Journal of Experimental Psychology: General

Is This My Group or Not? The Role of Ensemble Coding of Emotional Expressions in Group Categorization

Amit Goldenberg, Timothy D. Sweeny, Emmanuel Shpigel, and James J. Gross

Online First Publication, July 18, 2019. http://dx.doi.org/10.1037/xge0000651

American Psychological Association

Journal of Experimental Psychology: General

0096-3445/19/$12.00 http://dx.doi.org/10.1037/xge0000651

Is This My Group or Not? The Role of Ensemble Coding of Emotional Expressions in Group Categorization

Amit Goldenberg
Stanford University

Timothy D. Sweeny
University of Denver

Emmanuel Shpigel and James J. Gross
Stanford University

When exposed to others' emotional responses, people often make rapid decisions as to whether these others are members of their group or not. These group categorization decisions have been shown to be extremely important to understanding group behavior. Yet, despite their prevalence and importance, we know very little about the attributes that shape these categorization decisions. To address this issue, we took inspiration from ensemble coding research and developed a task designed to reveal the influence of the mean and variance of group members' emotions on participants' group categorization. In Study 1, we verified that group categorization decreases when the group's mean emotion is different from the participant's own emotional response. In Study 2, we established that people identify a group's mean emotion more accurately when its variance is low rather than high. In Studies 3 and 4, we showed that participants were more likely to self-categorize as members of groups with low emotional variance, even if their own emotions fell outside of the range of group emotions they saw, and that this preference is seen for judgments of both positive and negative group emotions. In Study 5, we showed that this unique preference for low group emotional variance is special to group categorization and does not appear in a more basic face categorization task. Our studies reveal unexplored and important tendencies in group categorization based on group emotions.

Keywords: categorization, ensemble coding, summary statistical perception, emotions, social cognition

Supplemental materials: http://dx.doi.org/10.1037/xge0000651.supp

Imagine yourself attending a party in your new place of work. You do not know anyone at the party. Right after you arrive, the host gives a toast, and during his speech, makes a political slur. You immediately scan the audience to read people's emotional responses, trying to take in their facial expressions to determine whether or not your new coworkers share your views and thus are "your group" or "not your group." This type of complex categorization process is as common as it is important. We often try to decipher information about our social environment by aggregating others' responses and then comparing them to our own.

Understanding how people do this requires an investigation of two different types of computations. The first involves the ability to take in collective information, in this case multiple expressions of emotions, and rapidly aggregate them into relevant summary statistics. This type of cognitive process has been researched under the heading of ensemble coding (Alvarez, 2011; Whitney, Haberman, & Sweeny, 2014; Whitney & Yamanashi Leib, 2018). The second involves a group categorization process in which one assesses whether certain social information and one's own response fall under the same or different social categories. This question has been investigated under the heading of group categorization (for a recent review, see Rule & Sutherland, 2017), and more specifically, in research related to self-categorization theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; Turner & Reynolds, 2011).

Although there has been increasing interest in investigating the intersection of ensemble coding and social psychology (Dannals & Miller, 2017; Lamer, Sweeny, Dyer, & Weisbuch, 2018; Phillips, Slepian, & Hughes, 2018), there have not been any studies, to the best of our knowledge, examining how people categorize these ensemble coding evaluations as either belonging to their group or not. This points to the importance of the current investigation, which highlights the social utility of ensemble coding while grounding social categorization in basic visual mechanisms.

Ensemble Coding of Emotions

Observing other people and making sense of their behavior plays an important role in social functioning. We navigate crowds

Amit Goldenberg, Department of Psychology, Stanford University; Timothy D. Sweeny, Department of Psychology, University of Denver; Emmanuel Shpigel and James J. Gross, Department of Psychology, Stanford University.

We thank Yiyue (Kiki) Cao, Sime Luketa, Maria Reyes and Naila Ebeid for their assistance in this project.

Correspondence concerning this article should be addressed to Amit Goldenberg, who is now at Negotiations, Organizations and Markets, Harvard Business School, Baker Library, Bloomberg, 437 Boston, MA 02163. E-mail: chagag@gmail.com

on our daily commute, observe groups' responses in shows, sporting events, and demonstrations, and incorporate this information into our evaluations of our environment. Consistent with the idea that group perception plays a particularly important role in social interactions, research shows that people gaze longer at groups compared to individuals (Woolhouse & Lai, 2014), and that people's attention is more impacted by seeing groups than by seeing individuals gazing in a specific direction (Milgram, Bickman, & Berkowitz, 1969; Woolhouse & Lai, 2014).

However, information from groups is very rich, and the capacity for visual representation is limited. This introduces a bottleneck for seeing information at the level of the group, particularly in contexts where real-time adjustments in one's own behavior are required. Two thousand years ago, Aristotle suggested that people create summary representations to understand visual information at the gist level (for review see Whitney et al., 2014). In modern psychology, this idea has been validated and systematically investigated under the heading of ensemble coding. Findings from this literature have shown that people can form summaries of many types of complex visual objects quickly and efficiently (Alvarez, 2011; Haberman, Brady, & Alvarez, 2015; Hubert-Wallander & Boynton, 2015).

Recently, there has been an increasing interest in ensemble coding of faces (Rhodes et al., 2018) and especially of faces that express emotions (Elias, Dyer, & Sweeny, 2017; Haberman & Whitney, 2009; Ji, Rossi, & Pourtois, 2018; Whitney et al., 2014). Emotions are a unique source of information as they provide relatively clear and strong signals of people's thoughts, intentions, and future actions. Indeed, people are highly sensitive to the emotions of others (Goldenberg, Garcia, Zaki, et al., 2019; Hatfield, Cacioppo, & Rapson, 1994; Parkinson, 2011) and often use these emotions to make decisions regarding the environment (Campos, Hiatt, Ramsay, Henderson, & Svejda, 1978) and other people (Van Kleef, 2009).

Much of the research done on identifying groups' emotions (and ensemble coding in general) has provided evidence of people's sensitivity to the mean emotional expressions of multiple faces, whether they are positive or negative. One clear expectation is that the mean of a group's emotions is important to any categorization process. A second feature that may play a role in identifying crowds' emotions is the variance of these emotions, and indeed there is some work showing that people are sensitive to the variance of complex visual features (Michael, de Gardelle, & Summerfield, 2014) and even emotions (Haberman, Lee, & Whitney, 2015). What is not yet clear, however, is how these summary statistics, either independently or together, influence group categorization.

Self-Categorization and Group Membership

As we navigate our social worlds, one pressing question is whether others are members of our group or not (Rule & Sutherland, 2017; Turner & Reynolds, 2011). Just like ensemble coding, group categorization requires that individuals process complex information in efficient and generalizable ways (Bruner, 1957). Consistent with the idea that categorization is a fundamental social process, research shows that group categorization emerges early in development (Kelly et al., 2007) and can occur unintentionally (Martin & Macrae, 2007).

Deciding whether certain people are part of one's ingroup or outgroup has important implications for a variety of social processes. People are motivated to perceive their group in a positive light, as suggested by research related to social identity theory (Tajfel & Turner, 1979; Turner & Oakes, 1986). As a result, categorizing oneself as a member of a certain group has motivational, emotional, and cognitive consequences such as increased attention to one's group (Ellis, Deregowski, & Shepherd, 1975), increased conformity (Abrams & Hogg, 1990) and ingroup bias (Mullen, Brown, & Smith, 1992).

In the context of emotions, people judge the emotional expressions of ingroup members as more positive than those of outgroup members (Lazerus, Ingbretsen, Stolier, Freeman, & Cikara, 2016). They also tend to express similar emotions to their ingroup (Bourgeois & Hess, 2008; Lin, Qu, & Telzer, 2018; Weisbuch & Ambady, 2008) and different emotions from their outgroup (Cikara, Bruneau, & Saxe, 2011; Cikara, Bruneau, Van Bavel, & Saxe, 2014; Cikara & Fiske, 2013; Lau, Morewedge, & Cikara, 2016). The ingroup–outgroup distinction plays a key role in determining individual emotional responses and therefore may be an important driver of social processes driven by emotions such as intergroup conflicts (Halperin, 2014), collective action (van Zomeren, Leach, & Spears, 2012), and polarization (Iyengar, Lelkes, Levendusky, Malhotra, & Westwood, 2018).

Although the consequences of the ingroup–outgroup categorizations are well known, how group categorization actually unfolds is much less clear. One theory that has focused on this question is self-categorization theory (SCT), which was developed out of questions related to categorization that arose in the wake of social identity theory (Turner et al., 1987; Turner & Reynolds, 2011). According to SCT, group categorization occurs frequently and at different levels of abstraction (e.g., family member, female, American). These levels of abstraction are emphasized based on the individual's fit to these categories and their cognitive accessibility (Oakes, 1987; Oakes, Turner, & Haslam, 1991).

What mechanisms underlie actual group categorization decisions? In addressing this question, Turner suggested the *metacontrast ratio*, estimated by dividing the average distance *between* a certain individual and a certain group by the average distances *within* that group (Turner et al., 1987, p. 47). Turner does not provide a detailed description of exactly how these distances are evaluated, or what the metacontrast values are that lead to categorization. The ratio is therefore a theoretical principle, contributing to the insight that smaller distances between an individual and their group are likely to increase the chance of categorization. Adopting this insight, a few studies have used the concept of the metacontrast ratio to conceptualize the degree of group categorization and its effect on various group-related outcomes (Hohman, Gaffney, & Hogg, 2017; Mummendey, Otten, Berger, & Kessler, 2000). Yet, an evaluation of the metacontrast ratio as a predictive measure of categorization decisions has not been conducted.

Although taking the metacontrast approach is helpful, it also raises a few important questions. Take, for example, a comparison of two cases in which the intensity of an individual's emotional response is exactly at the mean of two distributions of group emotional responses on a neutral-to-angry continuum (see Figure 1A). One of these distributions has low variance, whereas the other has high variance. If we calculate the metacontrast ratio by estimating the average distance using, for example, sum of squares,

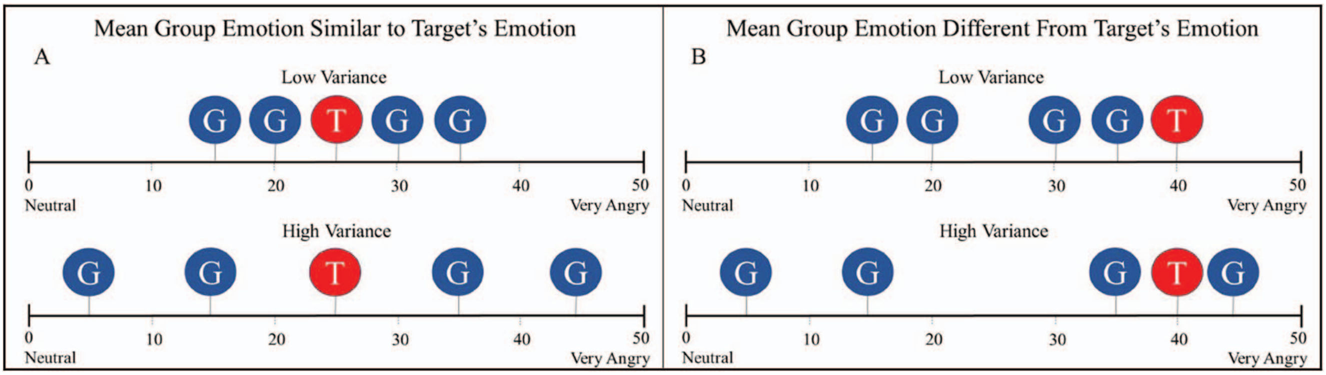

Figure 1. Situations for group-categorization that will be examined in Studies 1–4. The circles represent people's degree of anger in response to a stimulus. Target emotions are marked with T while the group emotions are marked with G. In Panel A we illustrate a case in which the target's emotion is similar to the average group emotion, and the variance group emotion is either low (top) or high (bottom). In Panel B, we illustrate a case in which the target's emotion is different from the average group emotion, and the variance of the group emotion is either low (top) or high (bottom). See the online article for the color version of this figure.

we find that the ratio is similar in both cases. However, it is intuitive to think that the lower variance case should increase the probability for group categorization. This example also reflects the intuition that not only is the average distance between a person's emotion and the group mean emotion important, but also that the variance should be an important signal for one's group categorization decisions.

Consistent with the emphasis on group variance, one factor that seems to play an important role in categorization is group entitativity (Campbell, 1958). *Entitativity* is generally described as the degree to which a collection of persons are perceived as being bonded together in a coherent unit (Lickel et al., 2000). In the context of group members' emotional responses, high entitativity can be understood as the degree of similarity between group members' emotions (the inverse of the standard deviation of group members' emotional responses). Research has established that group members prefer to self-categorize as members of groups with high entitativity, compared to low (Hogg et al., 2007; Lickel et al., 2000). A potential reason for this preference is that groups with high entitativity are easier to understand and predict and therefore provide the benefit of reducing subjective uncertainty (Hogg, 2004; McGregor, Zanna, Holmes, & Spencer, 2001), as suggested by uncertainty reduction theory (Hogg, 2000). Group members not only prefer to be in homogenous groups but also tend to judge their own ingroups as more homogenous (ingroup homogeneity effect, see Simon & Pettigrew, 1990). This preference emphasizes the importance of group variance in group categorization decisions.

Group entitativity seems to be an important factor for understanding group categorization. What is unclear is what happens if we pit the tendency to prefer entitativity against the tendency to prefer a mean close to one's own emotion. Take, for example, a case in which a person can choose to be either a member of a coherent group (low variance), but one that does not include the person's own emotions (Figure 1B, top), or a less coherent group (higher variance) but one that includes the person's emotions (Figure 1B, bottom). With which group would the person prefer to be categorized? Answering this question has important consequences for how people relate to groups that express emotions that are different from their own.

Is Group Categorization Just Another Form of Categorization?

The attempt to integrate the domains of ensemble coding and group categorization raises an important question: Is group categorization just a simple matter of matching one's own emotions with those of a group, or is there a difference between group-categorization decisions and other face-categorization decisions? We can imagine two seemingly similar tasks that would allow us to examine this question. The first is a *group categorization task*, in which participants provide their emotional response to a certain stimulus (e.g., by choosing a facial expression that reflects their emotion), then observe others' emotional responses to the same picture, and finally, based on these responses, make a decision about whether these other people are from their group or not. The second is a simple *face categorization task*. In this task, participants see a target facial expression, then they see a group of facial expressions and are asked to decide whether the target face was taken from the group or not. In both tasks, participants seemingly do the same thing, which is to categorize a facial expression into a group. We believe, however, that the two tasks should lead to quite different outcomes. Although the face categorization task is essentially a visual task, the group categorization task involves reacting emotionally to a stimulus and using this reaction as a basis for deciding which responses are included in one's group. Making these group categorization decisions reflects not only whether one is technically part of the group but also if one wants to be part of that group.

People's preference to be in certain groups compared to others can be driven by their affinity for coherent groups, as suggested by research on entitativity (Lickel et al., 2000). Coherent groups are easier to predict, as the behavior of their members is less variable, and therefore membership in those groups reduces uncertainty (Hogg, 2000, 2004; McGregor et al., 2001). In cases in which the mean group emotion is aligned with the individual's emotions, group coherence also means that more people in the group express emotion similarly to the individual. Take, for example, a case when participants' emotional expression is similar to that of certain groups (Figure 1A). We hypothesize that in such situations group members should prefer to categorize themselves as members of a group with lower variance compared to higher variance. However, when merely asked to indicate whether a face expressing emotion is part of a certain sample distributed around that emotion, it is not clear that participants should be more likely to associate the sample with the low-variance group compared to the high-variance group, as in both cases the target face is taken from within the range of each sample. In the different group emotion case (Figure 1B), the target face will only appear in the high variance group, compared to the low variance group, and therefore the target face should be more likely to be categorized as part of the group in the high variance condition. However, it is unclear what participants would choose in such a situation during a group categorization task. The current framework allows us to examine these points of divergence and convergence.

The Present Research

The goal of the present research was to examine the influence of the mean and variance of group emotions on group categorization. In Study 1, we examined the effect of a group's mean emotions on participants' group categorization and we verified that group categorization is more likely when the mean group emotion is closer to participants' own emotion. Study 2 laid the groundwork for manipulating variance by testing whether participants accurately evaluated the group's mean emotions when the group's variance was low versus high. Study 3 was the crucial test of this investigation, in which we manipulated both the group mean and variance and measured how these features influenced participants' group categorization. In Study 4 we examined whether the effects of Study 3 could be replicated with positive stimuli rather than with negative stimuli. Finally, Study 5 was designed to differentiate between our group categorization results and results of a simple face categorization task. We modified our categorization task such that participants did not respond emotionally to stimuli, but merely reported whether a single face was taken from a face sample or not.

These studies were motivated by several preregistered hypotheses (https://osf.io/tmc7f/). In terms of a group's mean emotions, we hypothesized that a larger distance between the individual's emotions and the group's mean emotions would lead to a reduction in the probability of group categorization. In terms of variance, predicting how it might impact endorsements of group membership seemed less straightforward. In cases in which the group's mean emotion is similar to that of the individual's emotion, we hypothesized that larger variance in that group's emotions would lead to a reduction in the probability of group categorization (Figure 1A). This is because such a group would include people whose emotions are more distant from the individual's emotions. However, when the group's mean is different from that of the individual, it was unclear to us whether an increase in group variance would lead to an increase or a decrease in categorization (Figure 1B). On the one hand, when group variance is high there would be more group members with emotion similar to that of the individual. On the other hand, there would also be more people with emotions that are different from that of the individual. Therefore, we had no clear hypothesis on the probability of group categorization in these cases. Finally, we hypothesized that people's group categorization decisions would be different than their categorization decisions in a face categorization task. We estimated that participants' preference of low variance should not influence their decisions in the face categorization task. We therefore hypothesized that in conditions in which the target face and group have the same mean emotion, we should see no difference in the likelihood of categorizing the target face as a member of the low and high variance groups. However, when the single face's emotion is different from the mean of the group, participants should be more likely to categorize the target face as a member of the high variance group, as this target face is indeed more likely to come from that group (Figure 1B).

Study 1: The Effect of Group Mean on Categorization

The goal of Study 1 was to examine whether participants' group categorization was influenced by the mean group emotion and its distance from participants' own emotions.

Method

Participants. As we did not know the effect size of the manipulation, we used the data from the first 10 participants from our sample to estimate the required sample size for the study (as indicated in our preregistration). For these 10 participants, we compared the difference in group categorization between the same-mean condition and different-mean condition (similar to the actual analysis, see below). We then used the results to conduct a simulated power analysis for 30 participants using the R package simr (Green & Macleod, 2016). Results suggested that using 30 participants would be enough to obtain more than 80% power for the study (see the online supplementary materials for power estimations based on each sample size). Based on this calculation, our final sample included 30 participants, which included the 10 participants that were used to evaluate the effect size (eight men, 22 women; $M_{\text{age}} = 25.96$, $SD = 13.44$). All of the participants were American citizens. Participants were recruited using the Stanford paid pool, which includes a mix of Stanford students and community members. Participants received $8 for their participation in the study. No participants were removed from the analysis.

Procedure. Along with all subsequent studies described here, Study 1 received research ethics committee approval (protocol number 7273, Stanford University) prior to the collection of data. Participants were told that they were taking part in a study testing how emotions influence group categorization. The procedure started with a minimal group paradigm, which is a classic manipulation designed to prime participants with the notion of group membership. Participants were asked to answer a few binary questions regarding general personal preferences. For example, participants were asked to decide whether they preferred cats or dogs, hamburgers or pizza, big cities or small towns, and so forth. This procedure was adapted from minimal group manipulations that were used in other studies (see e.g., Levy, Saguy, van Zomeren, & Halperin, 2017). After making their choices, participants were told that they had been assigned to one of two groups based on their preferences. Participants were also told that the two groups were equal in size and that the groups were comprised of all of the previous participants in the experiment. This explanation was designed to create a sense of an ingroup and an outgroup for participants.

Participants then completed our group categorization task (after instructions and a practice run) which was adapted from a previous ensemble coding task (Elias et al., 2017). The task included 50 trials. In each trial participants first observed a picture depicting a case of immoral behavior by an American official. For example, some of the pictures were taken from the Abu Ghraib incident in which American army personnel were caught abusing prisoners of war. Other pictures were of American soldiers threatening children and women in combat zones. Results from a previous study suggested that Americans who observed these pictures experienced strong anger in response to these pictures (Goldenberg, Garcia, Halperin, et al., 2019). Participants observed each picture for 5 s (Figure 2A). They then viewed a screen containing a face (Figure 2B). Moving the mouse from left to right gradually transformed the face's expression from neutral to angry (see sample of the scale in Figure 3).

Participants were asked to express their own anger in response to the picture by modifying the face. The neutral-to-angry continuum was created by concatenating 50 modifications of the same face, from no anger to extremely angry, effectively creating a 1–50 (translating to 1–100%) anger scale. The complete range of faces was created for a previous investigation of ensemble face perception (Elias et al., 2017) using exemplar faces from the NimStim face set (Tottenham et al., 2009). We used four different facial identities that were randomly chosen for each trial. Importantly, all of the faces were of White males (who made up the majority of our participants). However, analysis suggested that the measures were not affected by participants' race or gender (this was also consistent with previous studies that used the same sample, see the online supplementary materials). The ratings participants were allowed to select from to indicate their own emotions during the task was purposefully limited, and they ranged only from the 10th face to the 40th face in the 1–50 scale. The reason for this was to allow a normal distribution of faces to be formed around participants' rating, thereby enabling the groups to include individuals expressing more or less emotion than the face selected by the participant. Participants had no time limit when rating their anger in response to the picture. Although we assumed that participants' responses would be more likely to represent their emotional experience in response to the stimuli than their actual facial expression, we thought that using faces would allow us to measure participants' emotional responses and compare them to those of a group.

After providing their own ratings, participants saw 12 faces at once, each expressing different intensities of emotions. The identities of the 12 faces were identical to the identity of the face which was used by participants to rate the image from that same trial. The 12 faces appeared on the screen for 2 s (Figure 2C). We chose to use a 2-s time window, which is longer than some ensemble coding tasks, for a few reasons. First, as this was our first test of the effect of group emotion on group categorization, we wanted to use a longer exposure time to increase the chance of finding differences between conditions. Second, social cognition studies have often used longer exposure times as these studies generally have been less interested in finding the minimal visual conditions for group cognition than in detecting their impact (Dannals & Miller, 2017; Phillips et al., 2018). As this was our first study in this domain, we shared the same goals. Finally, determining how participants ar-

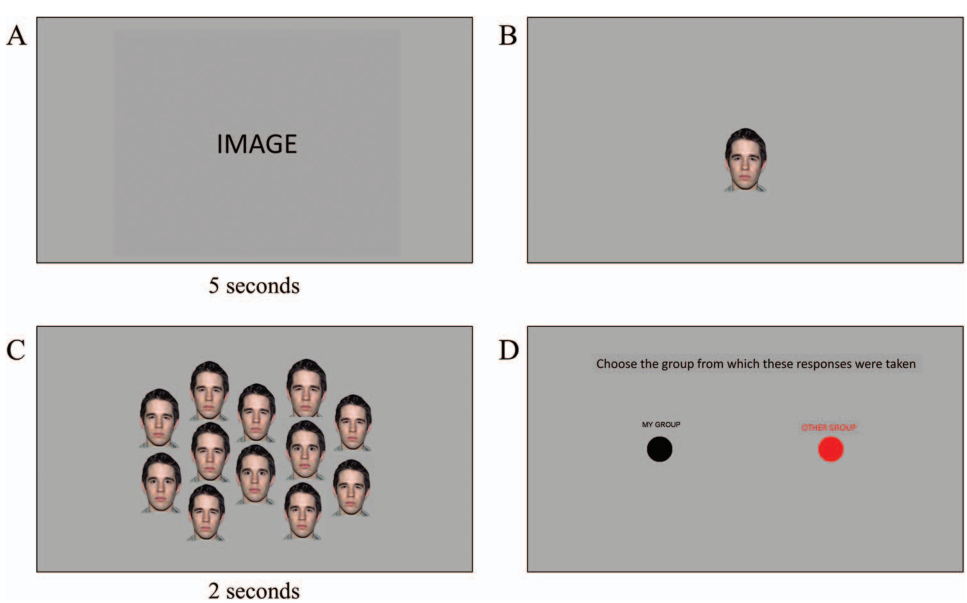

Figure 2. The group categorization task used in Study 1. Participants first saw a picture of Americans behaving immorally (A). They were then asked to rate their degree of anger in response to the picture by modifying the face on the screen (B). Participants then saw what they were told were 12 ratings from other participants (C) and were then asked to choose whether the faces were taken from their own group or from another group (D). Faces are from the the NimStim Set of Facial Expressions (Tottenham et al., 2009). See the online article for the color version of this figure.

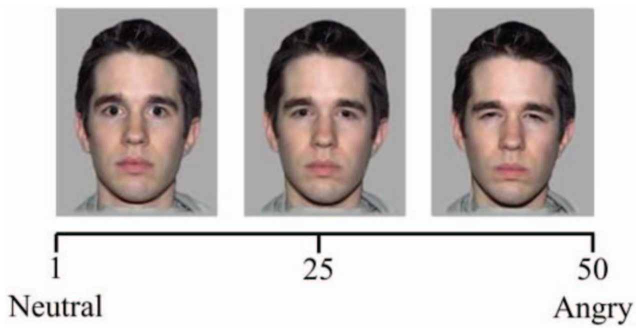

Figure 3. A sample of three faces from the anger scale that was used in the studies, from neutral (left) to angry (right). Values of 25 and 50 correspond to 50% and 100% intensities in our morph range. Faces are from the the NimStim Set of Facial Expressions (Tottenham et al., 2009). See the online article for the color version of this figure.

rived at their evaluations of a group's summary statistical properties was not our primary goal (although Study 2 does provide evidence that ensemble coding occurred). We were instead focused on the way participants used these summary representations to inform their own group categorizations.

The faces were presented against a uniform gray background (red/green/blue value = 170, 170, and 170, respectively; luminance = 27.5 cd/m2). The location of the 12 pictures was independently jittered by 1 to 15 pixels in either direction along the horizontal and vertical axes. Without taking into account this random position variation, the centroids of adjacent faces were 10.8° away from each other along the horizontal axis and 8.1° away from each other along the vertical axis. Participants were told that these 12 faces represented 12 ratings of participants who had already completed the task. They were also told (in the instructions to the task) that these 12 ratings were all taken from either their own group or from the other group (based on the minimal group paradigm).

The standard deviation of anger depicted in the faces on the screen was always 10 points (based on the 1–50 anger scale, 1 being neutral and 50 very angry). However, the mean intensity of anger of the faces was randomly assigned to be either the same as participants' own rating (the same-mean condition) or different from participants' own rating (either 10 points higher or lower; the different-mean condition). The standard deviation of each distribution was always 10 points and the distribution of the sample was designed to be uniform. Not all of participants' own ratings permitted their random assignment to either lower or higher ratings in the different-mean condition. This is because we wanted to make sure that the shape of the distribution for all conditions was close to normal. Therefore, if a participant's rating was higher than 30 on a given trial, they were randomly assigned either to (a) the same-mean group emotion condition or (b) the different-mean group emotion condition in which the group's mean emotion was lower than their own. If a participant's rating was lower than 20, they were randomly assigned to the same-mean condition or a variant of the different-mean condition where the group's mean emotion was higher than their own. Although we suspected that there might be differences in group categorization between the low and the high variants of the different-mean condition, we found no such differences (see the online supplementary materials). We therefore examined the different-mean condition with data collapsed across these variations.

After viewing the 12 faces for 2 s, participants were presented with group-categorization screen, in which they were instructed to "choose the group from which these responses were taken" (Figure 2D). This instruction, like the initial task instructions, was designed to remind participants that they were choosing from the two groups of participants who completed the minimal group paradigm—their group and the other group. Participants then made a binary choice between these two options. They had as much time as they needed to make this choice.

Measures. As previously mentioned, participants' ratings were based on 50 concatenated faces, from 1 (*neutral*) to 50 (*very angry*; see Figure 3). Participants group categorization was measured as a binary variable. In addition to these measures, participants were asked to indicate their age, gender, race, and education level. Finally, participants indicated their political affiliation on a 1 (*very liberal*) to 7 (*very conservative*) scale as well as a group identification scale. These scales were not used in the main analysis of the article (but are reported in the online supplementary materials).

Results and Discussion

We analyzed the data from our task by using a mixed generalized linear model, treating the group mean (similar or different) as our independent variable and participants' binary choice of group as a dependent variable. We also used a by-participant random variable. Using a by-stimulus random variable did not improve the model and was therefore not used in the analysis. Results suggested that the probability of group categorization significantly increased when the group's mean ratings were similar to participants' own ratings compared to when they were different ($b = -.68$ [$-.92$, $-.45$], $SE = .12$, $z = -5.69$, $p < .001$, $R^2 = .13$, intraclass correlation coefficient [ICC] = .10, Figure 3). R^2 value represents the conditional R^2 (including both fixed and random effects), based on recommendations from Nakagawa and Schielzeth (2013). ICC values were calculated using the R package sjstats (Lüdecke, 2015) based on recommendations from Nakagawa, Johnson, and Schielzeth (2017). These results supported our first hypothesis that increase in the distance between a person's own emotion and that of a group will reduce the probability of group categorization. Interestingly, the probability of group categorization in the different-mean condition was also significantly higher than chance (represented by the line in Figure 4, see the online supplemental material for analysis), suggesting that even in this case participants were more likely to categorize the group as their own group.

Results of Study 1 confirm that the difference between a group's mean emotion and one's own emotion influences the probability of group categorization. Our next goal was to examine whether the variance of the group may also play a role in group categorization. However, before we could evaluate our primary question regarding how a group's mean and variability interact to impact group categorization, we had to first examine how changes in variability influence evaluations of the mean.

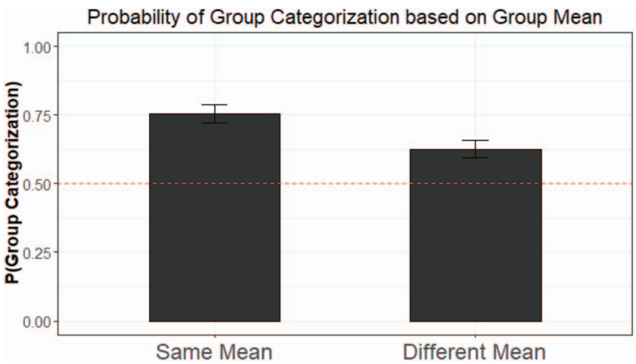

Figure 4. The probability of group categorization as a function of the group mean in relation to the individual in Study 1. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

Study 2: Identifying Mean Emotions of Groups With Different Standard Deviations

The goal of this study was to test how participants evaluated the group's mean emotion when both the mean and the variance of group emotion were manipulated. This was done to examine whether (and if so, how) these two summary statistics interact in the context of the current experimental design, with the hope of gaining further insight into the mechanisms of group categorization (tested in Study 3). Our hope was that participants' ability to evaluate the mean group emotions of certain groups would provide insight into whether participants would be more likely to prefer to be members of these groups. This idea is inspired by subjective uncertainty reduction theory (Hogg, 2000), which suggested that people prefer to be members of groups they can understand and predict.

Method

Participants. We recruited 30 participants for the study, as in Study 1 (14 men 14, 16 women; $M_{\text{age}} = 24.76$, $SD = 12.33$). Again, participants were recruited from the Stanford paid pool and received \$8 for their participation in the study. No participants were removed from the analysis.

Procedure. Participants were told that they were taking part in a study that was designed to test whether people can identify the mean emotional expressions of groups. Participants then completed an adapted version of our group categorization task used in Study 1 (after instructions and a practice run). The task included 50 trials. In each trial, participants first observed a picture depicting a case of immoral behavior by an American official, similar to those in Study 1. Participants observed each picture for five seconds. Participants then saw 12 faces expressing different intensities of anger for two seconds. The group's mean emotion was randomly generated to be between 10 and 40 (based on the 1–50 scale of angry faces, 1 being neutral and 50 very angry). We limited the range of the group means in order to allow distributions that were as close to uniform as possible within the 1–50 scale. We also manipulated the standard deviation of anger intensities depicted in the groups to be either 5 or 10 points (the *low* and *high* conditions, respectively). After viewing the group of 12 faces for two seconds, participants then viewed a single face presented on the screen. Moving the mouse from left to right transformed the face from a neutral face to an angry face (on a 1–50 scale, see Study 1). Participants were asked to adjust the face in order to capture the mean emotion expressed by the group. They had as much time as they needed to estimate each group's mean intensity of emotion.

Measures. As previously mentioned, participants' responses were based on 50 concatenated faces, from 1- neutral to 50 - very angry (see Figure 2 for a sample). In addition to these measures, participants were asked to indicate their age, gender, race and education level. Finally, participants indicated their political affiliation on a 1–7 scale (1 – very liberal, 7 – very conservative) as well as a group identification scale. These scales were not used in the main analysis of the article (results are reported in the online supplementary materials).

Results and Discussion

We analyzed the data from our task using a linear mixed-model. The group mean and standard deviation were our independent variables. As the dependent variable, we looked at the discrepancy between participants' evaluation of the group mean and the actual group mean emotion. As an index of this discrepancy, we created a difference score between participants' estimation of the mean and the actual mean group emotion. Positive values indicated an overevaluation of the group's average emotional intensity and negative values indicated an underevaluation of the group's average emotional intensity (we present these findings in Figure 4).

To examine the effect of group mean and variance, we evaluated the interaction between the group's mean emotion (10–40) and variance. Looking first at the intercept of the interaction equation, results suggested that at the middle of the scale, participants tended to evaluate the mean emotion at 2.89 [1.88, 3.42] points above the actual mean ($SE = .50$, $t(29) = 5.74$, $p < .001$), pointing to a general tendency to overestimate mean anger. This bias is congruent with previous findings that suggest a bias to report neutral faces as appearing angry (Neta & Whalen, 2010). Furthermore, results suggested a significant interaction between the mean group emotion and the standard deviation of the group, such that both extremes of the scale in the high variance trials were more likely to be evaluated as closer to the middle of the scale compared to the low variance trials ($b = -.08$ [$-.92$, $-.45$], $SE = .02$, $t(1467) = -3.82$, $p < .001$ $R^2 = .29$, ICC $= .11$, Figure 5A). In other words, when the intensity of the mean group emotion was low, participants overrated the mean group emotion, and the magnitude of this overrating was greater in the high variance condition. When the intensity of the mean group emotion was high, participants underrated the mean group emotion, and the strength of this underrating was greater in the high variance condition. A second way to examine these data is by comparing the absolute value of the difference between the estimated and real mean for both the high and low variance trials. We used the absolute value of the difference as an indication of the error in order to make sure that the overand underratings do not cancel each other out. Indeed, results suggested that error in the high variance trials was significantly greater than in the low variance trials ($b = -.90$ [-1.46, $-.35$], $SE = .28$, $t(1469) = -3.22$, $p < .001$, $R^2 = .17$, ICC $= .16$, Figure 5B).

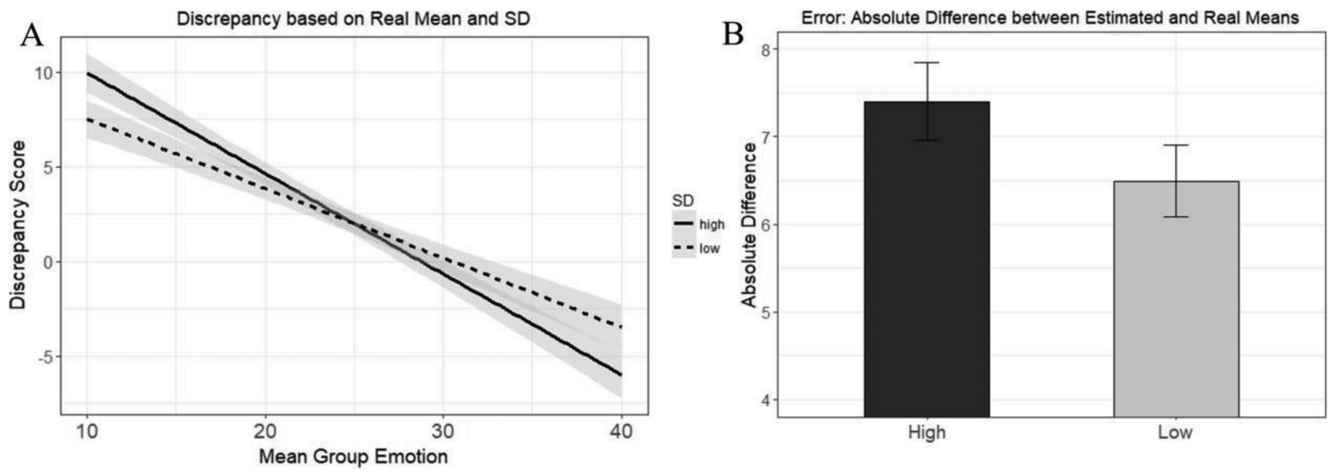

Figure 5. Panel A shows the effects of mean group emotional intensity (*x*-axis) and mean group *SD* (legend) on two related outcome variables: the discrepancy score, which represents the difference between participants' evaluation of the mean and the real mean. The gray area around the lines represents the standard error. Panel B shows the difference in error, calculated as the absolute difference between the estimated and read means. Error bars are 95% confidence intervals.

Results of Study 2 indicate that participants estimated the mean emotion of groups with low emotional variance more accurately than groups with high variance. This pattern is congruent with other examinations of ensemble coding (Sweeny, Haroz, & Whitney, 2013). In particular, we found that participants consistently erred in reporting the means of high-variance groups, underrating high mean emotions and overrating low mean emotions. This effect could have been perceptual, such that high-variance groups with extreme emotions actually appeared to be less intense. Alternatively, difficulty estimating the mean emotion of the high-variance groups may have compelled participants to default to using the middle of the response scale. These interpretations lead to two opposite predictions regarding the probability of group categorization in an experiment that includes a comparison between high and low variance in the different-mean trials. The first prediction is that for the different group emotion trials, participants would tend to perceive (and thus evaluate) high variance groups as closer to the middle of the scale. Such an evaluation would increase the probability that high variance groups would be self-categorized compared to low variance groups. Imagine a case in which a participant rates a certain picture as having an emotion intensity of 25 and is then assigned to the high group mean emotion condition (i.e., mean emotion of 35). In the high variance case, the mean would be perceived as lower compared to the low variance case, increasing the chance for group categorization. An alternative prediction—which is congruent with subjective uncertainty reduction theory (Hogg, 2000) – is that high variance will make it harder for participants to identify a group's mean emotion, and because of this uncertainty, they will tend to less frequently self-categorize as members of these groups. The goal of Study 3 was therefore to assess these two explanations by looking at how both mean and variance of group emotions may interact to influence group categorization.

Study 3: The Effect of Group Mean and Variance on Categorization

The goal of Study 3 was to test the effect of changes in group mean and variance on group categorization.

Method

Participants. We recruited 30 participants for the study, as in Study 1 (10 men, 20 women; $M_{\text{age}} = 27.36$, $SD = 11.36$). Again, participants were recruited from the Stanford paid pool and received \$8 for their participation in the study. No participants were removed from the analysis. Results of Study 3 were also replicated in an online sample of 100 participants (see the online supplementary materials).

Procedure. The procedure for Study 3 was similar to that of Study 1. However, in addition to manipulating the mean group emotion to be either similar (same group mean as participants' ratings) or different (10 points higher or lower than participants' ratings), we also manipulated the standard deviation of emotions expressed in the group to be either low ($SD = 5$) or high ($SD = 10$). On the different-mean group emotion trials, participants' ratings fell within the range of the distribution of faces when the standard deviation was high, but not when it was low. As in Study 1, the dependent variable was participants' group categorization decision.

Results and Discussion

We conducted a mixed generalized linear model looking at the interaction between the group mean (similar vs. different) and the group standard deviation (low vs. high) predicting participants' group categorization decision. We also used a by-participant random variable. Looking first at the interaction, results suggested that the interaction was nonsignificant ($b = .14$ [$-.63$, .35], $SE =$

.25, $z = .56$, $p = .57$, ICC = .14). We therefore further examined the main and simple effects (also see Figure 6).

Looking first at the main and simple effects of the group mean, results show that categorization was higher when the group mean was similar to that of participants, compared to when the group mean was different ($b = -.54$ [−1.35, −.69], $SE = .06$, $z = -8.53$, $p < .001$, $R^2 = .20$, ICC = .14). These results replicate our findings from Study 1. This effect was similar in the low standard deviation group ($b = -.58$ [−.77, −.39], $SE = .09$, $z = -6.12$, $p < .001$, $R^2 = .21$) and the high standard deviation group ($b = -.51$ [−.67, −.34], $SE = .08$, $z = -6.06$, $p < .001$, $R^2 = .18$).

Next, we examined the main effect of standard deviation on categorization decisions. We found a main effect for standard deviation such that categorization was higher for groups with a low standard deviation of emotional intensity regardless of the group mean ($b = -.30$ [−.43, −.18], $SE = .06$, $z = -4.86$, $p < .001$, $R^2 = .24$, ICC = .14). We further examined the simple effects of the differences between the low and high standard deviation conditions for both the similar and different-mean group emotion conditions. For the same-mean condition, group categorization was higher for the low standard deviation trials compared to the high standard deviation trials ($b = -.34$ [−.54, −.14], $SE = .10$, $z = -3.41$, $p < .001$, $R^2 = .11$). These results were congruent with our second hypothesis, which suggested that when the group's mean emotion is similar to participants' own emotions, participants would be more likely to self-categorize to groups with low variance. Importantly, results of the different group emotions trials showed that group categorization was higher in the low standard deviation trials as well, suggesting that even in the different-mean emotion trials, participants preferred to self-categorize themselves as members of more homogeneous groups ($b = -.27$ [−.42, −.12], $SE = .07$, $z = -3.54$, $p < .001$, $R^2 = .01$). Furthermore, it was only in the different-mean trials with high standard-deviations that participants group categorization was not different from chance (represented by the red dotted line, see the online supplementary material for analysis). This points to the possibility that participants were uncertain about whether this was their group on these trials.

Overall, the results of Study 3 provide evidence that a group's emotional variance can have an important influence on a viewer's group categorization decisions. Participants preferred to self-categorize into groups with low emotional variance, both when the group's average emotional intensity was similar or different from participants' own emotions. These results are especially interesting when looking at the different-mean group emotion trials. Participants preferred to self-categorize as members of the low standard deviation groups despite the fact that it was only in the high variance trials that participants' responses were actually located within the range of the distribution. In other words, participants paradoxically indicated that they felt more like a member of the low variance groups even though they would have been the most extreme member of those groups. Importantly, in a robustness test of these findings, results of Study 3 were replicated in a new online sample of 100 participants (Study 3A, see the online supplementary materials).

The findings of the different-mean trials can be explained based on the findings from Study 2, which point to the fact that participants tended to report the mean group emotion of the high variance groups as being closer to the middle of the scale compared to the low variance groups. Additional support for this interpretation lies in the fact that the probability of group categorization for the different-mean, high-variance trails was not different from chance, suggesting that participants struggled to evaluate their categorization in these trials. As argued by subjective uncertainty reduction theory (Hogg, 2000), group members prefer to be members of coherent groups, as these groups are easier to predict and understand. Our findings suggest that this preference holds even when participants' emotions fall outside of the range of the group's emotion.

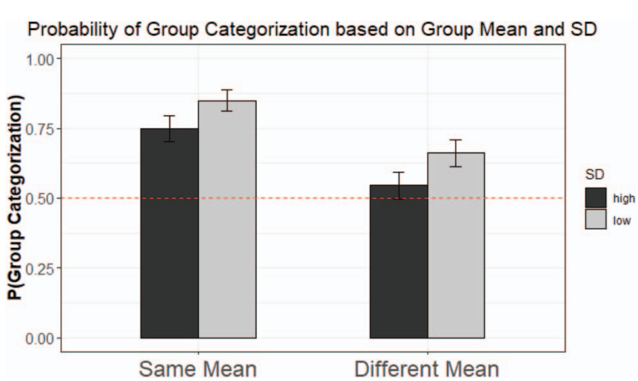

Figure 6. Group categorization based on both mean and standard deviation manipulation in Study 3. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

Study 4: The Effect of Group Mean and Variance on Categorization in Response to Positive Stimuli

The goal of Study 4 was to examine whether the effects found in Study 3 generalize to positive stimuli. Previous work has suggested that people may be more accurate in identifying variance in anger compared to happiness (Ackerman et al., 2006), however these studies examined perception of individual emotions. We therefore were interested to test whether participants' categorization decisions in response to positive emotions would be similar to the ones found in Study 3.

Method

Participants. We recruited 100 participants who completed the study online on Mechanical Turk. Sample size was matched to the one used in our online replication of Study 3. (See Study 3A in the online supplementary materials.) Participants were all Americans located in the United States who completed the task on a computer (rather than a smartphone). Selection criteria was at least 500 hits with a 95% success rate. Participants' received $2 for their participation (average completion time was 21.15 min). The study included a set of attention checks in which participants were asked to describe the content of the pictures that they responded to. Out of the 100 participants who completed the task, we removed one participant for describing the pictures incorrectly and one participant for refusing to fill out a few key demographics (such as location, gender, and age). Three additional participants were removed for making the same categorization decision in all 50 trials. Removing these participants did not change the difference between conditions but merely raised the categorization rate for all conditions (all three participants marked "my group" to all trials). Our final sample was therefore 95 participants (50 men, 45 women; $M_{\text{age}} = 35.41$, $SD = 10.46$).

Procedure. Prior to running the task, we conducted a pilot study to create a sample of positive emotions eliciting primarily happiness in response to group-related situations. These were pictures of people celebrating the 4th of July and of American athletes celebrating having received Olympic medals or winning international competitions. During the piloting phase, we kept pictures that were rated as positive by more than 95% of participants. When asked to choose from a series of distinct emotions, participants predominantly categorized the pictures as eliciting happiness, with excitement being the second most common emotion (for a detailed description, see the online supplementary materials).

With the new picture sample in hand, we then turned to conduct the actual task. The procedure for Study 4 was similar to that of Study 3. However, instead of providing an emotional response from neutral to negative, participants marked their emotional response using faces that depicted happiness on a 10–40 scale (similar to Studies 1 and 3). This scale was a subset of a 1–50 scale (see Figure 7) that was abridged to allow a group distribution to be formed around participants' ratings. Similar to Studies 1–3, participants then observed 12 ratings ostensibly made by other participants. These ranged from neutral to positive. Similar to Study 3, we manipulated the mean group emotion to be either similar (same group mean as participants' ratings) or different (10 points higher or lower than participants' ratings). We also manipulated the standard deviation of emotions expressed in the group to be either low ($SD = 5$) or high ($SD = 10$). On the different-mean group emotion trials, participants' ratings fell within the range of the distribution of faces when the standard deviation was high, but not when it was low. As in Studies 1 and 3, the dependent variable was participants' group categorization decision.

Results and Discussion

We conducted a mixed generalized linear model looking at the interaction between the group mean (similar vs. different) and the group standard deviation (low vs. high) predicting participants' group categorization decision. We also used a by-participant random variable. Looking first at the interaction, results suggested that the interaction was nonsignificant ($b = .01$ [$-.03$, $.07$], $SE = .02$, $z = .60$, $p = .54$, ICC = .12). We therefore further examined the main and simple effects (also see Figure 8).

Looking first at the main and simple effects of the group mean, results show that categorization was higher when the group mean was similar to that of participants, compared to when the group mean was different ($b = -.55$ [$-.62$, $-.48$], $SE = .03$, $z = -15.85$, $p < .001$, $R^2 = .18$, ICC = .12). These results replicate our findings from Studies 1 and 3. The significant difference between the same-mean and different-mean conditions was similar in the low standard deviation group ($b = -.57$ [$-.67$, $-.47$], $SE = .05$, $z = -11.43$, $p < .001$, $R^2 = .19$) and the high standard deviation group ($b = -.53$ [$-.63$, $-.44$], $SE = .04$, $z = -11.13$, $p < .001$, $R^2 = .16$).

Next, we examined the main effect of standard deviation on categorization decisions. We found a main effect for standard deviation such that categorization was lower for groups with a high standard deviation of emotional intensity regardless of the group mean ($b = -13.$ [$-.20$, $-.06$], $SE = .03$, $z = -3.80$, $p < .001$, $R^2 = .18$, ICC = .12). These results were similar to those of Study 3. We further examined the simple effects of the differences between the low and high standard deviation conditions for both the similar and different-mean conditions. For the same-mean condition, group categorization was higher for the low standard deviation trials compared to the high standard deviation trials ($b = -.15$ [$-.26$, $-.04$], $SE = .05$, $z = -2.75$, $p = .01$, $R^2 = .18$). These results were congruent with our second hypothesis and similar to those found in Study 3, which suggested that when the group's mean emotion is similar to participants' own emotions, participants would be more likely to self-categorize to groups with low variance. Importantly, results of the different group emotions trials showed that group categorization was higher in the low standard deviation trials as well, suggesting that even in the different-mean emotion trials, participants preferred to self-categorize themselves as members of more homogeneous groups ($b = -.11$ [$-.19$, $-.02$], $SE = .04$, $z = -2.66$, $p = .01$, $R^2 = .15$). These results were similar to those in Study 3 as well.

Overall, results of Study 4 reiterate the findings of Study 3 and provide further evidence that a group's emotional variance can have an important and surprising influence on a viewer's group categorization decisions. Participants preferred to self-categorize into groups with low emotional variance, both when the group's average emotional intensity was similar or different from participants' own emotions.

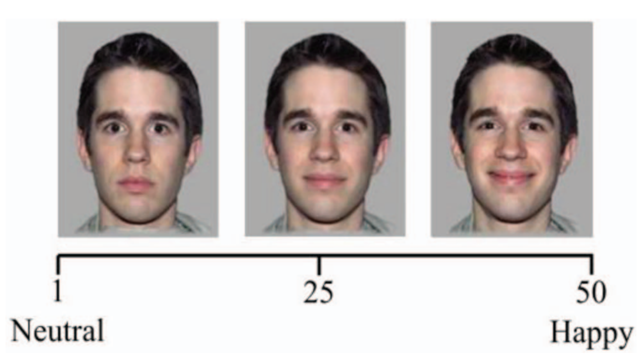

Figure 7. A sample of three faces from the scale that was used in Study 4, from neutral (left) to happy (right). Values of 25 and 50 correspond to 50% and 100% intensities in our morph range. Faces are from the the NimStim Set of Facial Expressions (Tottenham et al., 2009). See the online article for the color version of this figure.

Study 5: Is Group Categorization Just Another Form of Categorization?

One of the questions raised by the results of Studies 3 and 4 is whether our group categorization task is actually just a simple face categorization task in which participants are merely asked to estimate whether a certain facial expression is part of a set of faces or not. In our previous experiments, we assumed that people reflected about their own emotion in response to a stimulus, provided a rating of their emotion with a single response face, then made a judgment about their emotion relative to the group. Cer-

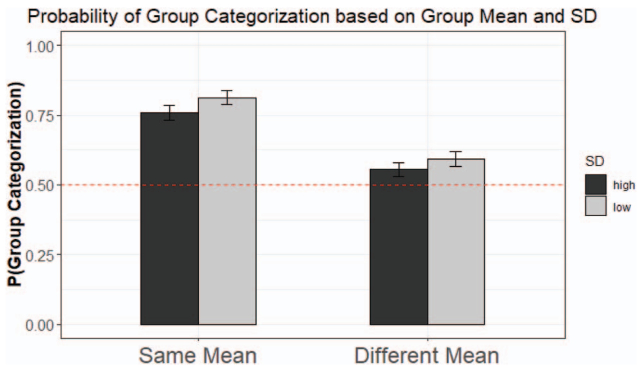

Figure 8. Group categorization based on both mean and standard deviation manipulation in Study 4. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

tainly, participants did reflect on their internal state and they adjusted the response face. But their evaluation of group membership may have simply reflected a purely visual comparison between the appearance of the response face and the appearance of the group, subsequent to and independent of their evaluation of their own emotional state. If this were the case, then a simple matching task between a single face and a group, without any sort of inquiry about the participant's emotional state, should produce the same pattern of results. In the current study, we modified our categorization task to allow us to examine this question. We hypothesized that in the same-mean group emotion condition, we should see no difference between the low and high variance groups, as in both cases the target face was sampled from the group. Furthermore, in the different-mean group emotion condition, we hypothesized that participants should be more likely to categorize the target face as member of the high variance group compared to the low variance, as this target face is indeed more likely to come from that group.

Method

Participants. We recruited 100 participants who completed the study online on Mechanical Turk. Sample size was matched to the one used in our online replication of Study 3 and to Study 4. Participants were all Americans located in the United States who completed the task on a computer (rather than a smartphone). Selection criteria was similar to previous studies. Participants' received \$2 for their participation (average time = 21.15 min, identical to Study 4). The study included a set of attention checks in which participants were asked type in words that appeared on the screen. Out of the 100 participants who completed the task, we removed seven participants for making the same categorization decision in all 50 trials. Removing these participants did not change the difference between the conditions but merely changed the categorization rate for all conditions (six out of the seven participants marked "my group" to all trials). Our final sample was therefore 93 participants (62 men, 31 women; $M_{\text{age}} = 35.17$, $SD = 11.23$).

Procedure. The current task was a modification of the task used for Studies 3 and 4. The task included 50 trials. In each trial participants first observed a target face that was randomly drawn from our 10–40 neutral-to-angry continuum (which was a subset of a 1–50 scale, the shortened scale was designed to allow a distribution to be formed around the target face, similar to all previous studies). The face appeared on the screen for 5 s. To make sure that participants indeed observed the target face, they were then asked to report the degree of emotion expressed by the target face. This was done using a scale of neutral to angry, similar to that of Study 3A, in which participants were asked to mark the location on the scale from which the face they just saw was drawn. After estimating the intensity of the emotion expressed by the target face they just saw, participants saw 12 faces at once, each expressing different intensities of emotions. The identities of the 12 faces were identical to the identity of the target face which they saw earlier. The distribution was also identical to the ones from Studies 3 and 4. The 12 faces appeared on the screen for 2 s, similar to Studies 1–4. We manipulated both the mean group emotion and the standard deviation of the 12 faces based on the target face that participants' first saw. Participants either saw 12 faces with a mean that was similar to the target face or different (10 points higher or lower than the target face). We also manipulated the standard deviation of emotions expressed in the group to be either low ($SD = 5$) or high ($SD = 10$). After viewing the 12 faces, participants were asked to estimate whether the target face was taken from the sample or not. Notice that the structure of the task was identical to that of the group categorization task with one difference: Instead of responding emotionally to a stimulus, participants saw a target face and were asked to report whether it was part of a group of faces or not.

Results and Discussion

Before starting our main analysis of participants categorization choice, we conducted a test to examine that participants were able to accurately estimate the intensity of the target face they saw at the beginning of each trial. We conducted a linear mixed model analysis, looking at the difference between the actual emotional intensity of the target face and the estimated rating from participants. Similar to our previous analysis, we used a by-participant random variable. Results suggested that there was no significant difference between participants' estimation of the actual intensity of the face and the estimated rating ($b = .11$ [$-.03$, $.07$], $SE = .23$, $t(9190) = .48$, $p = .62$, ICC = .16). These findings are encouraging as they suggest that participants' categorization decisions were not biased by their perception of the target face. It also means that participants paid attention to the target face.

To look at participants' categorization decisions, we conducted a mixed generalized linear model looking at the interaction between the group mean (similar vs. different) and the group standard deviation (low vs. high) predicting participants' categorization decisions. We also used a by-participant random variable. Looking first at the interaction, results suggested a significant interaction between group mean and standard deviation ($b = .13$ [.04, .21], $SE = .04$, $z = 3.00$, $p = .001$, $R^2 = .18$, ICC = .01). These results were different from Studies 3 and 4 in which the interaction was not significant. We therefore further examined the main and simple effects to further understand the interaction (also see Figure 9).

Looking first at the main and simple effects of the group mean, results show that group categorization was higher when the group

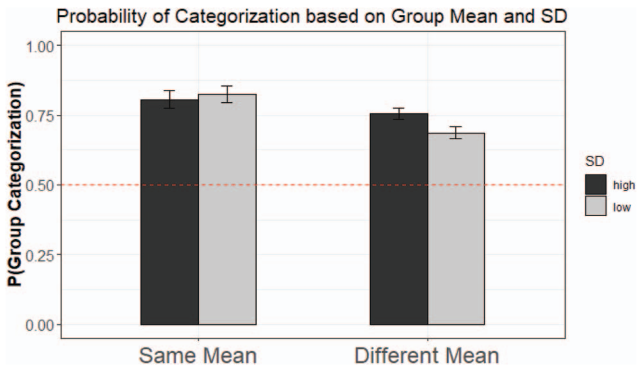

Figure 9. Group categorization based on both mean and standard deviation manipulation in Study 5. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

mean was similar to that of participants, compared to when the group mean was different ($b = -.29$ [$-.38$, $-.21$], $SE = .04$, $z = -6.85$, $p < .001$, $R^2 = .17$, ICC = .15). These results were similar to our findings from Studies 3 and 4 and appeared both when looking at the low standard deviation group ($b = -.42$ [$-.55$, $-.30$], $SE = .06$, $z = -6.92$, $p < .001$, $R^2 = .15$) and the high standard deviation group ($b = -.16$ [$-.29$, $-.04$], $SE = .06$, $z = -2.74$, $p = .01$, $R^2 = .21$). These results suggest that even when participants are categorizing faces (rather than their own emotional responses) their categorization decisions are influenced by the mean group emotion.

Next, we examined the main effect of standard deviation on categorization decisions. Unlike Studies 3 and 4, we did not find a significant main effect for standard deviation such that participants did not categorize the target face differently depending on the standard deviation of the sample ($b = .06$ [$-.02$, .15], $SE = .04$, $z = -1.50$, $p = .13$, ICC = .15). Looking at the simple effect revealed an even more interesting picture. When the group mean was similar to that of the target face, we found no difference in participants' categorization between the low and high variance ($b = -.06$ [$-.20$, $-.06$], $SE = .07$, $z = -.84$, $p = .39$). These results were different from our categorization task in which a significant difference was found. We then examined difference in categorization between high and low group standard deviation in the different-mean group emotion condition. Results suggested that participants were more likely to categorize the target face as part of the high variance group, compared to the low variance ($b = .19$ [.11, .27], $SE = .04$, $z = 4.81$, $p < .001$, $R^2 = .15$). These results are in the opposite direction to the ones found in Study 3 and 4. They point to the fact that participants indeed estimate that the target face was more likely to be within the high variance sample, compared to the low variance.

Overall, the results of Study 5 point to two important differences between our group categorization task and a simple categorization task of facial expression. In the same group emotion condition, the standard deviation of the group had no impact on whether the target face was evaluated as being part of the face sample in the simple categorization task. These findings are different from our group categorization task in which participants preferred to categorize themselves to the low-variance groups. In the different-mean group emotion condition, we found the opposite results relative to the group categorization task. Participants' were more likely to categorize the target face as being part of the high standard deviation group, compared to the low standard deviation group. These results make sense, as the target was actually more likely to be included in this high variance group. However, they also point to the striking difference between this study and our group categorization task, in which participants preferred to be categorized in the low variance group even when their own emotion fell outside of the distribution. We conclude from these findings that group categorization involves other processes than those of the simple face categorization task, and that one clear difference between the two is participants' preference to be members of coherent groups, whether the mean group emotion is similar to, or different from, their own emotion.

General Discussion

One crucial question regarding social perception is how we synthesize and organize complex social information and use this information to address important questions about group membership. Here, we took inspiration from the ensemble coding and group categorization literatures to examine the influence of the mean and variance of a group's emotions on group categorization. Results of Study 1 affirmed that the probability of group categorization decreased when the group mean was different from that of the individual's emotions (compared to when it was the same). Results of Study 2 laid the foundation for manipulating variance by showing that participants were less accurate in evaluating the mean emotion of groups with high variance. In line with participants' ability to understand the mean group emotions, results of Studies 3 and 4 suggested that participants were also more likely to self-categorize to groups with a lower standard deviation of emotional expression compared to a higher standard deviation, regardless of whether the group mean emotion was similar or different to participants' own emotions. Finally, results of Study 5 suggested that participants' preference for low group emotional variance, seen in Studies 3 and 4, was not found when participants completed a simple face categorization task that did not require them to report on their own emotions, thus revealing the uniqueness of group categorization in relation to other, more general categorization tasks. Combined, these studies provide novel and surprising insights into how people make complex and rapid assessments of group membership.

Mechanisms Underlying Group Categorization

Our findings suggest that participants preferred to self-categorize to groups with low emotional variance. This was especially interesting considering the fact that in the high variance condition, participants' own emotions were located within the range of the distribution of emotions expressed by the group, whereas in the low variance group participants' own emotions were outside the range of the distribution. Still, participants preferred to categorize themselves as members of the low variance groups. Based on the findings of Study 2, one mechanism that may play a role in this preference for low variance, especially when thinking about the different-group mean trials, is the fact that it was easier to estimate a group's mean emotion when that group had low variance. When people decide whether or not to self-categorize themselves into certain groups, they may simply prefer groups that are coherent as these groups are easier to understand and predict (Campbell, 1958; Hogg, 2000; Lickel et al., 2000).

Preference for homogeneity should not be inevitable, however. Previous work suggests that when group members—particularly those with low group identity—learned that their group was low status, they actually tended to interpret it as having high variability (Doosje, Ellemers, & Spears, 1995). Thinking of their group as heterogenous may make it easier for certain group members to maintain group membership while still thinking of themselves as different from their group (Brewer, 1991). Further research should be done to examine potential moderators such as group status and threat to group image that may affect our observed preference for group homogeneity.

Implications for Group Processes

The idea that people prefer to be members of coherent groups, even when their own emotions fall outside of the group emotional response, may have important implications for understanding intragroup dynamics. It is well established that extreme groups often present a more coherent message (Hogg, 2007; Hogg et al., 2007; Lickel et al., 2000). If people have a preference for group coherence, as our studies suggest, they may be more attracted to self-categorize as members of these extreme groups.

Attraction to extreme groups based on their coherence can explain processes such as polarization and escalation in group emotions (Goldenberg, Garcia, Halperin, et al., 2019; Iyengar et al., 2018). Group members' tendency to categorize themselves with more coherent groups will lead them to conform to these groups' emotions. Conformity to emotions of more extreme groups may play a role in emotional polarization, as situations that lead to polarized emotional responses will force categorization decisions and therefore conformity to extreme groups (Brady, Wills, Jost, Tucker, & Van Bavel, 2017; Goldenberg, Garcia, Halperin, et al., 2019; Goldenberg, Saguy, & Halperin, 2014). Even if the group emotional responses are not polarized, conforming to more extreme groups can lead to an increase in overall emotions and to escalation (Halperin, 2016; van Zomeren et al., 2012). Further work should examine the connection between this preference for coherence and these broader social processes.

Implications for Well-Being and Social Interactions

Based on the current findings, group categorization may vary even due to minor shifts in visual representations of the emotions of one's group. Considering the high frequency with which people evaluate their surroundings (Bargh, 1994), and considering the fact that one's social environment is in constant flux (e.g., other people's emotional expressions are likely to be dynamic), especially in the current social media era, one's sense of group membership may shift from moment to moment.

Literature on the psychological sense of belonging has suggested that lack of belonging is connected to a variety of negative effects on one's health and well-being (Walton & Brady, 2017). However, this research has often considered belonging as a long-term state of mind rather than a fleeting sensation that may be updated based on momentary perceptual changes (Walton & Cohen, 2007). Thinking about group categorization as a dynamic process raises interesting questions regarding individual differences in the stability of group categorization. We predict that low-stability in one's group categorization may be associated with similar negative outcomes as lack of belonging.

Once individuals self-categorize as members of a group, a variety of group-related processes such as conformity and ingroup favoritism become active (for a review, see Hornsey, 2008). Therefore, it would be interesting to see how the stability and frequency of a person's sense of belonging influence a variety of social processes such as conformity and ingroup bias. It may be that the tendency for frequent categorization leaves people insecure regarding their group membership, and therefore leads them to be more highly influenced by their group. However, it may be that constantly examining one's own group categorization makes it harder to feel part of a group and therefore decreases the probability of social influence processes. Future work should examine these questions.

Limitations and Future Directions

The current research provides important insights into how the mean and variance of a group's emotions influence group categorization. Nevertheless, these studies have limitations that suggest the value of several directions for future research.

First, in the current study, the faces that were presented to participants varied from neutral to very angry or neutral to very happy and did not include any faces that expressed emotions opposite to the expected response from the picture. This decision was motivated by the desire to increase the probability that mean and standard deviation would play a role in the categorization decision. Group categorization choices may be very different when including positive faces in response to negative stimuli or negative faces in response to positive stimuli. It is possible that people may weigh variance within the angry spectrum less when there are also faces expressing positive emotions and the opposite for positive emotions.

Second, our manipulation of mean and variance included a limited range of these two statistics. It may be possible that expanding the degree of variance or mean differences will reveal important new information about group categorization decisions. Learning more about the effects of mean and variance on group categorization will be important for developing a more complete theory.

More generally, the current research considered just one dimension of variation among faces (namely, in emotional expression). One interesting question is whether increasing the diversity in other dimensions such as gender or race would alter the results (Hess, Thibault, Adams, & Kleck, 2010; Johnson, Freeman, & Pauker, 2012). Do people treat variance on different dimensions in the same way, or are some aspects of variance more important than others? Further work should test how these features interact with the findings of this article.

References

Abrams, D., & Hogg, M. A. (1990). Social identification, self-categorization and social influence. *European Review of Social Psychology, 1,* 195–228. http://dx.doi.org/10.1080/14792779108401862

Ackerman, J. M., Shapiro, J. R., Neuberg, S. L., Kenrick, D. T., Becker, D. V., Griskevicius, V., . . . Schaller, M. (2006). They all look the same to me (unless they're angry): From out-group homogeneity to out-group heterogeneity. *Psychological Science, 17,* 836–840. http://dx.doi.org/10.1111/j.1467-9280.2006.01790.x

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15,* 122–131. https://www.sciencedirect.com/science/article/pii/S1364661311000040. http://dx.doi.org/10.1016/j.tics.2011.01.003

Bargh, J. A. (1994). The four horsemen of automaticity: Awareness, intention, efficiency, and control in social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognitions* (pp. 1–40). New York, NY: Psychology Press.

Bourgeois, P., & Hess, U. (2008). The impact of social context on mimicry. *Biological Psychology, 77,* 343–352. http://dx.doi.org/10.1016/j.biopsycho.2007.11.008

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences of the United States of America, 114,* 7313–7318. http://dx.doi.org/10.1073/pnas.1618923114

Brewer, M. B. (1991). The social self: On being the same and different at the same time. *Personality and Social Psychology Bulletin, 17,* 475–482. http://dx.doi.org/10.1177/0146167291175001

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64,* 123–152. http://dx.doi.org/10.1037/h0043805

Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science, 3,* 14–25. http://dx.doi.org/10.1002/bs.3830030103

Campos, J. J., Hiatt, S., Ramsay, D., Henderson, C., & Svejda, M. (1978). The emergence of fear on the visual cliff. In M. Lewis & L. A. Rosenblum (Eds.), *The development of affect* (pp. 149–182). Boston, MA: Springer. http://dx.doi.org/10.1007/978-1-4684-2616-8_6

Cikara, M., Bruneau, E. G., & Saxe, R. R. (2011). Us and them: Intergroup failures of empathy. *Current Directions in Psychological Science, 20,* 149–153. http://dx.doi.org/10.1177/0963721411408713

Cikara, M., Bruneau, E., Van Bavel, J. J., & Saxe, R. (2014). Their pain gives us pleasure: How intergroup dynamics shape empathic failures and counter-empathic responses. *Journal of Experimental Social Psychology, 55,* 110–125. http://dx.doi.org/10.1016/j.jesp.2014.06.007

Cikara, M., & Fiske, S. T. (2013). Their pain, our pleasure: Stereotype content and schadenfreude. *Annals of the New York Academy of Sciences, 1299,* 52–59. http://dx.doi.org/10.1111/nyas.12179

Dannals, J. E., & Miller, D. T. (2017). Social norm perception in groups with outliers. *Journal of Experimental Psychology: General, 146,* 1342–1359. http://dx.doi.org/10.1037/xge0000336

Doosje, B., Ellemers, N., & Spears, R. (1995). Perceived intragroup variability as a function of group status and identification. *Journal of Experimental Social Psychology, 31,* 410–436. http://dx.doi.org/10.1006/jesp.1995.1018

Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble perception of dynamic emotional groups. *Psychological Science, 28,* 193–203. http://dx.doi.org/10.1177/0956797616678188

Ellis, H. D., Deregowski, J. B., & Shepherd, J. W. (1975). Description of white and black faces by white and black subjects. *International Journal of Psychology, 10,* 119–123. http://dx.doi.org/10.1080/00207597508247325

Goldenberg, A., Garcia, D., Halperin, E., Zaki, J., Kong, D., Golarai, G., & Gross, J. J. (2019). Beyond emotional similarity: The role of situation-specific motives. *Journal of Experimental Psychology: General*. Advance online publication. http://dx.doi.org/10.1037/xge0000625

Goldenberg, A., Garcia, D., Zaki, J., Kong, D., Golarai, G., Halperin, E., & Gross, J. J. (2019). Beyond emotional similarity: The role of situation specific motives. *Journal of Experimental Psychology: General*. Advance online publication. http://dx.doi.org/10.1037/xge0000625

Goldenberg, A., Saguy, T., & Halperin, E. (2014). How group-based emotions are shaped by collective emotions: Evidence for emotional transfer and emotional burden. *Journal of Personality and Social Psychology, 107,* 581–596. http://dx.doi.org/10.1037/a0037462

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

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General, 144,* 432–446. http://dx.doi.org/10.1037/xge0000053

Haberman, J., Lee, P., & Whitney, D. (2015). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *Journal of Vision, 15*(4), 16. http://dx.doi.org/10.1167/15.4.16

Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 718–734. http://dx.doi.org/10.1037/a0013899

Halperin, E. (2014). Emotion, emotion regulation, and conflict resolution. *Emotion Review, 6,* 68–76. http://dx.doi.org/10.1177/1754073913491844

Halperin, E. (2016). *Emotions in conflict: Inhibitors and facilitators of peace making*. New York, NY: Routledge.

Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. New York, NY: Cambridge University Press.

Hess, U., Thibault, P., Adams, R. B., Jr., & Kleck, R. E. (2010). The influence of gender, social roles, and facial appearance on perceived emotionality. *European Journal of Social Psychology, 40,* 1310–1317. http://dx.doi.org/10.1002/ejsp.715

Hogg, M. A. (2000). Subjective uncertainty reduction through self-categorization: A motivational theory of social identity processes. *European Review of Social Psychology, 11,* 223–255. http://dx.doi.org/10.1080/14792772043000040

Hogg, M. A. (2004). Uncertainty and extremism: Identification with high entitativity groups under conditions of uncertainty. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 401–418). New York, NY: Psychology Press.

Hogg, M. A. (2007). Uncertainty–identity theory. *Advances in Experimental Social Psychology, 39,* 69–126. http://dx.doi.org/10.1016/S0065-2601(06)39002-8

Hogg, M. A., Sherman, D. K., Dierselhuis, J., Maitner, A. T., Moffitt, G., & Hogg, M. (2007). Uncertainty, entitativity and group identification. *Journal of Experimental Social Psychology, 43,* 135–142. http://dx.doi.org/10.1016/j.jesp.2005.12.008

Hohman, Z. P., Gaffney, A. M., & Hogg, M. A. (2017). Who am I if I am not like my group? Self-uncertainty and feeling peripheral in a group. *Journal of Experimental Social Psychology, 72,* 125–132. http://dx.doi.org/10.1016/j.jesp.2017.05.002

Hornsey, M. J. (2008). Social identity theory and self-categorization theory: A historical review. *Social and Personality Psychology Compass, 2,* 204–222. http://dx.doi.org/10.1111/j.1751-9004.2007.00066.x

Hubert-Wallander, B., & Boynton, G. M. (2015). Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of Vision, 15*(4), 5. http://dx.doi.org/10.1167/15.4.5

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2018). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science, 22,* 1–35.

Ji, L., Rossi, V., & Pourtois, G. (2018). Mean emotion from multiple facial expressions can be extracted with limited attention: Evidence from visual ERPs. *Neuropsychologia, 111,* 92–102. http://dx.doi.org/10.1016/j.neuropsychologia.2018.01.022

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

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18,* 1084–1089. http://dx.doi.org/10.1111/j.1467-9280.2007.02029.x

Lamer, S. A., Sweeny, T. D., Dyer, M. L., & Weisbuch, M. (2018). Rapid visual perception of interracial crowds: Racial category learning from emotional segregation. *Journal of Experimental Psychology: General, 147,* 683–701. http://dx.doi.org/10.1037/xge0000443

Lau, T., Morewedge, C. K., & Cikara, M. (2016). Overcorrection for social-categorization information moderates impact bias in affective forecasting. *Psychological Science, 27,* 1340–1351. http://dx.doi.org/10.1177/0956797616660292

Lazerus, T., Ingbretsen, Z. A., Stolier, R. M., Freeman, J. B., & Cikara, M. (2016). Positivity bias in judging ingroup members' emotional expressions. *Emotion, 16,* 1117–1125. http://dx.doi.org/10.1037/emo0000227

Levy, A., Saguy, T., van Zomeren, M., & Halperin, E. (2017). Ingroups, outgroups, and the gateway groups between: The potential of dual identities to improve intergroup relations. *Journal of Experimental Social Psychology, 70,* 260–271. http://dx.doi.org/10.1016/j.jesp.2016.09.011

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology, 78,* 223–246. http://dx.doi.org/10.1037/0022-3514.78.2.223

Lin, L. C., Qu, Y., & Telzer, E. H. (2018). Intergroup social influence on emotion processing in the brain. *Proceedings of the National Academy of Sciences of the United States of America, 115,* 10630–10635. http://dx.doi.org/10.1073/pnas.1802111115

Lüdecke, D. (2015). *Sjstats: Collection of convenient functions for common statistical computations*. Retrieved from https://cran.r-project.org/web/packages/sjstats/index.html

Martin, D., & Macrae, C. N. (2007). A face with a cue: Exploring the inevitability of person categorization. *European Journal of Social Psychology, 37,* 806–816. http://dx.doi.org/10.1002/ejsp.445

McGregor, I., Zanna, M. P., Holmes, J. G., & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology, 80,* 472–488. http://dx.doi.org/10.1037/0022-3514.80.3.472

Michael, E., de Gardelle, V., & Summerfield, C. (2014). Priming by the variability of visual information. *Proceedings of the National Academy of Sciences of the United States of America, 111,* 7873–7878. http://dx.doi.org/10.1073/pnas.1308674111

Milgram, S., Bickman, L., & Berkowitz, L. (1969). Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology, 13,* 79–82. http://dx.doi.org/10.1037/h0028070

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology, 22,* 103–122. http://dx.doi.org/10.1002/ejsp.2420220202

Mummendey, A., Otten, S., Berger, U., & Kessler, T. (2000). Positive-negative asymmetry in social discrimination: Valence of evaluation and salience of categorization. *Personality and Social Psychology Bulletin, 26,* 1258–1270. http://dx.doi.org/10.1177/0146167200262007

Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society, Interface*. Advance online publication. http://dx.doi.org/10.1098/rsif.2017.0213

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4,* 133–142. http://dx.doi.org/10.1111/j.2041-210x.2012.00261.x

Neta, M., & Whalen, P. J. (2010). The primacy of negative interpretations when resolving the valence of ambiguous facial expressions. *Psychological Science, 21,* 901–907. http://dx.doi.org/10.1177/0956797610373934

Oakes, P. J. (1987). The saliance of social categorization. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Rieche, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory* (pp. 117–141). Oxford, UK: Blackwell Publishers Inc.

Oakes, P. J., Turner, J. C., & Haslam, S. A. (1991). Perceiving people as group members: The role of fit in the salience of social categorizations. *British Journal of Social Psychology, 30,* 125–144. http://dx.doi.org/10.1111/j.2044-8309.1991.tb00930.x

Parkinson, B. (2011). Interpersonal emotion transfer: Contagion and social appraisal. *Social and Personality Psychology Compass, 5,* 428–439. http://dx.doi.org/10.1111/j.1751-9004.2011.00365.x

Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018). Perceiving groups: The people perception of diversity and hierarchy. *Journal of Personality and Social Psychology, 114,* 766–785. http://dx.doi.org/10.1037/pspi0000120

Rhodes, G., Neumann, M., Ewing, L., Bank, S., Read, A., Engfors, L. M., . . . Palermo, R. (2018). Ensemble coding of faces occurs in children and develops dissociably from coding of individual faces. *Developmental Science, 21*(2), e12540. http://dx.doi.org/10.1111/desc.12540

Rule, N. O., & Sutherland, S. L. (2017). Social categorization from faces: Evidence from obvious and ambiguous groups. *Current Directions in Psychological Science, 26,* 231–236. http://dx.doi.org/10.1177/0963721417697970

Simon, B., & Pettigrew, T. F. (1990). Social identity and perceived group homogeneity. *European Journal of Social Psychology, 20* (July 1989), 269–286.

Sweeny, T. D., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 329–337. http://dx.doi.org/10.1037/a0028712

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks Cole.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168,* 242–249. http://dx.doi.org/10.1016/j.psychres.2008.05.006

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA: Basil Blackwell.

Turner, J. C., & Oakes, P. J. (1986). The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British Journal of Social Psychology, 25,* 237–252. http://dx.doi.org/10.1111/j.2044-8309.1986.tb00732.x

Turner, J. C., & Reynolds, K. J. (2011). Self-categorization theory. In P. A. M. Van Lange, A. W. Kruglanski, & T. E. Higgins (Eds.), *Handbook of theories in social psychology* (pp. 399–417). London, UK: SAGE Publications.

Van Kleef, G. A. (2009). How emotions regulate social life the emotions as social information (EASI) model. *Current Directions in Psychological Science, 18,* 184–188. http://dx.doi.org/10.1111/j.1467-8721.2009.01633.x

van Zomeren, M., Leach, C. W., & Spears, R. (2012). Protesters as "passionate economists": A dynamic dual pathway model of approach coping with collective disadvantage. *Personality and Social Psychology Review, 16,* 180–199. http://dx.doi.org/10.1177/1088868311430835

Walton, G. M., & Brady, S. T. (2017). The many questions of belonging. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation* (pp. 272–293). New York, NY: Guilford Press.

Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology, 92,* 82–96. http://dx.doi.org/10.1037/0022-3514.92.1.82

Weisbuch, M., & Ambady, N. (2008). Affective divergence: Automatic responses to others' emotions depend on group membership. *Journal of Personality and Social Psychology, 95,* 1063–1079. http://dx.doi.org/10.1037/a0011993

Whitney, D., Haberman, J., & Sweeny, T. (2014). From textures to crowds: Multiple levels of summary statistical perception. In J. Werner & L. M. Chalupa (Eds.), *The new visual neuroscience* (pp. 695–709). Boston, MA: MIT Press.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology, 69,* 105–129. http://dx.doi.org/10.1146/annurev-psych-010416-044232

Woolhouse, M. H., & Lai, R. (2014). Traces across the body: Influence of music-dance synchrony on the observation of dance. *Frontiers in Human Neuroscience, 8,* 965. http://dx.doi.org/10.3389/fnhum.2014.00965

Received September 21, 2018
Revision received May 6, 2019
Accepted June 13, 2019 ■