comparisons found a significant difference between response times for White in-group (M = 117.04 ms, SD = 498.13 ms) versus White out-group targets (M = .154.95 ms, SD= 408.79 ms), F(1, 61) = 10.42, p = .002, $\eta^2 = .15$, 95% CI[.02, .30], and for Black in-group (M= 186.74 ms, SD = 418.85 ms) versus Black out-group targets (M = .112.18 ms, SD = 476.32ms), F(1, 61) = 5.63, p = .021, $\eta^2 = .084$, 95% CI[.001, .23]. Thus, the minimal out-group was dehumanized as indexed by both infrahumanization and implicit association with animals.

Exploratory Analyses

ERP analysis. To assess the effects of minimal group membership and race on ERPs related to early attention and motivated processing, we conducted a series of 2 (minimal group: in- or out-) x 2 (race: Black or White) ANOVAs on ERP mean amplitudes (see Table 1 for means and standard deviations and Figure 6 for a depiction of the ERP waveforms). ERPs related to early visual face processing and attention—N170 in both right and left hemispheres, P200, and N200—all showed the expected main effect of race (all ps < .008), no main effect of minimal group (all ps > .21), and no interaction between group and race (all ps > .27). P200 and both N170s responded more strongly to the Black targets and N200 responded more strongly to White targets. These robust results (all $\eta^2 s > .18$) match research showing early potential responses to only racial categorization in multiply categorized targets (Alonso-Prieto et al., 2015; Cassidy, Boutsen, Humphreys, & Quinn, 2014; Hehman et al., 2011b; Wiese, 2012).⁶ Interestingly, a different pattern emerged for the P3 reflective of motivated selective processing: the ANOVA revealed no significant main effect of race, F(1, 37) < .001, p = .99, $\eta^2 < .001$, or minimal group F(1, 37) = 1.44, p = .24, $\eta^2 = .04$, 95% CI[0, .20], but did show an interaction between race and minimal group F(1, 37) = 4.44, p = .042 $\eta^2 = .11$, 95% CI[0, .30]. Post-hoc simple comparisons found only a significant simple effect of minimal group for White targets F(1, 37) = 5.77, $p < 10^{-10}$.021, $\eta^2 = .14,95\%$ CI[.001, .33], with a larger P3 for White out-group members ($M = 1.68 \text{ \muV}$, $SD = 2.20 \text{ \muV}$) than White in-group members ($M = .69 \text{ \muV}$, $SD = 2.48 \text{ \muV}$). Please note, however, that these P3 effects do not remain significant after correction for multiple comparisons (using a Holms-Bonferroni corrected p-value of 0.005). Taken together, the ERP analysis suggest that although minimal group is driving behavior and self-report, early perceptional neural processing

⁶ Hehman et al. (2011b) did find an effect of orthogonal non-racial categorization in N200, but only at occipital electrodes, which we did not test.

Page 20 of 60

is driven by race and is not affected by minimal group. Whether minimal group does affect later more reflexive processing remains unclear.

Group identification. Finally, participants reported stronger identification with their racial groups (M = 4.23, SD = 1.02) than with their minimal groups (M = 3.88, SD = 1.31) on the four-item group commitment measures, t(61) = 2.01, p = .048, d = .30, and this identification correlated with Symbolic Racism, r = .29, 95% CI[.05, .51], t(60) = 2.38, p = .02—higher White identifiers evinced more modern racism. Minimal group identification also correlated with positivity bias IAT scores, r = .35, 95% CI[.11, .55], t(60) = 2.87, p = .0057, such that higher minimal group identifiers showed more implicit bias for their minimal group. See Table 2 for selected correlations between our variables. See the Supplementary Materials for exploratory moderation analyses showing minimal group bias was stronger among participants who identified more strongly with their minimal in-group, consistent with Social Identity Theory and previous work (Van Bavel & Cunningham, 2012, Study 2).

Discussion

These results add to the existing evidence that intergroup bias is influenced by salient social categorization, not only visual markers of group membership. In a paradigm where the only information given about two groups was their names and the participant's assignment to one, participants associated more strongly with their new in-group, showed implicit bias in favor of their in-group, and, for the first time, infrahumanized the minimal out-group. All of this occurred despite the presence of race, a visually salient, orthogonal social category. Interestingly, in contrast to this behavioral and self-reported focus on the minimal group, early neural processing was primarily driven by race. Early components of the ERP (the N170, P200, and the N200) were solely affected by target race. While these early ERP components are often

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING 21

influenced primarily by simple stimulus and task factors, later ERPs such as the P3 are influenced by more complex stimulus factors and contextual information (Willadsen-Jensen & Ito, 2006) including, in this study, minimal group membership. Race no longer drove processing at the time of the P3 and we saw a trending interaction effect suggesting that minimal group membership might come into play at this time, potentially giving rise to implicit and explicit behavioral responses, such as infrahumanization. Moreover, the simple effect showing a stronger response to White outgroup than White ingroup targets is consistent with P3 responding to complexity, but this is an interpretation requiring further research.

Implications

Minimal groups engender dehumanization. Two measures provide evidence that the minimal out-group is seen as less human than the in-group. For the first time, participants infrahumanized a minimal out-group, ascribing fewer uniquely human emotions to them despite the arbitrariness of the assignment and the orthogonal presence of multiple racial groups. Participants also implicitly dehumanized the minimal out-group, again despite the presence of a more traditionally meaningful social classifier. These results suggest that a small, salient categorization can cause a target to be perceived as less than fully human, an important finding given that dehumanization has been shown to uniquely predict societal issues like police violence (Goff, Jackson, Di Leone, Culotta, & DiTomasso, 2014; Kahn et al., 2015). Recent work has found an unexpected willingness on the part of participants to endorse blatant dehumanization of out-groups (e.g., Kteily, Bruneau, Waytz, & Cotterill, 2015; Kteily, Hodson, & Bruneau, 2016). Whether such explicit attitudes could be found toward minimal out-groups—let alone in the presence of orthogonal categories—remains to be seen, but the finding that

minimal out-groups are infrahumanized and implicitly dehumanized raises the possibility that they could be blatantly dehumanized as well.

In addition, the infrahumanization results combined with theoretical suggestions that dehumanizing others is "default" (Waytz, Schroeder, & Epley, 2014) suggest that the label of "in-group" might increase perceptions of humanity. Future research should differentiate whether grouping causes the in-group to be considered more human, the out-group to be considered less human, or both. Consensus holds that in-group favoritism predominates over out-group derogation in determining biased behaviors (Brewer, 1979; Hewstone et al., 2002; Kurzban, Tooby, & Cosmides, 2001), even in minimal group paradigms (Ratner et al., 2014). It is possible that perceptions of humanity are similarly partial to the in-group rather than biased against the out-group.

Additionally, participants showed implicit bias on the basis of minimal group (D = .35) similar to the typical racial bias found in hundreds of thousands of online participants ($M_{project}$ $_{implicit} = .28, 95\%$ CI[.281, .283]; Xu et al., 2016), meaning that just-introduced, just-memorized faces engendered at least as much bias as the historically and visually salient category of race, even though race differences were present. This conceptually replicates findings using other measures, supporting re-categorization as a powerful method to combat unconscious bias (e.g., Van Bavel & Cunningham, 2009, 2012).

Early processing was driven by race. The early ERP results match literature suggesting that race is quickly differentiated (c.f., Ito & Bartholow, 2009) and extend the literature confirming that race remains the driving factor of early perceptional processing even in the presence of an orthogonal minimal group category. Minimal group categorization has been shown to be reflected as early as 170 ms (Ratner & Amodio, 2013) in the absence of race

EFFECTS OF MINIMAL GROUPING ON SOCIAL PREFERENCES AND PROCESSING 23

differences, suggesting a potential disruptive effect of a culturally ingrained group category on minimal group categorization. Since minimal group membership does drive our behavioral effects, it must be processed eventually, but the exact timing remains unclear. The P3 was no longer modulated by race and we found a non-significant trend suggesting that minimal group might interact with race as early as 400 ms into processing. Since these effects are the result of exploratory, non-preregistered analyses, more research is needed to conclusively pinpoint the first time-point of modulation by group membership in cross-categorization designs.

Limitations

While we used symbolic racism as a measure of participants' prejudice, there is a possibility that our sample was not racially biased to begin with. Symbolic racism does not have a known mean, but online samples from our laboratory report more bias than the students in this study. Our design also lacks an IAT condition during which participants would have to sort based on race rather than group membership. Although race was implicit in our face stimuli, only group membership was made salient. Therefore, we cannot directly compare participants' racial bias to their minimal group bias. Previous research using response-window priming found that minimal group preference overrode racial bias for the in-group, but not for the out-group or for novel Black and White faces (Van Bavel & Cunningham, 2009). Given that our participants were never asked to categorize by race, we do not know whether they would have shown bias. However, we can still see minimal group bias and dehumanization despite the presence of orthogonal racial categorizations, including on measures not previously used in this paradigm.

Finally, our design did not include measures of discriminatory behavior, which would have allowed us to explore the connection of ecologically valid behaviors to these biases. It is one thing for an orthogonal categorization to erase widely-seen biases on widely-used measures.

It is another for that sort of re-categorization to have an effect on real world prejudice. Multiple categorization research (e.g. Albarello & Rubini, 2012; Crisp, Turner, & Hewstone, 2010) suggests ways of crossing actual rather than minimal categories, and shows promise as a means of changing cognition around stigmatized social groups. This research reinforces that effort to overturn biases with subtle manipulations which can hopefully be expanded into realistic interventions.

Conclusion

This study was the first to look at early representations of multiply categorized targets and to measure infrahumanization of those targets. Our findings suggest that intergroup bias is not inherently tuned to concrete prejudices toward specific groups, but rather follows salient social distinctions, even in the presence of notable orthogonal differences. We showed that participants were biased toward a minimal out-group, even infrahumanizing its members, despite the presence of racial out-group members in both minimal groups, and that while racial group membership seems to be coded within the first 200 ms of neural processing, minimal group membership may be processed in the next 200. Clearly in-group favoritism is alive and well, but it seems to be easily drawn even to the most arbitrary of groups. Invisible, novel categorizations can be as powerful as visually, historically important differences — at least on laboratory measures of implicit prejudice and infrahumanization — and in some cases seem able to override them. This suggests that category salience may be an effective means of combating prejudice.

Running head: Effects of Minimal Grouping on Social Perception and Processing.

References

Albarello, F., & Rubini, M. (2012). Reducing dehumanisation outcomes towards Blacks: The role of multiple categorisation and of human identity. *European Journal of Social Psychology*, 42(7), 875–882. http://doi.org/10.1002/ejsp.1902

Alonso-Prieto, E., Oruç, I., Rubino, C., Zhu, M., Handy, T., & Barton, J. J. S. (2015).
Interactions between the perception of age and ethnicity in faces: an event-related potential study. *Cognitive Neuropsychology*, *32*(6), 368–384.
http://doi.org/10.1080/02643294.2015.1061981

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(5), 789–799. http://doi.org/10.1037/0022-3514.81.5.789

Balas, B., & Nelson, C. A. (2010). The role of face shape and pigmentation in other-race face perception: An electrophysiological study. *Neuropsychologia*, 48(2), 498–506. http://doi.org/10.1016/j.neuropsychologia.2009.10.007

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The Cross-Category Effect. *Psychological Science*, *18*(8), 706–713. http://doi.org/10.1111/j.1467-9280.2007.01964.x

Bodenhausen, G. V., & Peery, D. (2009). Social Categorization and Stereotyping In vivo : The VUCA Challenge. *Social and Personality Psychology Compass*, 3(2), 133–151. http://doi.org/10.1111/j.1751-9004.2009.00167.x

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitivemotivational analysis. *Psychological Bulletin*, 86(2), 307–324. http://doi.org/10.1037/0033-2909.86.2.307

Brewer, M. B., Ho, H.-K., Lee, J.-Y., & Miller, N. (1987). Social Identity and Social Distance

Effects of Minimal Grouping on Social Perception and Processing.

Among Hong Kong Schoolchildren. *Personality and Social Psychology Bulletin*, 13(2), 156–165.

- Buckels, E. E., & Trapnell, P. D. (2013). Disgust facilitates outgroup dehumanization. *Group Processes & Intergroup Relations*, 20(10), 1–10. http://doi.org/10.1177/1368430212471738
- Capozza, D., Andrighetto, L., Di Bernardo, G. A., & Falvo, R. (2012). Does status affect intergroup perceptions of humanity? *Group Processes & Intergroup Relations*, 15(3), 363– 377. http://doi.org/10.1177/1368430211426733
- Capozza, D., Boccato, G., Andrighetto, L., & Falvo, R. (2009). Categorization of Ambiguous Human/Ape Faces: Protection of Ingroup but Not Outgroup Humanity. *Group Processes & Intergroup Relations*, 12(6), 777–787. http://doi.org/10.1177/1368430209344868
- Cassidy, K. D., Boutsen, L., Humphreys, G. W., & Quinn, K. A. (2014). Ingroup categorization affects the structural encoding of other-race faces: Evidence from the N170 event-related potential. *Social Neuroscience*, 9(3), 235–248. http://doi.org/10.1080/17470919.2014.884981
- Cehajic, S., Brown, R., & González, R. (2009). What do I Care? Perceived Ingroup
 Responsibility and Dehumanization as Predictors of Empathy Felt for the Victim Group. *Group Processes & Intergroup Relations*, 12(6), 715–729.
 http://doi.org/10.1177/1368430209347727
- Cikara, M., & Van Bavel, J. J. (2014). The Neuroscience of Intergroup Relations: An Integrative Review. *Perspectives on Psychological Science*, *9*(3), 245–274. http://doi.org/10.1177/1745691614527464
- Cortes, B. P., Demoulin, S., Rodriguez, R. T., Rodriguez, A. P., & Leyens, J. (2005). Infrahumanization or Familiarity? Attribution of Uniquely Human Emotions to the Self, the

Ingroup, and the Outgroup. *Personality and Social Psychology Bulletin*, *31*(2), 243–253. http://doi.org/10.1177/0146167204271421

- Crisp, R. J., & Hewstone, M. (2007). Multiple Social Categorization. In Advances in Experimental Social Psychology (Vol. 39, pp. 163–254). http://doi.org/10.1016/S0065-2601(06)39004-1
- Crisp, R. J., Turner, R. N., & Hewstone, M. (2010). Common ingroups and complex identities:
 Routes to reducing bias in multiple category contexts. *Group Dynamics: Theory, Research, and Practice*, 14(1), 32–46. http://doi.org/10.1037/a0017303
- Cunningham, W. A., Van Bavel, J. J., Arbuckle, N. L., Packer, D. J., & Waggoner, A. S. (2012).
 Rapid social perception is flexible: approach and avoidance motivational states shape P100 responses to other-race faces. *Frontiers in Human Neuroscience*, 6(May), 1–7.
 http://doi.org/10.3389/fnhum.2012.00140
- Decety, J. (2011). Dissecting the Neural Mechanisms Mediating Empathy. *Emotion Review*, *3*(1), 92–108. http://doi.org/10.1177/1754073910374662

Demoulin, S., Cortes, B. P., Viki, G. T., Rodriguez, A. P., Rodriguez, R. T., Paladino, M. P., & Leyens, J.-P. (2009). The role of in-group identification in infra-humanization.
 International Journal of Psychology, 44(1), 4–11.

http://doi.org/10.1080/00207590802057654

Demoulin, S., Leyens, J.-P., Paladino, M., Rodriguez-Torres, R., Rodriguez-Perez, A., & Dovidio, J. (2004). Dimensions of "uniquely" and "non-uniquely" human emotions. *Cognition & Emotion*, 18(1), 71–96. http://doi.org/10.1080/02699930244000444

DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science*, *15*(5), 319–324.

Effects of Minimal Grouping on Social Perception and Processing.

http://doi.org/10.1111/j.0956-7976.2004.00676.x

- Eimer, M. (1997). An event-related potential (ERP) study of transient and sustained visual attention to color and form. *Biological Psychology*, 44(3), 143–160. http://doi.org/10.1016/S0301-0511(96)05217-9
- Eyssel, F., & Ribas, X. (2012). How to be good (or bad): On the fakeability of dehumanization and prejudice against outgroups. *Group Processes and Intergroup Relations*, 15(6), 804–812. http://doi.org/10.1177/1368430212447204
- Fiske, S. T., & Neuberg, S. L. (1990). A Continuum of Impression Formation, From Category-Based to Individuating Processes: Influences of Information and Motivation on Attention and Interpretation. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (pp. 1–62).
- Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N.,
 Vanderwert, R. E., ... Van Ijzendoorn, M. H. (2015). Assessing Human Mirror Activity
 With EEG Mu Rhythm: A Meta-Analysis. *Psychological Bulletin*.
 http://doi.org/10.1037/bul0000031
- Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review*, *118*(2), 247–279. http://doi.org/10.1037/a0022327
- Freeman, J. B., & Ambady, N. (2014). The dynamic interactive model of person construal:
 Coordinating sensory and social processes. In J. W. Sherman, B. Gawronski, & Y. Trope (Eds.), *Dual-process theories of the social mind* (pp. 235–248). Guilford Press.
- Freeman, J. B., & Johnson, K. L. (2016). More Than Meets the Eye: Split-Second Social Perception. *Trends in Cognitive Sciences*, *1548*(5), 1–13.

http://doi.org/10.1016/j.tics.2016.03.003

Effects of Minimal Grouping on Social Perception and Processing.

- Gajewski, P. D., Schlegel, K., & Stoerig, P. (2008). Effects of Human Race and Face Inversion on the N170. *Journal of Psychophysiology*, 22(4), 157–165. http://doi.org/10.1027/0269-8803.22.4.157
- Goff, P. A., Eberhardt, J. L., Williams, M. J., & Jackson, M. C. (2008). Not yet human: implicit knowledge, historical dehumanization, and contemporary consequences. *Journal of Personality and Social Psychology*, *94*(2), 292–306. http://doi.org/10.1037/0022-3514.94.2.292
- Goff, P. A., Jackson, M. C., Di Leone, B. A. L., Culotta, C. M., & DiTomasso, N. A. (2014). The essence of innocence: Consequences of dehumanizing Black children. *Journal of Personality and Social Psychology*, *106*(4), 526–545. http://doi.org/10.1037/a0035663
- Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically Small Effects of the Implicit Association Test Can Have Societally Large Effects. *Journal of Personality and Social Psychology*, 108(4), 553–561. http://doi.org/10.1037/pspa0000016
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality* and Social Psychology, 74(6), 1464–1480. http://doi.org/10.1037/0022-3514.74.6.1464
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85(2), 197–216. http://doi.org/10.1037/0022-3514.85.2.197
- Gutsell, J. N., & Inzlicht, M. (2013). Using EEG Mu-Suppression to Explore Group Biases in Motor Resonance. In B. Derks, D. Scheepers, & N. Ellemers (Eds.), *Neuroscience of Prejudice and Intergroup Relations* (pp. 278–298). London: Psychology Press.

Hackel, L. M., Looser, C. E., & Van Bavel, J. J. (2014). Group membership alters the threshold

Effects of Minimal Grouping on Social Perception and Processing.

for mind perception: The role of social identity, collective identification, and intergroup threat. *Journal of Experimental Social Psychology*, *52*, 15–23. http://doi.org/10.1016/j.jesp.2013.12.001

Hager, B. M., Yang, A. C., & Gutsell, J. (2018). Measuring Brain Complexity During Neural Motor Resonance. *Frontiers in Neuroscience*, 12(October), 758. http://doi.org/10.3389/FNINS.2018.00758

- Haslam, N., & Loughnan, S. (2012). Prejudice and Dehumanization. In J. Dixon & M. Levin (Eds.), *Beyond Prejudice* (pp. 89–104). London: Psychology Press.
- Haslam, N., & Loughnan, S. (2014). Dehumanization and infrahumanization. Annual Review of Psychology, 65, 399–423. http://doi.org/10.1146/annurev-psych-010213-115045

Hehman, E., Stanley, E. M., Gaertner, S. L., & Simons, R. F. (2011a). Multiple group membership influences face-recognition : Recall and neurological evidence. *Journal of Experimental Social Psychology*, 47(6), 1262–1268.

http://doi.org/10.1016/j.jesp.2011.05.014

Hehman, E., Stanley, E. M., Gaertner, S. L., & Simons, R. F. (2011b). Multiple group membership influences face-recognition: Recall and neurological evidence. *Journal of Experimental Social Psychology*, 47(6), 1262–1268.

http://doi.org/10.1016/j.jesp.2011.05.014

- Henry, P. J., & Sears, D. O. (2002). The Symbolic Racism 2000 Scale. *Political Psychology*, 23(2), 253–283. http://doi.org/10.1111/0162-895X.00281
- Hewstone, M., Islam, M. R., & Judd, C. M. (1993). Models of cross categorization and intergroup relations. *Journal of Personality and Social Psychology*, *64*(5), 779–793. http://doi.org/10.1037//0022-3514.64.5.779

Effects of Minimal Grouping on Social Perception and Processing.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup Bias. *Annual Review of Psychology*, 53(1), 575–604. http://doi.org/10.1146/annurev.psych.53.100901.135109

Ito, T. A., & Bartholow, B. D. (2009). The neural correlates of race. *Trends in Cognitive Sciences*, *13*(12), 524–531. http://doi.org/10.1016/j.tics.2009.10.002

Ito, T. A., & Senholzi, K. B. (2013). Us versus them: Understanding the process of race perception with event-related brain potentials. *Visual Cognition*. http://doi.org/10.1080/13506285.2013.821430

Ito, T. A., & Urland, G. (2005). The Influence of Processing Objectives on the Perception of Faces : An ERP Study of Race and Gender Perception. *Cognitive, Affective, & Behavioral Neuroscience*, 5(1), 21–36. http://doi.org/10.3758/CABN.5.1.21

- Ito, T. A., & Urland, G. R. (2003). Race and Gender on the Brain : Electrocortical Measures of Attention to the Race and Gender of Multiply Categorizable Individuals. *Journal of Personality and Social Psychology*, 85(4), 616–626. http://doi.org/10.1037/0022-3514.85.4.616
- Kahn, K. B., Goff, P. A., & McMahon, J. M. (2015). Intersections of Prejudice and Dehumanization: Charting a Research Trajectory. In W. D. Hund, Ch. W. Mills, & S.
 Sebastiani (Eds.), *Simianization: Apes, Gender, Class, and Race* (pp. 223–241). Münster: LIT Verlag.
- Klauer, K. C., Ehrenberg, K., & Wegener, I. (2003). Crossed categorization and stereotyping:
 Structural analyses, effect patterns, and dissociative effects of context relevance. *Journal of Experimental Social Psychology*, *39*(4), 332–354. http://doi.org/10.1016/S0022-1031(03)00017-9

Kteily, N., Bruneau, E., Waytz, A., & Cotterill, S. (2015). The "Ascent of Man": A theoretical

Effects of Minimal Grouping on Social Perception and Processing.

and empirical case for blatant dehumanization. *Journal of Personality and Social Psychology*, *109*(5), 901–931. http://doi.org/10.1037/pspp0000048

Kteily, N., Hodson, G., & Bruneau, E. (2016). They See Us as Less Than Human:
Metadehumanization Predicts Intergroup Conflict via Reciprocal Dehumanization. *Journal* of Personality and Social Psychology, 110(3), 343–370.

http://doi.org/10.1037/pspa0000044

- Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences of the United States* of America, 98(26), 15387–15392. http://doi.org/10.1073/pnas.251541498
- Leyens, J.-P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez-Perez, A., & Gaunt, R. (2000). The Emotional Side of Prejudice: The Attribution of Secondary Emotions to Ingroups and Outgroups. *Personality and Social Psychology Review*, 4(2), 186–197. http://doi.org/10.1207/S15327957PSPR0402_06
- Leyens, J.-P., Rodriguez-Perez, A., Rodriguez-Torres, R., Gaunt, R., Paladino, M.-P., Vaes, J., & Demoulin, S. (2001). Psychological essentialism and the differential attribution of uniquely human emotions to ingroups and outgroups. *European Journal of Social Psychology*, *31*(4), 395–411. http://doi.org/10.1002/ejsp.50
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47, 1122–1135. http://doi.org/10.3758/s13428-014-0532-5
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, 51, 93–120.

http://doi.org/10.1146/annurev.psych.51.1.93

Effects of Minimal Grouping on Social Perception and Processing.

- McCaslin, M. J. (2011). Is outgroup prejudice fundamental? Exploring intergroup bias in the minimal group paradigm. *Dissertation Abstracts International, B: Sciences and Engineering*, 71(12), 7781.
- Mekawi, Y., Bresin, K., & Hunter, C. D. (2015). White Fear, Dehumanization, and Low
 Empathy: Lethal Combinations for Shooting Biases. *Cultural Diversity and Ethnic Minority Psychology*.
- Molenberghs, P., & Morrison, S. (2014). The role of the medial prefrontal cortex in social categorization. *Social Cognitive and Affective Neuroscience*, 9(3), 292–296. http://doi.org/10.1093/scan/nss135
- Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status an integration. *European Journal of Social Psychology*.
 http://doi.org/10.1002/ejsp.2420220202
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at Age 7:
 A Methodological and Conceptual Review. Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes, 265–292.

http://doi.org/10.1016/j.mrfmmm.2009.01.007

Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology*, *105*(2), 171–92. http://doi.org/10.1037/a0032734

Pinter, B., & Greenwald, A. G. (2011). A comparison of minimal group induction procedures. Group Processes & Intergroup Relations, 14(1), 81–98. http://doi.org/10.1177/1368430210375251

Postmes, T., Haslam, S. A., & Jans, L. (2012). A single-item measure of social identification:

Effects of Minimal Grouping on Social Perception and Processing.

Reliability, validity, and utility. *The British Journal of Social Psychology*, *52*, 597–617. http://doi.org/10.1111/bjso.12006

- Prati, F., Crisp, R. J., Meleady, R., & Rubini, M. (2016). Humanizing Outgroups Through Multiple Categorization. *Personality and Social Psychology Bulletin*, 42(4), 526–539. http://doi.org/10.1177/0146167216636624
- 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(2), 298–301. http://doi.org/10.1016/j.jesp.2012.10.017
- Ratner, K. G., Dotsch, R., Wigboldus, D. H. J., Knippenberg, A. Van, & Amodio, D. M. (2014).
 Visualizing Minimal Ingroup and Outgroup Faces: Implications for Impressions, Attitudes, and Behavior. *Journal of Personality and Social Psychology*, *106*(6), 897–911.
 http://doi.org/10.1037/a0036498
- 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(7), 750–755. http://doi.org/10.1093/scan/nss063
- Rossion, B., & Caharel, S. (2011). ERP evidence for the speed of face categorization in the human brain : Disentangling the contribution of low-level visual cues from face perception. *Vision Research*, *51*(12), 1297–1311. http://doi.org/10.1016/j.visres.2011.04.003
- Rudman, L. A., & Mescher, K. (2012). Of Animals and Objects: Men's Implicit Dehumanization of Women and Likelihood of Sexual Aggression. *Personality and Social Psychology Bulletin*, 38(6), 734–746. http://doi.org/10.1177/0146167212436401
- Scroggins, W. A., Mackie, D. M., Allen, T. J., & Sherman, J. W. (2016). Reducing Prejudice With Labels: Shared Group Memberships Attenuate Implicit Bias and Expand Implicit

Effects of Minimal Grouping on Social Perception and Processing.	11
Group Boundaries. Personality and Social Psychology Bulletin, 42(2), 219-229.	
http://doi.org/10.1177/0146167215621048	
Senholzi, K. B., & Ito, T. A. (2013). Structural face encoding: How task affects the N170's	
sensitivity to race. Social Cognitive and Affective Neuroscience, 8(8), 937–942.	
http://doi.org/10.1093/scan/nss091	
Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and	
intergroup behaviour. European Journal of Social Psychology.	
http://doi.org/10.1002/ejsp.2420010202	
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(3), 321-335. http://doi.org/10.1177/0146167208327743	
Van Bavel, J. J., & Cunningham, W. A. (2010). A social neuroscience approach to self and so	ocial
categorisation: A new look at an old issue. European Review of Social Psychology, 21(1),
237-284. http://doi.org/10.1080/10463283.2010.543314	
Van Bavel, J. J., & Cunningham, W. A. (2012). A Social Identity Approach to Person Memor	y:
Group Membership, Collective Identification, and Social Role Shape Attention and	
Memory. Personality and Social Psychology Bulletin, 38(12), 1566–1578.	
http://doi.org/10.1177/0146167212455829	
Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The Neural Substrates of In-Gr	oup
Bias. Psychological Science, 19(11), 1131–1139. http://doi.org/10.1111/j.1467-	
9280.2008.02214.x	
Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2011). Modulation of the Fusiform Fa	ce
Area following Minimal Exposure to Motivationally Relevant Faces: Evidence of In-gro	up

Effects of Minimal Grouping on Social Perception and Processing.

Enhancement (Not Out-group Disregard). *Journal of Cognitive Neuroscience*, 23(11), 3343–3354. http://doi.org/10.1162/jocn a 00016

- Waytz, A., Schroeder, J., & Epley, N. (2014). The Lesser Minds Problem. In P. Bain, J. Vaes, &J. Leyens (Eds.), *Humanness and Dehumanization* (pp. 49–67). New York, NY: Psychology Press. Retrieved from
 - http://books.google.com/books?hl=en&lr=&id=Nb7eAQAAQBAJ&oi=fnd&pg=PA49&dq =The+Lesser+Minds+Problem+waytz&ots=bqL8Je6jYF&sig=JOQ5MXbqqcZ21_HK6hxlalQl2s%5Cnpapers3://publication/uuid/1F8C1FFA-403A-40C1-9330-38EA2029C330
- Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological Psychology*, 89(1), 137–147. http://doi.org/10.1016/j.biopsycho.2011.10.002
- Wilde, V. K., Martin, K. D., & Goff, P. A. (2014). Dehumanization as a Distinct Form of Prejudice. *TPM*, 21(3), 301–307.
- Willadsen-Jensen, E. C., & Ito, T. A. (2006). Ambiguity and the timecourse of racial perception. *Social Cognition*, 24(5), 580–606. http://doi.org/https://doi.org/10.1521/soco.2006.24.5.580
- Xu, K., Nosek, B. A., Greenwald, A. G., Ratliff, K. A., Bar-Anan, Y., Umansky, E., Banaji, M.,
 Lofaro, N., & Smith, C. (2016, August 15). Project Implicit Demo Website Datasets.
 Retrieved from osf.io/y9hiq

Good

or Asteroid



Figure 1. Representative trials of positivity bias IAT (faces are from the Chicago Face Database, but were not included in the study).





- 58
- 59 60



Figure 2. Order and duration of stimulus presentation in one trial. Numbers above indicate inter-stimulus time with a fixation "x" on the screen (times varied randomly between 500 and 800 milliseconds. Each face was presented in random order once per block. There were four blocks.

Black Outgroup

White Ingroup

White Outgroup



http://mc.manuscriptcentral.com/gpir







Figure 5. Difference scores between response times, with higher values indicating greater association of that stimulus with human concepts. Figure created using Raincloud Plots (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2018).

256x172mm (72 x 72 DPI)



Figure 6. Grand averaged ERP waveforms at electrode Cz, Pz, and left and right posterior electrodes and Oz for all group/race combinations.

ERP	Time (ms)	Electrodes	Target	M	SD
N170 - right	150-250	TP9, P7, PO9	Black Ingroup	-2.31	2.84
			Black Outgroup	-2.55	2.66
			White Ingroup	-1.09	2.62
			White Outgroup	-1.25	2.98
N170 - left	150-250	TP10, P8, PO10	Black Ingroup	-0.97	2.62
			Black Outgroup	-1.4	2.61
			White Ingroup	-0.36	2.78
			White Outgroup	-0.25	2.97
P200	150-250	Cz	Black Ingroup	0.67	2.51
			Black Outgroup	0.91	2.21
			White Ingroup	-0.13	2.32
			White Outgroup	0.29	2.52
N200	200-350	Cz	Black Ingroup	0.078	2.53
			Black Outgroup	0.28	2.39
			White Ingroup	-0.73	2.28
			White Outgroup	-0.65	2.52
P3	350-700	Pz	Black Ingroup	1.31	2.61
			Black Outgroup	1.07	2.86
			White Ingroup	0.69	2.48
			White Outgroup	1.68	2.2

Pages griptine statistics for Event Balgood Bates atial inter Baby Relations

Table 2

Means, standard deviations, and correlations with confidence intervals

Variable	М	SD	1	2	3	4	5	6	7	8
1. Symbolic Racism	1.16	0.50								
2. Minimal Group	3.88	1.31	.37**							
Commitment			[.14, .57]							
3. Outgroup	0.52	0.04	20*	24						
Emotions	0.53	0.94	29*	24						
			[51,05]	[46, .01]						
4. Positivity	0.35	0.34	.10	.35**	10					
D-30010			[15, .34]	[.11, .55]	[34, .16]					
5 Dehumanization										
D-score	0.22	0.42	15	.03	02	01				
			[38, .11]	[22, .28]	[27, .23]	[26, .24]				
6. P300 Black	1.31	2.61	41*	13	.11	20	01			
Ingroup			[65,11]	[43, .20]	[22, .41]	[49, .13]	[33, .31]			
7 D200 D1 1										
7. P300 Black Outgroup	1.07	2.86	50**	24	.04	23	02	.51**		
6 1			[70,21]	[52, .09]	[29, .35]	[52, .09]	[34, .30]	[.22, .71]		
8. P300 White	0.69	2.48	22	01	.21	11	.04	.53**	.62**	
Ingroup			[51, .10]	[33, .31]	[11, .50]	[41, .22]	[29, .35]	[.26, .73]	[.37, .78]	
0. 0200 011 4										
9. P300 White Outgroup	1.68	2.20	23	12	.06	19	08	.33*	.51**	.41*
			[51, .10]	[42, .21]	[26, .37]	[48, .14]	[39, .24]	[.01, .59]	[.23, .71]	[.10, .64]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates p < .05. ** indicates p < .01.

Supplemental Materials

Neural Resonance

Background

The similar neural activation patterns for actual and observed experiences—called neural resonance—has been linked to better performance in tasks that require affect sharing and basic forms of empathy (Carr & Winkielman, 2014; Gutsell & Inzlicht, 2012; Pineda & Hecht, 2009), understanding actions and emotions (Avenanti, Candidi, & Urgesi, 2013; Borgomaneri, Gazzola, & Avenanti, 2014), and interpersonal coordination, cooperation and helping (Hein, Silani, Preuschoff, Batson, & Singer, 2010; Hoenen, Schain, & Pause, 2013; Knoblich & Sebanz, 2006). However, people show diminished resonance to out-group relative to in-group members' experiences (Avenanti, Sirigu, & Aglioti, 2010; Gutsell & Inzlicht, 2010; Xu, Zuo, Wang, & Han, 2009), which have been linked to biases in helping behavior (Hein et al., 2010) and hostile social interactions (Levy et al., 2016). Recent results suggest that this resonance bias is not purely a result of perceived physical differences, but is affected by context, targets' invisible traits, and observer's own attitudes (Aragón, Sharer, Bargh, & Pineda, 2014; Gutsell & Inzlicht, 2010, 2012, 2013; Hogeveen, Chartrand, & Obhi, 2014; Simon, Styczynski, & Gutsell, under review; Varnum, Blais, & Brewer, 2016). Neural resonance in the sensory motor system is commonly measured using EEG, specifically suppression of the mu-rhythm at central electrode sites (see Fox et al., 2015, for a meta-analysis of 85 studies and Hobson & Bishop, 2016, for a critical perspective). EEG mu-suppression seems to be influenced by group membership but has yet to be used in a minimal group paradigm. Therefore, we do not know whether this basic

neural activity underlying social processing and empathy responds more to visually apparent or memorized groupings.

Analysis

To obtain EEG mu-power during video observation, artifact-free epochs of 2 seconds where extracted and then divided into smaller segments of .25s overlapping by 25% to minimize data loss. We performed fast Fourier transformation using a Hamming window and calculated power in the 8-13 Hz band. We created mu-suppression index scores for each condition by calculating difference scores of the natural log of mu-power during that condition minus the natural log of mu-power during baseline, which was the average mu during the two-second white noise presentation before each face.

Results

One-sample t-tests showed significant suppression below baseline in mu-power for all group and race combinations at C3 (ps < .001), suggesting that motor resonance occurred during the ball squeezing videos (average d = 1.32).

A 2 (race: Black or White) x 2 (minimal group: in- or out-) repeated-measures ANOVA found only a marginal main effect of race, F(1, 38) = 3.00, p = .091, $\eta^2 = .073$, 95% CI[0, .26], no effect of minimal group (p = .55), and no interaction (p = .92), failing to confirm our preregistered hypothesis. The lack of significant race difference is still notable relative to previous findings (Gutsell & Inzlicht, 2010, 2013), and participants resonated marginally more to actions attributed to Black faces ($M = -.19 \ \mu V^2$, SD = .14) than to White faces ($M = -.18 \ \mu V^2$, SD = .15). We also tested whether suppression was unique to the left sensorimotor cortex using an unregistered 2 (race: Black or White) x 2 (minimal group: in- or out-) x 2 (electrode posteriority: central or occipital) x 3 (electrode lateralization: left, central, or right) repeated measures

ANOVA using mu-suppression difference scores at C3, Cz, C4, O1, Oz, and O2 (adapted from Hobson & Bishop, 2016). This test showed a significant race effect, F(1, 38) = 5.25, p = .028, $\eta^2 = .12$, 95% CI[0, .32], showing more suppression for Black ($M = -.19 \mu V^2$) than White faces ($M = -.18 \mu V^2$), but no minimal group effect, F(1, 38) = 1.83, p = .18, $\eta^2 = .046$, 95% CI[0, .22], or race by minimal group interaction, F(1, 38) = .12. p = .74, $\eta^2 = .003$, 95% CI[0, .11]. The effect of lateralization was significant, F(2, 76) = 7.28, p = .001, $\eta^2 = .16$, 95% CI[.03, .30], as was the interaction between posteriority and lateralization, F(2, 76) = 3.70, p = .029, $\eta^2 = .089$, 95% CI[0, .21]. The effect of posteriority was not significant, p > .73, nor were the other interactions (all ps > .15). Examination of the means shows the greatest suppression of the 8-13 Hz band at C4, suggesting that our effect did not lateralize as well as expected but that the effect we see most likely emanates from the sensorimotor cortex.

Symbolic racism correlated with mu-power below baseline at C3 for White targets (r = .37, 95% CI[.06, .62], t(37) = 2.44, p = .02) and in-group targets (r = .35, 95% CI[.04, .60], t(37) = 2.30, p = .027), with more prejudiced participants resonating less. Simple linear regressions found that symbolic racism was a significant predictor of mu-suppression for White in-group faces, $\beta = .084, p = .013$, adjusted $r^2 = .13$.

Discussion

Participants did not show a racial bias in neural resonance. Instead, participants resonated with Black targets, potentially more than with White targets, an unlikely finding in the absence of a manipulation effect. There was also a marginally significant interaction of group and race, calling for further investigation. While we did not see a direct effect of grouping, the absence (and even inversion) of racial bias suggests that knowledge of social categories could affect or potentially supersede perceptual differences. Previous work, however, used samples in different

contexts and richer social stimuli (Gutsell & Inzlicht, 2010, 2013), or measured responses to pain (Levy et al., 2016; Perry, Bentin, Bartal, Lamm, & Decety, 2010), which may be another reason that our results differ. The poor lateralization of our effect, meanwhile, was likely due to the unstandardized nature of the ball-squeezing videos: the target hands in the videos are extended at various angles, making it difficult to identify the hands as right or left.

Self/Other IAT

The Self/Other IAT tested the association of in- and out-group photos with words related to the self (*"I," "me," "myself," "self"*) and words related to humans (*"other," "them," "they," "themselves"*; Conner & Barrett, 2005). This was intended to be a manipulation check that participants implicitly associated themselves with their minimal groups. Participants categorized words as "Other" or "Self" and faces as "Asteroid" or "Comet."

Results

The ANOVA showed the expected significant main effect of minimal group, F(1, 61) =37.04, p < .001, $\eta^2 = .38$, 95% CI[.19 .52], no effect of race, F(1, 61) = 2.53, p = .12, $\eta^2 = .040$, 95% CI[0, .17], and no interaction between group and race, F(1, 61) = .61, p = .44, $\eta^2 = .010$, 95% CI[0, .11]. Simple comparisons found a significant difference between response times for White in-group (M = 183.95 ms, SD = 354.20 ms) versus White out-group targets (M = .203.88ms, SD = 326.10 ms), F(1, 61) = 29.24, p < .001, $\eta^2 = .32$, 95% CI[.14, .48], and for Black ingroup (M = 133.30 ms, SD = 352.83 ms) versus Black out-group targets (M = .217.14 ms, SD =368.80 ms), F(1, 61) = 27.27, p < .001, $\eta^2 = .31$, 95% CI[.13, .46]. Thus the minimal in-group was seen as more implicitly associated with the self.

Moderation

To investigate whether individual commitment to the group was driving participants' minimal group biases despite orthogonal race categorizations, we tested the moderating effect of minimal and ethnic group commitment. While these tests are exploratory, our 62 participants do exceed recommendations for 50 level-two observations for adequate power in multilevel modeling (Maas & Hox, 2005). However, sensitivity analysis of our positivity bias IAT using simulated data (Lane & Hennes, 2018) based on the observed effect size (β = .32 for the interaction of minimal group commitment and minimal group) and observed variance of subject slopes and within-subject residuals showed us to have only .71 power at our sample size. We therefore seem to be underpowered for these analyses and results should be treated cautiously.

To test moderation, we fit hierarchical linear models for IAT response times and infrahumanization, treating each repeated measure as a level one variable and each subject as level two. First, we fitted a null model, then added minimal group and race (or group and emotion type for infrahumanization) as dummy-coded fixed effects ('1' or '0'). We then tested whether these effects varied randomly within subject, finding that minimal group was random for both IATs, but not infrahumanization. We then tested the interaction of minimal group and race, then tried interactions with each type of commitment, which were mean-centered. Ethnic group commitment did not predict any of our dependent variables. At each step of modeling we conducted a linear hypothesis test to test whether the new model was an improvement, arriving at an optimal model. See Tables S1 and S2 for details on the models at each step. We z-scored all continuous variables in the models to report standardized coefficients. We estimated predictors' significance levels via the Satterthwaite method in the 'ImerTest' package in R (Kuznetsova, Brokhoff, & Christensen, 2017). For infrahumanization, due to the large number of zeroes in the data we treated emotion attribution as a binary variable, with any assigned emotions designated

'1' versus '0' when none were assigned. We then fit a binomial logit model and estimated predictors' significance levels with the lht() function in the 'car' package (Fox & Weisberg, 2011). Symbolic Racism did not moderate any of our dependent variables.

Results

Minimal group (β = .56, *t*(61) = 3.24, *p* = .0019) significantly predicted implicit dehumanization, but neither minimal group nor race commitment moderated the effect.

The optimal model to predict implicit bias found significant effects of minimal group (β = .98, t(92.12) = 5.81, p < .001), race (β = -.22, t(122) = -1.99, p = .049), and minimal group commitment (β = -.21, t(60) = -2.23, p = .030), a marginal interaction of minimal group and race (β = .28, t(122) = 1.82, p = .072), and a significant interaction of minimal group commitment and minimal group (β = .32, t(60) = 2.15, p = .036). Minimal group commitment therefore moderated implicit bias, such that participants with stronger minimal group commitment more strongly associated their minimal in-group with positive words (see Figure S2).

We did not conduct a moderation analysis on P3 because we do not have a sufficient sample size. All of the moderation results should be seen as suggestive evidence requiring further research as they were not pre-registered and *p* values are between .005 and .05 (Lakens et al., 2018).

Discussion

Commitment to one's group increases bias in the group's favor, at least for implicit associations and infrahumanization. Minimal group commitment also correlated with racial group commitment, r = .30, t(60) = 2.46, p = .017, lending further evidence to the theory that some people may be high-identifiers (e.g. Jetten, Spears, & Postmes, 2004). While the minimal group memberships of our targets were salient, participants' personal commitment to their own minimal group determined reactions on the basis of that categorization. This suggests that the interaction of category salience and category commitment influences behavior toward a multiply-categorized social target and aligns with Social Identity Theory's suggestion that internalizing the group in one's self-concept is required for bias. It also matches results from similar paradigms (Van Bavel & Cunningham, 2012). Re-categorizing thus seems to be a way to spread favoritism to a different set of targets, rather than expanding whom is disliked. These findings are promising for anti-bias research because they suggest that re-categorization creates eziez preferences, not antipathies.

Supplemental References:

- Aragón, O. R., Sharer, E. A., Bargh, J. A., & Pineda, J. A. (2014). Modulations of mirroring activity by desire for social connection and relevance of movement. *Social Cognitive and Affective Neuroscience*, 9(11), 1762–1769. http://doi.org/10.1093/scan/nst172
- Avenanti, A., Candidi, M., & Urgesi, C. (2013). Vicarious motor activation during action perception: beyond correlational evidence. *Frontiers in Human Neuroscience*, 7(May), 1–8. http://doi.org/10.3389/fnhum.2013.00185
- Avenanti, A., Sirigu, A., & Aglioti, S. M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current Biology*, 20(11), 1018–1022. http://doi.org/10.1016/j.cub.2010.03.071
- Borgomaneri, S., Gazzola, V., & Avenanti, A. (2014). Transcranial magnetic stimulation reveals two functionally distinct stages of motor cortex involvement during perception of emotional body language. *Brain Structure & Function*. http://doi.org/10.1007/s00429-014-0825-6
- Carr, E. W., & Winkielman, P. (2014). When mirroring is both simple and "smart": how mimicry can be embodied, adaptive, and non-representational. *Frontiers in Human Neuroscience*, 8. http://doi.org/10.3389/fnhum.2014.00505
- Conner, T., & Barrett, L. F. (2005). Implicit self-attitudes predict spontaneous affect in daily life. *Emotion*, *5*(4), 476–488. http://doi.org/10.1037/1528-3542.5.4.476

Fox, N. A., Bakermans-Kranenburg, M. J., Yoo, K. H., Bowman, L. C., Cannon, E. N.,
Vanderwert, R. E., ... Van Ijzendoorn, M. H. (2015). Assessing Human Mirror Activity
With EEG Mu Rhythm: A Meta-Analysis. *Psychological Bulletin*.
http://doi.org/10.1037/bul0000031

Gutsell, J. N., & Inzlicht, M. (2010). Empathy constrained: Prejudice predicts reduced mental

simulation of actions during observation of outgroups. *Journal of Experimental Social Psychology*, *46*(5), 841–845. http://doi.org/10.1016/j.jesp.2010.03.011

- Gutsell, J. N., & Inzlicht, M. (2012). Intergroup differences in the sharing of emotive states: Neural evidence of an empathy gap. *Social Cognitive and Affective Neuroscience*, 7(5), 596–603. http://doi.org/10.1093/scan/nsr035
- Gutsell, J. N., & Inzlicht, M. (2013). Using EEG Mu-Suppression to Explore Group Biases in
 Motor Resonance. In B. Derks, D. Scheepers, & N. Ellemers (Eds.), *Neuroscience of Prejudice and Intergroup Relations* (pp. 278–298). London: Psychology Press.
- Hein, G., Silani, G., Preuschoff, K., Batson, C. D., & Singer, T. (2010). Neural Responses to Ingroup and Outgroup Members' Suffering Predict Individual Differences in Costly Helping. *Neuron*, 68(1), 149–160. http://doi.org/10.1016/j.neuron.2010.09.003
- Hobson, H. M., & Bishop, D. V. M. (2016). Mu suppression A good measure of the human mirror neuron system? *Cortex*, *82*, 290–310. http://doi.org/10.1016/j.cortex.2016.03.019
- Hoenen, M., Schain, C., & Pause, B. M. (2013). Down-modulation of mu-activity through empathic top-down processes. *Social Neuroscience*, 8(5), 515–524. http://doi.org/10.1080/17470919.2013.833550

Hogeveen, J., Chartrand, T. L., & Obhi, S. S. (2015). Social Mimicry Enhances Mu-Suppression During Action Observation. *Cerebral Cortex*, (August), 1–7. http://doi.org/10.1093/cercor/bhu016

Jetten, J., Spears, R., & Postmes, T. (2004). Intergroup distinctiveness and differentiation: A meta-analytic integration. *Journal of Personality and Social Psychology*, 86(6), 862–879. http://doi.org/10.1037/0022-3514.86.6.862

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. Current

Directions in Psychological Science, *15*(3), 99–104. http://doi.org/10.1111/j.0963-7214.2006.00415.x

Lakens, D., Adolfi, F. G., Albers, C. J., Anvari, F., Apps, M. A. J., Argamon, S. E., ... Zwaan, R.A. (2018). Justify Your Alpha. *Nature Human Behaviour*, *2*, 168–171.

Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships*, 35(1), 7–31. http://doi.org/10.1177/0265407517710342

Levy, J., Goldstein, A., Influs, M., Masalha, S., Zagoory-Sharon, O., & Feldman, R. (2016).
 Adolescents growing up amidst intractable conflict attenuate brain response to pain of outgroup. *Proceedings of the National Academy of Sciences*, *113*(48), 13696–13701.
 http://doi.org/10.1073/pnas.1612903113

Lüdecke, D (2018). _sjPlot: Data Visualization for Statistics in Social Science_. doi: 10.5281/zenodo.1308157 (URL: http://doi.org/10.5281/zenodo.1308157), R package version 2.6.1, <URL: https://CRAN.R-project.org/package=sjPlot>.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient Sample Sizes for Multilevel Modeling. *Methodology*, 1(3), 86–92. http://doi.org/10.1027/1614-1881.1.3.86

Perry, A., Bentin, S., Bartal, I. B.-A., Lamm, C., & Decety, J. (2010). "Feeling" the pain of those who are different from us: Modulation of EEG in the mu / alpha range. *Cognitive, Affective, & Behavioral Neuroscience, 10*(4), 493–504. http://doi.org/10.3758/CABN.10.4.493

Pineda, J. A., & Hecht, E. (2009). Mirroring and mu rhythm involvement in social cognition: Are there dissociable subcomponents of theory of mind? *Biological Psychology*, 80(3), 306– 314. http://doi.org/10.1016/j.biopsycho.2008.11.003

Van Bavel, J. J., & Cunningham, W. A. (2012). A Social Identity Approach to Person Memory:

-		
7	L	
_		

1	
3	Group Membership, Collective Identification, and Social Role Shape Attention and
5	Memory. Personality and Social Psychology Bulletin, 38(12), 1566–1578.
7 8	http://doi.org/10.1177/0146167212455829
9 10	Varnum, M. E. W., Blais, C., & Brewer, G. A. (2016). Social class affects Mu-suppression
11 12	during action observation. Social Neuroscience, $11(A)$, 140 , 154
13 14	during action observation. Social Iven oscience, 11(4), 449–454.
15 16	http://doi.org/10.1080/17470919.2015.1105865
17 18	Xu, X., Zuo, X., Wang, X., & Han, S. (2009). Do you feel my pain? Racial group membership
19 20	modulates empathic neural responses. The Journal of Neuroscience, 29(26), 8525-8529.
21 22	http://doi.org/10.1523/JNEUROSCI.2418-09.2009
23 24	
25 26	
27 28	
29 30	
31 32	
33 34	
35 36	
37 38	
39 40	
41 42	
43 44	
45 46	
47	
49	
50 51	
52 53	
54 55	
56 57	
58	
59 60	http://mc.manuscriptcentral.com/gpir

Supplemental Figures and Tables



Figure S1. Log mu power difference scores from average C3 baseline for the four target types at electrode C3. More negative numbers indicate greater suppression of mu wave oscillations and therefore greater resonance. All four are significantly different from baseline, suggesting the presence of mu suppression in response to observed action.



Positivity Bias IAT Response Time by Minimal Group Commitment for each Target Type

Figure S2. Positivity Bias IAT Response Time by Minimal Group Commitment for each Target Type. Colors indicate target type for both the raw data points and the model fit lines. Response times are differences between good pairings and bad pairings, but times were inverse-rooted so higher values indicate greater implicit bias. Minimal group commitment is mean centered.



Figure S3. Secondary Emotions by Minimal Group Commitment Controlling for Symbolic Racism. Circle size indicates the number of observations at that point; color indicates whether the assignment is to the ingroup or the outgroup, both for the plotted raw data and the model fit lines. Secondary emotions are just tallied from the three possible choices for each group; primary emotions are omitted from the figure because the pattern is the same and secondary emotions are those of interest. Minimal group commitment is mean centered.

Secondary Emotions by Minimal Group Commitment Controlling for Symbolic Racism

		Main Effects			Interaction			Moderation
Predictors	Estimates	CI	р	Estimates	CI	р	Estimates	CI
Intercept	-0.52	-0.730.32	<0.001	-0.45	-0.670.23	<0.001	-0.45	-0.670.24
White	-0.08	-0.23 - 0.08	0.3 <mark>2</mark> 1	-0.22	-0.430.00	0.046	-0.22	-0.430.00
Outgroup	1.12	0.82 - 1.43	<0.001	0.98	0.64 - 1.32	<0.001	0.98	0.65 - 1.31
White Outgroup				0.28	-0.02 - 0.58	0.069	0.28	-0.02 - 0.58
Minimal Group Commitment							-0.21	-0.400.03
Minimal Group Commitment x Outgroup							0.32	0.03 - 0.62
Random Effects								
σ2	0.37			0.37			0.37	
τ ₀₀	0.41 Subj	ect		0.41 Subj	ect		0.37 _{Subj}	ect
τ ₁₁	1.12 Subj	ect.groupOut		1.13 Subj	ect.groupOut		1.04 Subj	ect.groupOut
Q ₀₁	-0.96 Sub	ject		-0.96 Sub	ject		-0.96 Sub	ject
ICC	0.52 Subj	ect		0.53 Subj	ect		0.51 Subj	ect
Observations	248			248			248	
Marginal R2 / Conditional R2	0.315/0	.630		0.320/0	.637		0.346/0	.640

Table S1. Linear modeling of response times on Positivity Bias IAT. Table created using или или 1.

"sjPlot" package in R (Lüdecke, 2018).

http://mc.manuscriptcentral.com/gpir

	M	ain Effects	Moderation			
Predictors	Odds Ratios	CI	р	Odds Ratios	CI	р
Intercept	0.77	0.33 - 1.77	0.534	0.75	0.31 – 1.79	0.51
Secondary Emotions	0.69	0.35 - 1.38	0.298	0.68	0.34 - 1.38	0.28
Outgroup	0.48	0.24 - 0.97	0.042	0.48	0.23 - 0.98	0.04
Minimal Group Commitment				1.61	0.73 – 3.58	0.23
Minimal Group Commitment x Outgroup)			0.41	0.20 - 0.87	0.02
Random Effects						
σ2	3.29			3.29		
τ ₀₀	4.99 Subject			5.47 Subject		
ICC	0.60 Subject			0.62 Subject		
Observations	248			248		
Marginal R2 / Conditional R2	0.020 / 0.61	1		0.040 / 0.64	10	

Table S2. Binomial logit modeling of emotional attribution with dependent variable being '1' =emotion assigned, '0' = emotion not assigned. Table created using "sjPlot" package in R (Lüdecke, 2018).