Journal of Personality and Social Psychology

Perceiving Groups: The People Perception of Diversity and Hierarchy

L. Taylor Phillips, Michael L. Slepian, and Brent L. Hughes

Online First Publication, January 15, 2018. http://dx.doi.org/10.1037/pspi0000120

CITATION

Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018, January 15). Perceiving Groups: The People Perception of Diversity and Hierarchy. *Journal of Personality and Social Psychology*. Advance online publication. http://dx.doi.org/10.1037/pspi0000120

Journal of Personality and Social Psychology

0022-3514/18/$12.00 http://dx.doi.org/10.1037/pspi0000120

Perceiving Groups: The People Perception of Diversity and Hierarchy

L. Taylor Phillips
New York University

Michael L. Slepian
Columbia University

Brent L. Hughes
University of California, Riverside

The visual perception of individuals has received considerable attention (visual *person* perception), but little social psychological work has examined the processes underlying the visual perception of groups of people (visual *people* perception). Ensemble-coding is a visual mechanism that automatically extracts summary statistics (e.g., average size) of lower-level sets of stimuli (e.g., geometric figures), and also extends to the visual perception of groups of faces. Here, we consider whether ensemble-coding supports *people perception,* allowing individuals to form rapid, accurate impressions about groups of people. Across nine studies, we demonstrate that people visually extract high-level properties (e.g., diversity, hierarchy) that are unique to social groups, as opposed to individual persons. Observers rapidly and accurately perceived group diversity and hierarchy, or variance across race, gender, and dominance (Studies 1–3). Further, results persist when observers are given very short display times, backward pattern masks, color- and contrast-controlled stimuli, and absolute versus relative response options (Studies 4a–7b), suggesting robust effects supported specifically by ensemble-coding mechanisms. Together, we show that humans can rapidly and accurately perceive not only individual persons, but also emergent social information unique to groups of people. These people perception findings demonstrate the importance of visual processes for enabling people to perceive social groups and behave effectively in group-based social interactions.

Keywords: person perception, groups, diversity, hierarchy, social vision

Supplemental materials: http://dx.doi.org/10.1037/pspi0000120.supp

Humans frequently perceive, judge, and interact with groups of people. Teachers perceive and interact with their classes, managers their employees, and athletes their teammates, as well as opponents. Groups of people are important for social and organizational life; we work in squads, depend on teams, and must frequently evaluate other groups of people to inform our social decisions and behavior (Hackman & Katz, 2010; Neuberg, Kenrick, & Schaller, 2010). Which group we should join, which we should hire, and which we may be likely to beat in competition are among the many examples of decisions and behaviors that rely on impressions of groups as wholes. And, people often form visual impressions before any other (Balcetis & Lassiter, 2010; Ito & Urland, 2003; McArthur & Baron, 1983). Whereas visual perception of individual persons has received considerable attention (Adams, Ambady, Nakayama, & Shimojo, 2011; Balcetis & Lassiter, 2010; Hugenberg, Young, Bernstein, & Sacco, 2010; Macrae & Quadflieg, 2010), visual perception of groups of people has rarely been investigated. Do people quickly and accurately form visual impressions of groups, as they do for individuals?

Recent advances in vision sciences identify mechanisms by which observers are able to perceive sets of complex stimuli, including human faces, as coherent wholes (see Alvarez, 2011; Phillips, Weisbuch, & Ambady, 2014; Whitney, Haberman, & Sweeny, 2014 for reviews). For example, ensemble-coding mechanisms allow observers to quickly and accurately extract summary statistics, such as averages across a group of stimuli (Ariely, 2001; Haberman & Whitney, 2009) — an important requirement for *people perception*, as compared with person perception.

Ensemble-coding mechanisms also provide fertile ground for exploring a second requirement for people perception: whether observers can accurately detect characteristics that emerge uniquely within group contexts. Many such properties are based on higher moments that have not been examined by perception research. For instance, while “happiness” can characterize both an

---

L. Taylor Phillips, Stern School of Business, New York University; Michael L. Slepian, Columbia Business School, Columbia University; Brent L. Hughes, Department of Psychology, University of California, Riverside.

In loving memory of Nalini Ambady, whose creative thinking, vitalizing mentorship, and constant encouragement made this work possible. This material is based upon work supported by National Science Foundation Graduate Research Fellowship grant DGE-114747. We are grateful to Jason Haberman, Schinria Islam, Erik Madsen, Sean Malahy, Michelle Peretz, Timothy Sweeny, Alex Todorov, Chris Young, and the members of Stanford SPA, IPC, and DPER Labs for helpful comments, materials, and research assistance.

Correspondence concerning this article should be addressed to L. Taylor Phillips, Stern School of Business, New York University, 44 West 4th Street, New York, NY 10012. E-mail: ltaylorphillips@stern.nyu.edu

individual (single unit) and a group (sum, count, or average unit), "emotional convergence" can characterize *only* groups (variance among the units; Barsade & Knight, 2015). Finally, these visual computations would need to shape adaptive social interaction and behavior. Can people extract such variance characteristics and, if so, do people use this information to inform their social interactions and behavior?

Here, we integrate research from vision science, cognitive science, social psychology, and organizational behavior to provide an examination of people perception. We examine whether observers quickly and accurately perceive uniquely group-level characteristics of social groups. We then explore the robustness of these effects, using a variety of stringent tests to identify whether ensemble-coding supports group perceptions, and to compare ensemble-coding methods with other visual thin-slice methods. Finally, we consider implications for both intragroup dynamics and intergroup interactions.

Social Vision: Person Perception and Group Cognition

Many theories across social psychology and organizational behavior hinge on observers' ability to form quick and accurate visual impressions of groups (Phillips et al., 2014). Theories of status, for example, suggest the importance of accurately detecting group hierarchy, which can endear one to leaders and protect one from punishment for acting out of place (Anderson & Brown, 2010). Theories of diversity suggest that people's satisfaction with team interactions depends in part on their beliefs about how diverse the team is (Homan, Greer, Jehn, & Koning, 2010). Theories of leadership emphasize that successful leaders should be attuned to their group's emotional variance, and adjust their motivational tactics accordingly (DeRue, 2011; Sanchez-Burks & Huy, 2009). Despite the important theoretical role of group-based impressions, whether and how people can visually perceive such group-level characteristics that are foundational to social life (e.g., diversity, hierarchy) remains relatively unexplored (for impressions of collective affect, see Sanchez-Burks, Bartel, Rees, & Huy, 2015; for nonvisual impressions of diversity, see Shemla, Meyer, Greer, & Jehn, 2014).

Person Perception

Decades of research have helped build extensive models of person perception (e.g., Abelson, Dasgupta, Park, & Banaji, 1998; Hamilton, Sherman, Way, & Percy, 2015; Lickel et al., 2000). More recently, a growing literature on social vision has focused on integrating social, cognitive, and vision sciences to model person perception and social interaction (Adams et al., 2011; Balcetis & Lassiter, 2010). These models highlight the primacy of visual information when forming impressions of individuals. For example, perceivers detect the gender of a target face in under 200 ms (Ito & Urland, 2003).

Moreover, models of person perception highlight the interactive nature of visual perception and social–cognitive processes in forming impressions of individuals (Adams et al., 2011). For instance, people dynamically process visual information (e.g., facial features) and social information (e.g., category labels) as they categorize gender (Freeman & Ambady, 2011) and race (Pauker et al., 2009). People's own social characteristics, such as dominance, can affect how they process visual information, such as target facial dominance (Watkins, Jones, & DeBruine, 2010). And when judging a central target individual embedded in a group of nontargets, observers' cultural background shapes whether nontarget expressions influence judgments of the target's expression (Masuda et al., 2008; see also Walker & Vul, 2014). Taken together, extant research highlights the importance of understanding not only how observers think about individual social targets, but also how observers visually perceive targets to form impressions of them. However, this work has not been extended to visual perceptions of groups themselves; can people perceive group characteristics that cannot emerge in single individuals?

Group Cognition

Decades of research on group cognition examines how people's social attitudes and beliefs (e.g., stereotypes) and judged characteristics of groups (e.g., entitativity) combine to affect judgments about individual group members or entire groups themselves (Hamilton, 2005, 2007; Hamilton et al., 2015; Fiske & Taylor, 2013). For example, stereotypes about gender roles restrict impressions of women's suitability for positions in science or in leadership roles (Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012; Rudman & Glick, 2001). Further, the more people believe race is based on biology, the more they avoid contact with outgroups (Williams & Eberhardt, 2008). Past work has also shown that the more individuals think that an outgroup is heterogeneous, as compared with homogenous, the less discriminatory they are toward outgroup members (Brauer & Er-rafiy, 2011). And, when people believe a group is especially dominant, they tend to believe that group to also be more competent, and then act accordingly (Cuddy, Fiske, & Glick, 2008).

The literature on group cognition thus demonstrates that individuals' judgments of broad social categories (e.g., Canadians, women) and their properties (e.g., diversity, hierarchy) have important effects on interactions with members of those categories (Linville, 1998; Park & Judd, 1990). However, this literature has not explored the role that rapid *visual* processes play in perceptions of groups of people as wholes. While there are instances in which broad social category groups are "perceived" on paper (e.g., judging "binders full of women" based on a stack of resumes), observers also frequently form quick, visual impressions of specific groups of people (e.g., a project team in the office). In turn, such visual impressions may combine with prior knowledge and stereotypes to inform judgments, as well as their subsequent behaviors toward the group or group members.

People Perception

Whereas extant work on person perception and group cognition provides an important foundation for building models of social vision, these models remain incomplete. Much of our social lives involve perceiving and interacting with groups, rather than individuals alone—watching a basketball game, assembling a task force, working with a team. One important next step in advancing models of social vision—models that consider the interaction of vision and social psychology—is to account for how visual perception of *groups* is achieved. Here, we argue that a model of visual people perception will complement and extend existing models of social vision, making them more truly "social" by accounting for impressions of and interactions with groups, rather than only individuals. The current work adopts paradigms from vision sciences to examine this social vision of groups, or *people perception*.

Recent theorizing generates novel predictions about people perception (Phillips et al., 2014; see also Alvarez, 2011; Whitney et al., 2014). First, people are able to form quick and accurate visual impressions of groups. Second, these group-based impressions should shape downstream social decisions and behavior. The selection, extraction, application (SEA) model (Phillips et al., 2014) provides one theoretical, but currently untested, account of people perception. The SEA model suggests that the visual system first *selects* individuals for inclusion in the group or not, and then *extracts* social information about the group. Finally, this visually extracted information feeds forward and is *applied* to social decision-making and behavior.

Here, we test the specific hypothesis that observers can detect uniquely social properties from groups of people, such as diversity and hierarchy. Importantly, perceiving diversity and hierarchy in groups of people hinges on the ability to extract *variance* information along relevant social dimensions. For instance, visual perceptions of diversity depend upon the extraction of variance information about race or gender. Similarly, visual perceptions of hierarchy depend upon the extraction of variance information about dominance or status. The ability to detect variance along these social dimensions is a critical test of observers' people perception abilities. Visual impressions of groups should not rely simply on summations of person perceptions, but must also enable perceivers to detect uniquely group-level properties—those that emerge only in group contexts, such as properties based on variance. This information should in turn influence social behavior. We explore people perception by testing whether observers visually extract information (variance) unique to groups as opposed to individuals, and whether they then apply this information to their social decision-making.

Ensemble-Coding

Ensemble-coding from the vision sciences provides a lens through which to examine and test people perception processes (Alvarez, 2011; Ariely, 2001, 2008; Chong & Treisman, 2003; Haberman & Whitney, 2012). Previous ensemble-coding research demonstrates that observers detect average size from object sets, such as circles (Ariely, 2001). More recently, researchers have shown that ensemble-coding mechanisms extract group average information for more complex sets, like groups of faces. For instance, observers can detect overall motion direction of multiple walking point-light displays (Sweeny, Haroz, & Whitney, 2013), and morphological averages (gender, emotional expression) of groups of faces composed of the same individual identity (Haberman & Whitney, 2009). Further, observers are able to detect average facial morphology (deFockert & Wolfenstein, 2009), emotional expression (Won & Jiang, 2013), and attractiveness (Walker & Vul, 2014), even for groups of faces composed of different identities.

Although people perception processes that extract group averages may be useful for informing subsequent social decisions and behavior, their importance could be limited without also providing variance information about groups (e.g., Chan, 1998). That is, knowing the average dominance or gender of a group provides an incomplete picture of the dynamic of the group. For example, two groups that are equally dominant on average may in fact represent one high variance and one low variance group. Such variance information could suggest one group is a well-oiled competitive machine, while the other is fraught with internal conflict and ripe for the conquering (e.g., Halevy, Chou, Galinsky, & Murninghan, 2012). Work on social identity, psychological safety, and belonging suggests that surmising gender or racial means as well as variance (diversity) may help observers estimate whether they are welcome or belong in certain contexts, such as in computer science classrooms (e.g., Murphy, Steele, & Gross, 2007). For perceivers to engage in people perception, detecting both average and variance of group characteristics would offer more complete information, and thus be useful for quickly and accurately appraising a group.

Further, the extraction of averages in group contexts represent arguably similar properties that can be found in individual contexts: Individuals may have an emotional expression, and groups may have an average emotion. However, there are new properties that emerge only in groups that are not possible in a single individual, including diversity and hierarchy. For instance, a group can have a steep hierarchy (one person at the top, many at the bottom) or a flat hierarchy (more equally distributed power), but a single person cannot have either; multiple people are needed to form a hierarchy. A people perception, rather than person perception, approach suggests that these uniquely group-emergent properties should be readily perceivable.

Whereas existing work demonstrates that observers can perceive group facial averages along certain dimensions, there are as of yet no empirical demonstrations that observers can perceive even higher-order moments, such as variance, from groups of people (Haberman & Whitney, 2012; Phillips et al., 2014). Recent findings suggest that ensemble-coding could support such perceptual abilities. For instance, ensemble-coding mechanisms discount or ignore stimulus outliers when extracting averages (Haberman & Whitney, 2010; see also Dannals & Miller, 2017), implying that outliers are noted as outside the central variance of the set. Further, across several ensemble-coding data sets testing stimuli from simple lines to moving figures, group variability influences accuracy for perceptions of the mean, suggesting it somehow impinges on the perceptual process (Dakin, Mareschal, & Bex, 2005; Haberman, Brady, & Alvarez, 2015; Sweeny et al., 2013). Does ensemble-coding support the perception of variance along social dimensions in groups of complex stimuli, such as human faces?

The Current Research: People Perception

The processes underlying people perception, much like those involved in person perception, must solve a whole host of challenges. Perceptions of groups inherently involve greater complexity and diversity than perceptions of individuals. To perceive uniquely group-level characteristics—such as those that depend upon variance (e.g., diversity, hierarchy)—observers must move beyond summation or averaging processes (see also Ariely, 2008; Myczek & Simons, 2008). A people perception deployment of ensemble-coding would allow observers to quickly identify meaningful social dimensions within a group and compare group members along that dimension, even in the face of underlying heterogeneity of group members along other dimensions. Thus, the ability to perceive variance is a critical test for determining the uniqueness of people perception in relation to person perception.

First, can people quickly compare faces within a group in meaningful ways? Dystopian futures aside, groups entail variety. While ensemble-coding enables people to average across groups of objects or faces, averages are not alone in their importance. As discussed, observers must also extract variance information to understand the dynamics that emerge within the group. Here, we examine whether the visual system extracts variance within a group, thereby allowing observers to quickly glean group-level features fundamental to social life—diversity and hierarchy (Gruenfeld & Tiedens, 2010). We address this question by testing whether people can rapidly extract variance in race and gender (i.e., diversity: Studies 1 and 2) and dominance (i.e., hierarchy: Study 3) within a group of diverse facial identities.

Second, is ensemble-coding the underlying perceptual mechanism? To the extent that ensemble-coding helps observers accomplish perceptions of variance across complex and diverse social stimuli, it should support perception of critical group properties, including diversity and hierarchy. Therefore, we additionally test the robustness of our effects, by using (a) extremely short temporal intervals and backward pattern masks that make serial processing impossible, (b) color- and contrast-controlled stimuli that isolate morphology from other cues, and (c) absolute versus relative judgment measures to ensure accuracy is not contingent on the paradigm used (Studies 4a–7b). In sum, we provide the first exploration of whether observers are able to perceive uniquely group-level properties, and whether ensemble-coding supports this perceptual process. Such information promises to deepen models of social vision by incorporating perceptions of emergent, group properties that cannot be gleaned by person perception alone.

Methodological Approach

Stimuli

Face group stimuli were randomly generated from sets of 70 faces, representing seven standardized levels of race for 10 unique male facial identities (Studies 1, 4a, 5, 6, 7a; Singular Inversions, 2006), seven standardized levels of gender for 10 unique White facial identities (Studies 2, 4a, 5, 6; Singular Inversions, 2006), or seven standardized levels of facial dominance for 10 unique White male facial identities (Studies 3, 4b, 5, 7b; Oosterhof & Todorov, 2008; see also Johnson, Freeman, & Pauker, 2012). The computer model that generates these faces is based on a database of three-dimensional laser scanned human faces. Laser scans of faces (coded by sex and race) were then subjected to multiple regression analyses that identified the line in multidimensional face-space (i.e., constellations of facial features) that best predicts the sex, race, and other consensually perceived properties of the scanned faces (Oosterhof & Todorov, 2008; Singular Inversions, 2006). From this data-driven model, we generate faces that represent seven standardized levels of the critical dimension manipulated in each study. For instance, in the case of race, faces continuously represent extremely dark through extremely light morphology. In the case of gender, faces continuously represent extremely feminine through extremely masculine morphology. In the case of dominance, faces continuously represent extremely submissive through extremely dominant morphology (Oosterhof & Todorov, 2008; Figure 1).

Task Design

Across studies, we adapt an ensemble-coding protocol to test participants' people perception abilities (Ariely, 2001; Haberman & Whitney, 2009). To do so, we presented participants with multiple trials of unique face displays. Each trial briefly presented a grid of four faces, and subsequently presented a second set of four faces. Participants then judged whether the second set of faces was more or less diverse (Studies 1, 2, 4, 5, 6) or hierarchical (Studies 3, 4, 5) than the first set (Study 7 takes a different approach, described later). On each trial, one face was randomly drawn from each of eight (out of 10) randomly selected pools of faces, creating a trial-set of eight randomly selected faces. From this trial-set, the faces were divided into two groups of four faces, each arranged in a tight grid such that faces were tightly clustered as a single group (see Figure 2). One grid (target group) was presented first, followed by the second grid of faces (comparison group), which was displayed until participants made a response indicating "more" or "less" ("diverse" Studies 1, 2, 4a, 5, 6; or "hierarchical" Studies 3, 4b, 5).

For each set of two groups displayed, variance was never equivalent, and varied across trials. To assess accuracy, we calculated the objective morphological difference in variance between the two comparison groups, and tested whether participants' responses predicted this true objective variance. Accuracy rates above chance indicate participants, to some degree, successfully visually extracted group variance (Ariely, 2001; Haberman & Whitney, 2009).

Vision science paradigms that capitalize on stable similarities across human visual systems typically rely on high numbers of trials per participant and treat trial, not participant, as the unit of analysis (e.g., Ariely, 2001; Palmer, 1999). In a pilot study using our face stimuli, we followed this approach and based our initial sample size on similar paradigms (e.g., $n = 660$ trials; Ariely, 2001). Although in this pilot study, we found a large effect size (Cohen's h = .92), we then performed sample size analyses using a far more conservative estimate of a small effect size Cohen's h = .10, $\alpha = .05$, and power = .8, which suggested $n \geq 784$ trials was needed (two-sided pwr.p.test in the R software package). Therefore, in each study we collect more than 784 trials to ensure that our sample sizes yield high-powered designs.

Analytical Approach

Across studies, we analyzed participant accuracy using multiple methods. First, we use a multilevel generalized linear mixed model analysis (glmer for R Statistical Package), which allows for the inclusion of both trial- and participant-level controls (to explore robustness), as well as delineation of random versus fixed effects. We treat participant as a random effect (by-participant random-intercepts model), effectively controlling for differences across participants that may affect responses (e.g., between-participants variance; reported in Tables). This allows us to analyze each trial while still accounting for participant variance, and allows us to

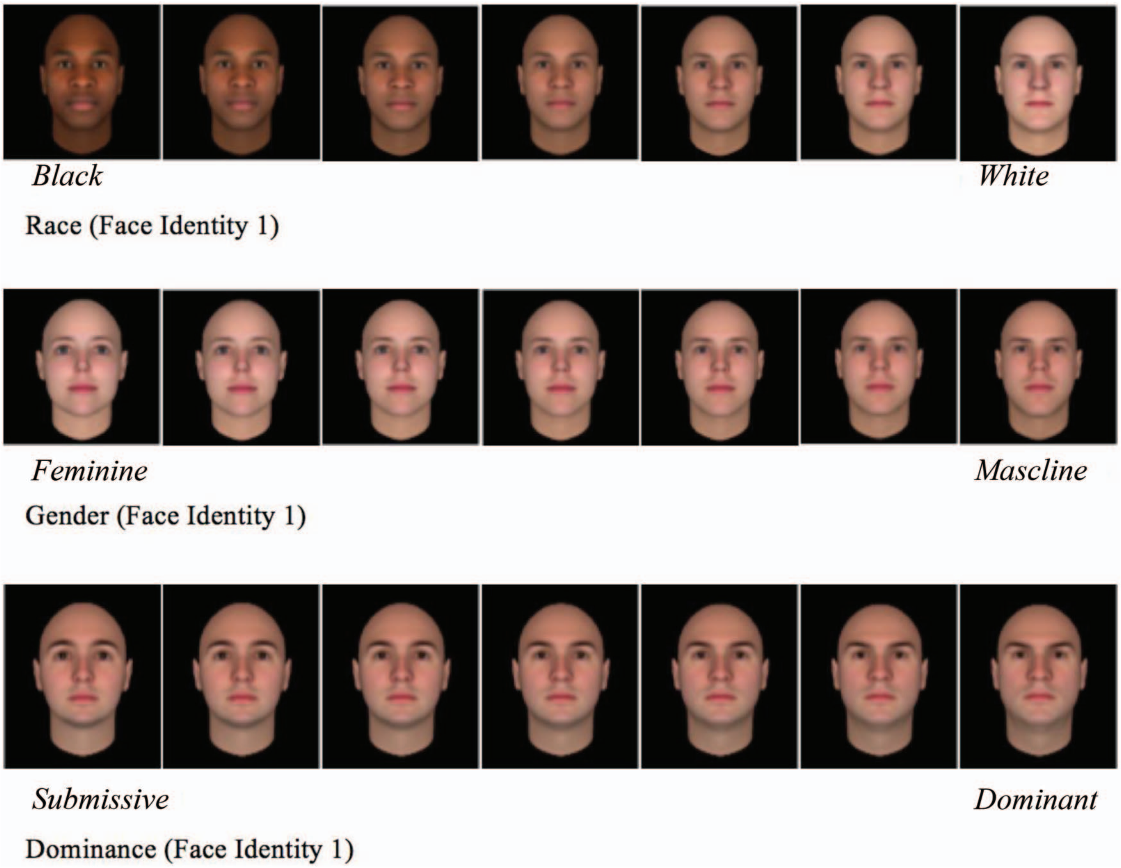

Figure 1. Face stimuli examples-seven levels per facial identity. Racial diversity, Study 1. Gender diversity, Study 2. Hierarchy, Study 3. See the online article for the color version of this figure.

generalize to unsampled participants.[1] We additionally include ICC's for each study, comparing between-participants variance to the overall variance; overall, we find that this variance generally approaches zero (see tables).

Further, the intercept of this regression compares the log-odds-ratio of response correct to zero (chance), and can thus be used to determine whether accuracy significantly differed from chance. This intercept can also be transformed into a probability, demonstrating a readily interpretable estimated likelihood of correct responses.

Second, we also use additional analytic methods, including (a) proportion testing, with trial as the level of analysis and no controls; and (b) binomial logistic regression, with trial as the level of analysis and trial-level controls. Results persist with these analyses, without controls or random effects (see online supplemental materials).

Part 1: Visual Extraction of Variance (Studies 1–3)

The first three studies examine whether perceivers can visually, rapidly, and accurately, perceive variance in a visual display of people (i.e., diversity in facial morphology) across race (Study 1), gender (Study 2), and dominance (Study 3). We then move to robustness studies, probing the role of the ensemble-coding mechanism even more explicitly, in Part 2.

Study 1: Racial Diversity

We first tested whether observers extract racial diversity information from groups. Race is considered one of the critical "master status" characteristics, in that race is perceived automatically from individual faces (Ito & Urland, 2003; Stangor, Lynch, Duan, & Glas, 1992). Such automaticity may be due, in part, to the evolutionary significance of distinguishing between in-group and out-group members (Kurzban & Leary, 2001). A range of indicators likely played a role in helping people make quick group membership determinations, including targets' size, language/vocal cues, physical adornment, facial morphology, and skin-tone; social constructions of race have historically been based, in part, upon the latter of these (Blair, Judd, & Fallman, 2004; Hunter, 2007; Smedley & Smedley, 2005; Omi & Winant, 2014). Today, perceptions

---

[1] Disagreement persists regarding the appropriate calculation of degrees of freedom for generalized linear mixed model analyses, and so we provide information on number of participants, number of trials, and residual degrees of freedom in lieu of traditional degrees of freedom by variable. For linear mixed models, we provide degrees of freedom by variable (rounded), as estimated using lmeTest for R Statistical Package. Further, for each analysis, we also provide more traditional, simpler analyses in the online supplemental material, which include traditional degrees of freedom.

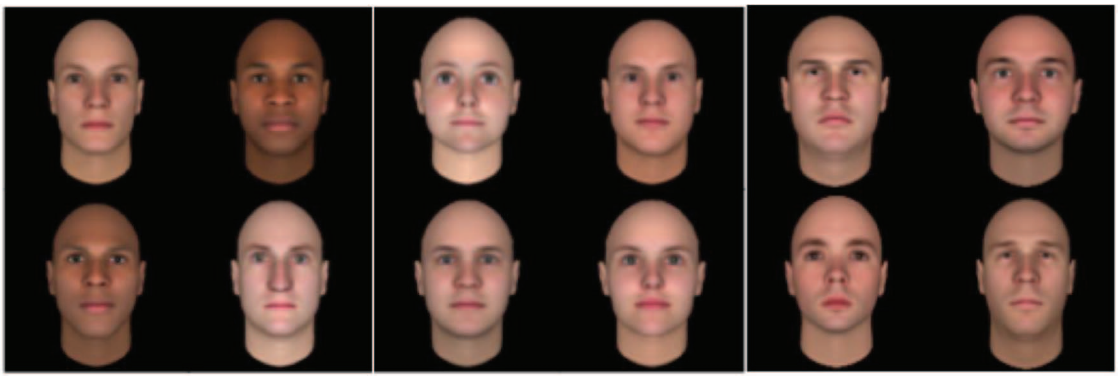

Figure 2. Face display examples. Racial diversity, Study 1. Gender diversity, Study 2. Hierarchy, Study 3. See the online article for the color version of this figure.

of racial diversity might help observers determine how welcome they may feel in contexts from classrooms to workplaces (e.g., Purdie-Vaughns, Steele, Davies, Ditlmann, & Crosby, 2008).

A person's race is a complex social identity, determined by sociocultural, historical, and personal factors. However, in today's society, a person's visually perceived race is in part influenced by a combination of face morphology and skin tone, which exists on a continuum, whether or not those perceptions accurately reflect targets' self-determined racial identity (e.g., Freeman, Pauker, Apfelbaum, & Ambady, 2010; Pauker et al., 2009). Therefore, we rely on both morphological and skin tone differences as racial diversity cues. Based on previous ensemble-coding work, we can infer that observers can extract average cues to diversity, including average facial morphology from groups of faces (de Fockert & Wolfenstein, 2009; Won & Jiang, 2013) and average color from groups of simpler stimuli like circles (Brady & Alvarez, 2011). Are observers able to quickly and accurately glean racial diversity—that is, variance in perceived race—from groups of diverse faces?

Method

Thirty student volunteers (20 female; eight Asian, three Black, two Latino/a, nine "Other," eight White) participated in an in-lab, within-subjects study in exchange for course credit, completing 100 trials each ($N = 3000$ trials). Participants were instructed to think of "diverse" as "how different are the people in the group from one another." We were interested in people's basic ability to visually extract variance rather than probing activation of motivated social concepts of what diversity means and represents (e.g., Unzueta, Knowles, & Ho, 2012). Thus, we further instructed participants to focus on racial diversity, reminding them "For example, if a group has all Black people, that group is not diverse. If a group has all White people, that group is not diverse. If a group has a mixture of people, that group is diverse." The task design was also explained, emphasizing that target and comparison groups display quickly, and that participants should rely on their gut feelings or first impressions to make their comparison judgments.

For each trial, participants viewed a set of four faces displayed for 2 s (target group). After a 100 ms blank screen, participants were then shown a second set of four new faces (comparison group), and asked if this group was more or less diverse than the previous group. Racially diverse face stimuli were used (for more information on stimuli, see Methodological Approach above).

Measures

Group diversity. Group diversity was measured by taking the standard deviation of the standardized race levels of the four faces in each group, thus mathematically representing the objective visual diversity in raced facial characteristics for the overall stimulus (i.e., the grid of faces). Each face had a score that ranged from -3 (Black) to $+3$ (White), with 0 as the average morphological face and a score of $+/-1$ moving one *SD* in White/Black morphology. We calculated this score for both the target group (*target* group diversity) and the comparison group (*comparison* group diversity).

Group diversity distance. Group diversity distance was measured by taking the absolute value of the difference between the group diversity for each pair of display grids, and serves as a measure of task difficulty. The larger the difference between target and comparison stimuli, the more notable the difference will likely be, with smaller distance likely yielding more difficulty in discriminating whether the second grid of faces was more or less diverse.

Group mean. Group mean was measured by averaging the standardized race levels of the four faces in each group, yielding *target* group mean and *comparison* group mean. Group morphology means may affect accuracy for extractions of group morphology variance, due to the fact that "diversity" (and in later studies, "hierarchy") may connote different things to different people, despite our clarifying instructions. For instance, many Americans think of diversity as non-White, rather than as "high variance" (Plaut, Garnett, Buffardi, & Sanchez-Burks, 2011). Therefore, we analyze data with and without means as control variables, to ensure that effects are based on stimuli variance, above and beyond stimuli means.

Response time. Response time indicated how long participants took to respond to each trial. If participants who took longer performed better on the task, this may suggest the underlying mechanism is more deliberative, rather than automatic.

Trial number. Trial number was identified for each response. Participants may have performed better on later trials, after some initial practice period, especially if the underlying mechanism is more deliberative than automatic.

Response correct. Response correct indicated whether participants' "less diverse" or "more diverse" selections were accurate (1 = correct, 0 = incorrect). If participants have more correct responses than would be expected by guessing (chance), it suggests participants are successfully extracting variance information from the visual displays.

Results

As hypothesized, we found that participants accurately perceived group racial diversity. We specified a binomial regression of response correct on random-intercept of participant. We found that participants performed significantly better than chance: intercept log-odds = .58, SE = .05, z = 10.85, $p < .001$; probability response correct 64.14% (see Table 1).

Results persisted when we additionally included random-intercept of trial number, as well as fixed effects of response time, group diversity distance, target group mean, and comparison group mean (each centered; see Table 2): intercept log-odds = .61, SE = .05, z = 11.61, $p < .001$; probability response correct 64.75%. Of the predictors in the regression, only group diversity distance had a significant effect: log-odds = .79, SE = .08, z = 9.71, $p < .001$. This suggests that task difficulty influences observers' accuracy, with greater differences between grids being more accurately perceived. In particular, participants were more accurate when the diversity between of comparison and target groups was more dissimilar, and less accurate when diversity of comparison and target groups was more similar. Overall, the results suggest that participants' accurately visually perceived group racial diversity from groups of different faces.

Study 2: Gender Diversity

Study 1 demonstrates that observers are able to accurately detect group racial diversity. Is this skill specific to racial diversity, or are other forms of diversity also detected? We next tested whether observers extract gender diversity information from groups. Like race, gender is considered a "master status" characteristic, in that race is perceived automatically from individual faces (Ito & Urland, 2003; Stangor et al., 1992). This may be due, in part, to the social and evolutionary significance of gleaning another person's sex (Macrae, Alnwick, Milne, & Schloerscheidt, 2002). Importantly, objective gender identity is self-determined and can include orthogonal feminine and masculine elements (Auster & Ohm, 2000). However, visually perceived gender identity often varies along a single continuum, from feminine to masculine, and is influenced by gendered facial morphology (Freeman, Rule, Adams, & Ambady, 2010). When we use the term gender in the current work, we mean it as existing along a continuum (both with respect to identity, but more critically, with respect to visual facial morphology).

As with racial diversity, previous ensemble-coding work demonstrates that observers can extract average cues to gender diversity from groups of faces, including average gender from a group of faces in which each has the same underlying facial identity (morphed continuously from male to female; Haberman & Whitney, 2009). However, observers may also need to extract variance in gender—or gender diversity within a group—from groups of different faces. While average gender is an important characteristic of groups, gender variance provides additional information that may prove crucial, or at least informative, in certain contexts (e.g., mate selection, college major decision, Finnish sauna entry). For example, women's impressions of gender diversity in STEM classrooms can help them decide whether they feel welcome (e.g., Murphy et al., 2007). Thus, we hypothesize that individuals can rapidly and accurately perceive gender diversity—or variance in the degree to which faces represent masculine and feminine morphology—in groups of different faces.

Method

Thirty student volunteers (19 female; eight Asian, three Black, two Latino/a, nine "Other," eight White) participated in an in-lab, within-subjects study in exchange for course credit, completing 100 trials each (N = 3,000 trials). Participants were first instructed to think of "diverse" as "how different are the people in the group from one another." As in Study 1, we further explained the meaning of diversity, using gender instead of race examples. Procedure then followed than of Study 1, but using gender diverse faces. Measures were the same as in Study 1, except standardized levels of gender were used instead of race.

Results

As hypothesized, we found that participants accurately perceived group gender diversity. We specified a binomial regression of response correct on random-intercept of participant. We found that participants performed significantly better than chance: intercept log-odds = .43, SE = .04, z = 11.42, $p < .001$; probability response correct 60.50% (see Table 1).

Table 1

Accuracy: Summary of Mixed-Model Binomial Logistic Regression Analyses (Studies 1–3)

	Racial diversity (Study 1) N = 3,000; residual df = 2,998				Gender diversity (Study 2) N = 3,000; residual df = 2,998				Hierarchy (Study 3) N = 3,000; residual df = 2,998			
Variable (fixed)	b	SE	z	P(correct)	b	SE	z	P(correct)	b	SE	z	P(correct)
Intercept (log-odds)	.58	.05	10.85**	.64	.43	.04	11.42**	.60	.17	.04	4.00**	.54
Variable (random)	SD^2		ICC		SD^2		ICC		SD^2		ICC	
Participant	.04		.02		<.001		.007		.02		.01	

** $p < .001$.

Table 2
Accuracy: Summary of Mixed-Model Binomial Logistic Regression Analyses With Control Variables (Studies 1–3)

	Racial diversity (Study 1)				Gender diversity (Study 2)				Hierarchy (Study 3)			
	$N = 3{,}000$; residual $df = 2{,}993$				$N = 3{,}000$; residual $df = 2{,}993$				$N = 3{,}000$; residual $df = 2{,}993$			
Variable (fixed)	*b*	*SE*	*z*	*p*	*b*	*SE*	*z*	*p*	*b*	*SE*	*z*	*p*
Intercept (log-odds)	.61	.05	11.61	<.001	.44	.04	11.60	<.001	.17	.04	3.95	<.002
Response time	−.005	.01	−.54	.59	−.01	.01	−1.61	.11	−.001	.01	−.14	.89
Group *SD* distance	.79	.08	9.71	<.001	.55	.07	7.30	<.001	.32	.07	4.47	<.001
Target group mean	.04	.04	1.00	.32	.03	.04	.68	.50	−.04	.04	−1.21	.23
Comparison group mean	.07	.04	1.68	.09	.19	.04	4.89	<.001	−.06	.04	−1.74	.08

Variable (random)	SD^2	*ICC*	SD^2	*ICC*	SD^2	*ICC*
Participant	.04	.02	<.001	.007	.02	.01
Trial number	<.001	.03	<.001	.02	<.001	.03

Note. Response time, group SD distance (S1 and S2 diversity; S3 hierarchy), target group mean (S1 Whiteness, S2 masculinity, S3 dominance), and comparison group mean (S1 Whiteness, S2 masculinity, S3 dominance) each centered at their means.

Results persisted when we additionally included random-intercept of trial number, as well as fixed effects of response time, group diversity distance, target group mean, and comparison group mean (each centered; see Table 2): intercept log-odds $= .44$, $SE = .04$, $z = 11.60$, $p < .001$; probability response correct 60.85%. Of the predictors in the regression, group diversity distance had a significant effect: log-odds $= .55$, $SE = .07$, $z = 7.30$, $p < .001$. This again suggests that task difficulty influences observers' accuracy.

Comparison group mean also had a significant effect, such that more masculine comparison groups were associated with increased observer accuracy for diversity perceptions. This suggests that observers' accuracy was also swayed by the average gender of the comparison group. This effect may be due to participants' layering of subjective understandings of "diversity" onto the task (e.g., women might be perceived as more "diverse" given underrepresentation in various fields; see General Discussion). Importantly, participants accurately perceived gender variance, above and beyond any influence of mean gender. Altogether, our results suggest that just as participants accurately extracted and perceived racial diversity in Study 1, they successfully did so with gender in Study 2.

Study 3: Hierarchy

While classic and recent theorizing suggests that diversity (or a lack thereof) is among the most important defining characteristics of human groups, hierarchy is equal in its importance (Gruenfeld & Tiedens, 2010). Just as variance is a critical ingredient for the emergence of group diversity, it is again critical for the emergence of group hierarchy. As such, hierarchy is a social dimension that emerges only within group contexts and cannot be perceived within a single individual. Hierarchy can be defined as the variance in dominance or power within a group, and is paramount to social functioning. For instance, knowing one's place in a group's hierarchy and acting accordingly has been important for maintaining social ties, collecting resources, and ultimately surviving across cultures throughout history (Anderson & Brown, 2010; Boehm, 1999; Gruenfeld & Tiedens, 2010). Failing to detect hierarchy can cause maladaptive social behavior that may ultimately result in ostracism, expulsion, and even death (Anderson, Ames, & Gosling, 2008; Neuberg et al., 2010).

Variance in social dominance within a group serves as a cue for hierarchy, with increasing variance suggesting steeper group hierarchy (Magee & Galinsky, 2008; Sidanius & Pratto, 1999). To the extent ensemble-coding mechanisms support perception of average facial morphology (Won & Jiang, 2013), as existing work suggests, then we might infer that ensemble-coding supports perceptions of average facial dominance across a group of diverse individuals (see also online supplemental materials).[2] However, perception of hierarchy may depend on successfully perceiving variance in dominance cues across individuals within a group.

Extending prior work, we examine the rapid extraction of an interpersonal trait dimension (dominance) in order to support perceptions of group hierarchy. Previous work shows that people more fluently process hierarchical (vs. egalitarian) team diagrams (Zitek & Tiedens, 2012), and accurately perceive dyads' hierarchical relationships by gleaning each individual's power from nonverbal behavioral streams (Hall & Friedman, 1999; Tiedens & Fragale, 2003). Further, person perception work with single faces finds that perceivers often assume interpersonal dominance from facial dominance alone (Hehman, Flake, & Freeman, 2015; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). Relying on this past work, we use facial dominance as a subtle, nonverbal cue, and expect that people will interpret this as indicative of interpersonal dominance and thus hierarchical position.

---

[2] In order to successfully perceive group hierarchy, observers need to detect the cues or bases of hierarchy, as they do for racial and gender diversity. Thus, we ran an additional study testing whether perceivers extract average facial dominance (known to correspond to behavioral dominance in many situations; Carré & McCormick, 2008; Hehman et al., 2015), before we then moved on to extractions of variance. Previous research demonstrates that observers quickly, and with consensus, perceive dominance from individual faces (Oosterhof & Todorov, 2008; Rule et al., 2012). We found that participants accurately perceived the average dominance of facial grids significantly greater than chance (binomial regression response correct on random-intercept effect of participant: intercept log-odds $= 1.14$, $SE = .13$, $z = 8.78$, $p < .001$; probability response correct 75.69%; see online supplemental materials).

However, in order to test whether people can rapidly visually perceive hierarchy in a group of faces, rather than who is simply more powerful in a dyad, Study 3 explored whether individuals accurately extract variance in facial dominance within a group of different facial identities. Can the visual system extract variance in dominance to support perceptions of the hierarchy of a group of diverse individuals?

Method

Thirty student volunteers (14 female; seven Asian, five Black, two Latino/a, one Native, four "Other," 11 White) participated in an in-lab, within-subjects study in exchange for course credit, completing 100 trials each ($N = 3{,}000$ trials). Participants were instructed to think of "hierarchical" as "how much hierarchy do you think was in the group" or "how unequal was the group." Procedure then followed than of Study 1, but using dominance diverse faces. Measures were the same as in Study 1, except standardized levels of dominance were used instead of race, and diversity was replaced with hierarchy.

Results

As hypothesized, participants accurately perceived group hierarchy. We specified a binomial regression of response correct on random-intercept of participant. We found that participants performed significantly better than chance: intercept log-odds = .17, $SE = .04$, $z = 4.00$, $p < .001$; probability response correct 54.32% (see Table 1).

Results persisted when we additionally included random-intercept of trial number, as well as fixed effects of response time, group hierarchy distance, target group mean, and comparison group mean (each centered; see Table 2): log-odds = .17, $SE = .04$, $z = 3.95$, $p < .001$; probability response correct 54.36%. Of the predictors in the regression, only group hierarchy distance had a significant effect: log-odds = .32, $SE = .07$, $z = 4.47$, $p < .001$. This again suggests that task difficulty influences observers' accuracy. Overall, we find that people extract information regarding group hierarchy, or variance in dominance, just as they extract diversity, or variance in race and gender.[3]

Discussion and Pooled Analysis: Visual Extraction of Variance (Studies 1–3)

Together, these results suggest that observers are able to accurately perceive variance in individuals' facial dominance, facial morphology, and skin-tone from groups of faces. And, people are able to accurately glean this variance information from groups of faces displayed a mere two seconds. Further, in a pooled analysis ($N = 9{,}000$ trials, 90 participants), we find that observers are able to accurately perceive group variance across different dimensions, intercept log-odds = .40, $SE = .03$, $z = 15.57$, $p < .001$; probability response correct 59.90%. While we do find variation in accuracy across studies (see Table 3), overall our results suggest observers successfully extract group-level summary statistics, allowing them to generate impressions of groups along interpersonal trait dimensions (such as dominance), as well as definitional features (such as facial morphology or skin-tone). Specifically, we find that people accurately extract variance in race, gender, and dominance from groups of faces, and use such information to accurately differentiate group diversity and hierarchy.

Table 3
Accuracy: Summary of Mixed-Model Binomial Logistic Regression Analysis (Pooled Studies 1–3)

	Pooled (Studies 1–3) $N = 9{,}000$; residual $df = 8{,}991$			
Variable (fixed)	*b*	*SE*	*z*	*p*
Intercept (log-odds)	.40	.03	15.56	<.001
Response time	−.03	.02	−1.41	.16
Group SD distance	.28	.02	12.37	<.001
Target group mean	.005	.02	.21	.84
Comparison group mean	.06	.02	2.75	.006
Study (linear contrast)	.08	.03	2.82	.005
Study (quadratic contrast)	−.11	.02	−7.33	<.001
Variable (random)	SD^2		*ICC*	
Participant	.006		<.001	
Trial Number	<.001		—	

Note. Response time, group *SD* distance (S1 and S2 diversity; S3 hierarchy), target group mean (S1 Whiteness, S2 masculinity, S3 dominance), and comparison group mean (S1 Whiteness, S2 masculinity, S3 dominance) each scaled and centered at their means within study. Study contrasts: linear (S1 Racial Diversity = 1, S2 Gender Diversity = −1, S3 Hierarchy = 0); quadratic (S1 Racial Diversity = −1, S2 Gender Diversity = −1, S3 Hierarchy = 2).

Part 2: Robustness Checks and Underlying Mechanisms (Studies 4–7)

Overall, the results of Studies 1–3 suggest that observers are able to accurately perceive the variance in facial morphology from a group of faces, supporting visual perceptions of racial and gender diversity as well as hierarchy. Yet it is important to distinguish whether participants are truly extracting these variances using parallel (ensemble-coding) processing mechanisms or instead are using serial (attention-driven) processes (for discussion, see Ariely, 2008; Myczek & Simons, 2008). That is, simply looking at each face in a serial fashion would be more akin to sequential person perception. However, we suggest perceivers are actually engaging in *people perception*: judging the group as a whole by processing faces in parallel. We thus confirm a role for ensemble-

---

[3] While we find medium-to-large effect sizes for racial diversity, gender diversity, and dominance mean perceptions, we find a smaller effect size for hierarchy perceptions. Hierarchy may be complicated by its multiply determined nature (e.g., status, rank, power) and ranging definitions. We use variance in facial dominance as a cue to group hierarchy, which is less direct and more conservative than using variance in racial and gender morphology as cues to group demographic diversity. As such, we ran an in-lab replication study with a slightly different design, testing 99 participants who each completed 50 trials (see online supplemental materials). We again found a significant yet small effect (binomial regression response correct on random-intercept of participant: intercept log-odds = .13, $SE = .03$, $z = 4.12$, $p < .001$; probability response correct 53.32%), suggesting participants are able to perceive group hierarchy based on variance in facial dominance (replicating across Studies 3, 4b, and Supplemental Study 2). The presence of even stronger cues to hierarchy—akin to facial masculinity/femininity as a cue to diversity—may increase effect sizes and should be investigated in future work (see General Discussion).

coding using a variety of study designs, procedures, and stimuli in Part 2.

Studies 4a–b: Parallel Processing and Short Presentation Times

Ensemble-coding is the rapid, automatic generation of summary statistics of a group or scene, and is achieved via parallel processing (Haberman & Whitney, 2012; Whitney et al., 2014). We tested whether ensemble-coding is indeed the mechanism underlying our earlier effects by running additional studies using target group display times so short (200 ms) that serial processing would be impossible (Ariely, 2001, 2008; Haberman & Whitney, 2009). If we find accuracy under these conditions, it cannot be due to serial processing, and therefore should be an outcome of parallel processing via ensemble-coding.

Method

Study 4a: Racial and gender diversity. Fifteen student volunteers (nine female; four Asian, two Black, four Latino/a, one Native, three "Other," one White) participated in an in-lab, within-subjects study in exchange for course credit, completing 200 trials each ($N_{RACE} = 3{,}000$ trials; $N_{GENDER} = 3000$ trials). Protocol and measures followed that of Studies 1 (racial diversity) and 2 (gender diversity), except that target groups were displayed for only 200 ms, and blank screens between target and comparison were 1 s. Half the participants completed the race study first, and half the participants completed the gender study first; after a short break, participants then read new instructions and began the second study. The use of the same participants across different visual tasks is common in vision science paradigms, again capitalizing on similarity across individual visual systems (Ariely, 2001; Palmer, 1999). For racial diversity, completing the race study first improved accuracy: log-odds = .12, $SE = .04$, $z = 3.17$, $p = .002$); for gender diversity, there was no effect of counterbalance condition on accuracy (log-odds = .04, $SE = .04$, $z = 1.11$, $p = .27$). Results persist with counterbalance as a fixed effect (see online supplemental materials).

Study 4b: Hierarchy. Fifteen student volunteers (eight female; four Asian, one Latino/a, one Native, five "Other," four White) participated in an in-lab, within-subjects study in exchange for course credit, completing 200 trials each ($N_{HIERARCHY} = 3{,}000$ trials). Protocol and measures followed that of Study 3 (hierarchy), except that target groups were displayed for only 200 ms, and blank screens between target and comparison were 1 s. Half the participants completed the hierarchy study first, and half the participants completed a dominance (mean) study first (see online supplemental materials). After a short break, participants then read new instructions and began the second study. There was no effect of counterbalance condition on accuracy (log-odds = .002, $SE = .04$, $z = .05$, $p = .96$), and results persist with counterbalance as a fixed effect (see online supplemental materials).

Results

Study 4a: Racial and gender diversity. We specified a binomial regression of response correct on random-intercept of participant. We found that participants performed significantly better than chance for both racial diversity (intercept log-odds = .46, $SE = .05$, $z = 9.10$, $p < .001$; probability response correct 61.28%; see Table 4) and gender diversity (intercept log-odds = .21, $SE = .04$, $z = 5.94$, $p < .001$; probability response correct 55.43%). Given presentation times were too short to allow for serial processing, this suggests parallel processing via ensemble-coding is in fact at play.

Study 4b: Hierarchy. We specified a binomial regression of response correct on random-intercept of participant. We found that participants performed significantly better than chance for hierarchy (intercept log-odds = .15, $SE = .05$, $z = 3.00$, $p = .003$; probability response correct 53.78%; see Table 4), again providing evidence for an ensemble-coding mechanism.

Overall, we find that even under conditions that should prohibit serial processing, observers' impressions of the variance of groups of faces predicts, better than chance, the objective variance of those groups. Thus, these impressions must be generated via parallel processing (ensemble-coding).

Study 5: Parallel Processing and Backward Masking

Study 5 further tests whether ensemble-coding underlies perceivers' ability to extract group variance information by pairing very short display times (200 ms) with backward pattern masking. It is worth noting that this display time (as in Studies 4a–b) prevents multiple fixations and so precludes serial processing, given that saccade initiation and completion requires approximately 200 ms (Thorpe, Fize, & Marlot, 1996). However, previous work demonstrates that observers are able to generate impressions of complex scenes from single fixations, without making saccades, via ensemble-coding—a parallel processing mechanism. Thus,

Table 4
Accuracy: Summary of Mixed-Model Binomial Logistic Regression Analysis With 200 ms Displays (Studies 4a and 4b)

	Racial diversity (Study 4a) $N = 3{,}000$; residual *df* = 2,998				Gender diversity (Study 4a) $N = 3{,}000$; residual *df* = 2,998				Hierarchy (Study 4b) $N = 3{,}000$; residual *df* = 2,998			
Variable (fixed)	*b*	*SE*	*z*	*P*(correct)	*b*	*SE*	*z*	*P*(correct)	*b*	*SE*	*z*	*P*(correct)
Intercept (log-odds)	.46	.05	9.10^{**}	.61	.22	.04	5.94^{**}	.55	.15	.05	3.00^{*}	.54
Variable (random)	SD^2		*ICC*		SD^2		*ICC*		SD^2		*ICC*	
Participant	.02		.009		<.001		.005		.02		.01	

* $p < .01$. ** $p < .001$.

while single faces from the group are not encoded at such speeds, and thus cannot be serially processed, properties of the entire group can be encoded via ensemble-coding (Cohen, Dennett, & Kanwisher, 2016).

Nevertheless, perhaps some portion of the scene or task itself is residually processed or processed in visual short-term memory (VSTM), after the display is removed. By including a backward pattern mask, such residual processing is prevented, interrupted by the mask. Thus, if we find accuracy under these conditions, it cannot be due to serial processing—even postdisplay in VSTM—and therefore any accuracy can be interpreted as due to ensemble-coding.

Method

Eighteen student volunteers (five female; five Asian, one Black, one Latino/a, four "Other", four White) participated in an in-lab, within-subjects study in exchange for $15, completing 200 trials each per stimulus set ($N_{RACE} = 3{,}000$ trials; $N_{GENDER} = 3{,}200$ trials; $N_{HIERARCHY} = 3{,}600$ trials).[4] Protocol and measures followed Studies 4a–b. However, each target group was displayed for 200ms, followed immediately by a backward pattern mask displayed for 200 ms (stimulus onset asynchrony 200 ms), which was then followed immediately by the comparison group, displayed until the participant entered a response. For the mask, we used a scrambled image: a random sample face grid cut into 10 × 10 pixel squares and randomly rearranged. Participants completed the race, gender, and hierarchy studies in random order. Participants completed the first study, then after a short break, read new instructions and began the second study, and again for the third study.

Results

For each study, we specified a binomial regression of response correct on random-intercept of participant. We found that participants performed significantly better than chance for both racial diversity (intercept log-odds = .53, $SE = .06$, $z = 9.54$, $p < .001$; probability response correct 62.91%; see Table 5) and gender diversity (intercept log-odds = .17, $SE = .04$, $z = 4.91$, $p < .001$; probability response correct 54.34%; see Table 5). However, we found that participants' performance was not significantly different from chance for hierarchy (intercept log-odds = .05, $SE = .04$, $z = 1.21$, $p = .23$; probability response correct 51.17%; see Table 5).

Under conditions that prohibit serial processing, and further prohibit residual processing, observers continue to form impressions of group diversity that are accurate, compared to chance. But performance was no better than chance in the case of hierarchy. In light of the results of Study 4b, the results here suggest the possibility that additional processing in VSTM helps support impression formation of group hierarchy. Our test of hierarchy is more conservative than for diversity: our cues to hierarchy are more minimal, and the term itself is loaded. Specifically, we manipulate hierarchy using only facial dominance (i.e., variations in morphology of a small set of cues, such as brow ridge and shape of jaw), while race and gender are manipulated using many facial features, including morphology and pigmentation. Thus, while parallel processing is still likely at play (given accuracy in Study 4b with very short presentation times), participants might recruit additional serial processing strategies to aid accurate judgments (see Footnote 3 and General Discussion).

Together, Studies 4a and 4b provide evidence for an ensemble-coding mechanism, and Study 5 suggests provides further evidence of this mechanism. At least in the case of diversity, even when preventing additional processing in VSTM, participants can form accurate group impressions from parallel processing (ensemble-coding).

Study 6: Impoverished Stimuli

In real-life groups of diverse individuals, faces vary across a variety of dimensions, including both morphology and skin tone. Our results suggest perceivers are able to judge diversity in such complex contexts. However, given the pigmentation variation built into our computer modeled stimuli, our gender and racial diversity findings may be driven by participants' ensemble-coding based on low-level color contrast alone (e.g., Brady & Alvarez, 2011; Haberman et al., 2015). While this may support perceptions of diversity, we were interested to see if observers are able to still perceive diversity when relying only on facial morphology cues. Thus, we ran additional studies using gray-scaled, brightness standardized versions of the same faces.

Method

Sixteen student volunteers (six female; three Asian, one Black, two Latino/a, six "Other," four White) participated in an in-lab, within-subjects study in exchange for course credit, completing 100 trials each ($N_{RACE} = 1{,}500$ trials; $N_{GENDER} = 1{,}600$ trials).[5] Protocol and measures followed that of Studies 1 (racial diversity) and 2 (gender diversity), except that gray-scaled, brightness standardized versions of the face stimuli were used. Half the participants completed the race study first, and half the participants completed the gender study first; after a short break, participants then read new instructions and began the second study. There was no effect of counterbalance condition on accuracy (racial diversity: log-odds = −.09, $SE = .05$, $z = -1.69$, $p = .09$; gender diversity: log-odds = −.01, $SE = .05$, $z = -.10$, $p = .92$). Results persist with counterbalance as a fixed effect (see online supplemental materials).

Results

We specified a binomial regression of response correct on random-intercept of participant. We found that participants performed significantly better than chance for both racial diversity (intercept log-odds = .36, $SE = .06$, $z = 6.24$, $p < .001$; proba-

---

[4] Three of the 15 originally recruited participants were unable to complete all three stimuli set studies due to computer malfunction. Thus, we recruited three additional participants who completed every study. In total, 15 completed the racial diversity study, 16 completed the gender diversity study, and 18 completed the hierarchy study.

[5] One of the 15 originally recruited participants was unable to complete the racial diversity study due to computer malfunction. Thus, we recruited one additional participant who completed both studies. In total, 15 completed the racial diversity study, while 16 completed the gender diversity study.

Table 5

Accuracy: Summary of Mixed-Model Binomial Logistic Regression Analysis With 200 ms Displays and Backward Pattern Mask (Study 5)

	Racial diversity (Study 5) $N = 3{,}000$; residual $df = 2{,}998$				Gender dversity (Study 5) $N = 3{,}200$; residual $df = 3{,}198$				Hierarchy (Study 5) $N = 3{,}600$; residual $df = 3{,}598$			
Variable (fixed)	*b*	*SE*	*z*	*P*(correct)	*b*	*SE*	*z*	*P*(correct)	*b*	*SE*	*z*	*P*(correct)
Intercept (log-odds)	.53	.06	9.54**	.63	.17	.04	4.91**	.54	.05	.04	1.21	.51
Variable (random)	SD^2		*ICC*		SD^2		*ICC*		SD^2		*ICC*	
Participant	.02		.01		<.001		.005		.007		.007	

** $p < .001$.

bility response correct 58.89%; see Table 6) and gender diversity (intercept log-odds = .12, $SE = .06$, $z = 2.17$, $p = .03$; probability response correct 53.01%). This suggests observers' ability persists in the face of impoverished stimuli with fewer cues to diversity: even when hue and brightness cues are held constant, observers are able to extract group diversity information (variance in race and gender).

Studies 7a–b: Perceiver Ratings

Ensemble-coding paradigms typically force observers to choose whether a stimulus is more or less representative of a critical dimension. As such, these paradigms capitalize on relative comparisons as well as mathematical properties of the stimuli to create an objective measure of perceiver accuracy, rather than relying on more subjective rating measures that are typical in social psychology paradigms (see Ariely, 2001). However, we additionally tested whether group racial diversity (Study 7a) or group hierarchy (Study 7b) affected observers' impressions of each group by using a correlational rating design, more typical to thin-slice paradigms. This methodology allows us to test whether participants are extracting visual information about group-level properties that then inform the accuracy of their judgments of the group, without relying on a comparative process. That is, instead of assuming that observers extract variance information about two groups to compare them, here we measure more directly the impressions individuals form of each target group.

Method

Study 7a: Racial diversity. One hundred twelve student volunteers (69 female; 33 Asian, six Black, seven Latino/a, 24 "Other," 37 White) participated in an in-lab, within-subjects study as part of a paid mass-testing session, completing 50 trials each ($N = 5{,}600$). The protocol followed that of Study 1 (racial diversity). However, for each trial, participants viewed only a single target group (grid of four faces) for 2 s, then rated the group on perceived diversity (0 = *not at all* to 7 = *very much*; composite of "diverse" and "different from each other;" $r = .67$) and perceived average race (0 = *not at all* to 7 = *very much*; composite of "White" and reversed "Black;" $r = .55$).

Study 7b: Hierarchy. One hundred forty adult volunteers from Mechanical Turk (64 female; 10 Asian, eight Black, eight Latino/a, four "Other," 110 White) participated in an online, within-subjects study as part of a paid mass-testing session, completing 50 trials each ($N = 7{,}000$). The protocol followed that of Study 3 (hierarchy). However, for each trial, participants viewed a single target group for 2 s, then rated the group on perceived hierarchy (0 = *not at all* to 7 = *very much*; composite of "hierarchical" and "unequal;" $r = .57$) and perceived dominance (0 = *not at all* to 7 = *very much*; composite of "dominant" and "strong;" $r = .77$).

Results

Study 7a: Racial diversity. First, we examined our focal outcome, perceived racial diversity. Unlike in previous studies, we found participant variance (random-intercept model) to be generally above zero; therefore we include both random-intercept and random-slopes effects in our model here. We specified a linear regression of perceived diversity on random intercept of participant, random slopes (by participant) of target diversity and target mean, and fixed effects of target diversity and target mean (each centered). We found that both objective target diversity (standard

Table 6

Accuracy: Summary of Mixed-Model Binomial Logistic Regression Analysis With Gray-Scale Stimuli (Study 6)

	Racial diversity (Study 6) $N = 1{,}500$; residual $df = 1{,}498$				Gender diversity (Study 6) $N = 1{,}600$; residual $df = 1{,}598$			
Variable (fixed)	*b*	*SE*	*z*	*P*(correct)	*b*	*SE*	*z*	*P*(correct)
Intercept (log-odds)	.36	.06	6.24**	.59	.12	.06	2.17*	.53
Variable (random)	SD^2		*ICC*		SD^2		*ICC*	
Participant	.01		.01		.01		.01	

* $p < .05$. ** $p < .001$.

deviation of stimulus; $b = .78$, 95% CI [.69, .87], $SE = .04$, $t(110) = 17.80$, $p < .001$) and objective target mean (mean of stimulus; $b = -.12$, 95% CI [−.17, −.07], $SE = .02$, $t(108) = -4.73$, $p < .001$) significantly and independently predicted perceived racial diversity (see Table 7). Importantly, objective target diversity more strongly predicted perceived racial diversity than did objective target mean race (i.e., nonoverlapping absolute value of confidence intervals).

We next conducted a supplementary analysis, predicting perceived average race. We found when entering objective target mean ($b = 1.21$, 95% CI [1.19, 1.24], $SE = .03$, $t(110) = 41.57$, $p < .001$) and objective target diversity ($b = -.01$, 95% CI [−.06, .02], $SE = .03$, $t(107) = -.54$, $p = .59$), only the former predicted perceived average race.

These results provide convergent evidence that perceivers are indeed able to quickly extract impressions of diversity from groups of faces, and that these perceptions of variance are separate from perceptions of mean. When judging diversity of a group, objective diversity played a bigger role than did objective mean race; when judging mean race of a group, objective mean race mattered instead.

Study 7b: Hierarchy. First, we examined our focal outcome, perceived hierarchy. Unlike in previous studies, we found participant variance (random-intercept model) to be generally above zero; therefore we include both random-intercept and random-slopes effects in our model here. We regressed perceived hierarchy on random intercept of participant, random slopes (by participant) of target hierarchy and target mean, and fixed effects of target hierarchy and target mean (each centered). We found that both objective target hierarchy (standard deviation of stimulus; $b = .32$, 95% CI [.26, .39], $SE = .03$, $t(138) = 9.74$, $p < .001$) and objective target mean (mean of stimulus; $b = .19$, 95% CI [.14, .25], $SE = .03$, $t(137) = 6.94$, $p < .001$) significantly and independently predicted perceived hierarchy (see Table 7). Importantly, objective target hierarchy more strongly predicted perceived hierarchy than did objective target mean (i.e., nonoverlapping absolute value of confidence intervals).

We next conducted a supplementary analysis, predicting perceived dominance. We found when entering objective target mean ($b = .41$, 95% CI [.36, .47], $SE = .03$, $t(137) = 15.37$, $p < .001$) and objective target hierarchy ($b = .09$, 95% CI [.04, .13], $SE = .02$, $t(138) = 3.77$, $p < .001$), the former was a stronger predictor of perceived average dominance.

These results provide convergent evidence that perceivers are indeed able to quickly extract impressions of hierarchy from groups of faces, and that these perceptions of variance are separate from perceptions of mean. When judging hierarchy of a group, objective hierarchy played a bigger role than did objective dominance; when judging dominance of a group, objective dominance mattered more instead.

Discussion: Robustness Checks and Underlying Mechanisms (Studies 4–7)

Overall, Studies 4–7 studies provide a more critical test of observers' ability to extract variance information from groups of faces, by providing only minimal display times (200 ms; Studies 4a–4b), backward pattern masks (Study 5), minimal stimuli cues (color- and contrast-controlled; Study 6), and minimal response indices (absolute judgments; Studies 7a–7b). Further, these robustness tests suggest that ensemble-coding does indeed underlie observers' perceptual ability to extract variance information. That is, our results cannot be attributed to processing faces in a serial order (Studies 4a–5b), and thus instead are a function of processing them in parallel (although we do find that residual processing may support impression formation of group hierarchy). Additionally, these results are not a function of mere color of the faces (Study 6), nor are they a function of the forced-choice nature of the comparative method (Studies 7a–7b). Finally, these studies replicate and extend our findings from Studies 1–3: ensemble-coding mechanisms enable observers' accurate perceptions of the variance in social features from a group of faces, supporting visual perceptions of racial diversity, gender diversity, and hierarchy.

General Discussion

Humans continually perceive and interact with groups of people: crowds, classrooms, teams, boards, audiences, and more.

Table 7

Subjectively Perceived Variance and Mean: Summary of Mixed-Model Linear Regression Analyses With Rating Design (Studies 7a and 7b)

	Perceived racial diversity (Study 7a) $N = 5{,}600$				Perceived average race (Study 7a) $N = 5{,}600$				Perceived hierarchy (Study 7b) $N = 6{,}994$				Perceived dominance (Study 7b) $N = 7{,}000$			
Variable (fixed)	*b*	*SE*	*t* (*df*)	*p*	*b*	*SE*	*t* (*df*)	*p*	*b*	*SE*	*t* (*df*)	*p*	*b*	*SE*	*t* (*df*)	*p*
Intercept	3.10	.07	42.89 (113)	<.001	3.92	.03	121.75 (112)	<.001	3.00	.08	35.51 (139)	<.001	3.47	.08	41.42 (139)	<.001
Target group *SD*	.78	.04	17.80 (110)	<.001	−.01	.03	−.54 (107)	.59	.32	.03	9.74 (138)	<.001	.09	.02	3.77 (138)	<.001
Target group mean	−.12	.02	−4.73 (108)	<.001	1.21	.03	41.57 (110)	<.001	.19	.03	6.94 (137)	<.001	.41	.03	15.37 (137)	<.001
Variable (random)	SD^2		*ICC*		SD^2		*ICC*		SD^2		*ICC*		SD^2		*ICC*	
P-intercept (total)	.55		.24		.10		.06		.98		.41		.96		.42	
P-slope group *SD*	.14		—		.03		—		.09		—		.02		—	
P-slope group mean	.04		—		.07		—		.08		—		.08		—	
Residual	1.53		—		.95		—		1.28		—		1.11		—	

Note. Target group SD (objective; S7a diversity, S7b hierarchy) and target group mean (objective; S7a Whiteness, S7b dominance) each centered at their means. Trials with missing ratings were dropped.

Accurate judgments of group social characteristics (e.g., diversity, hierarchy) facilitate successful interactions with such groups (e.g., Anderson et al., 2008; Shemla et al., 2014). Our results suggest that ensemble-coding supports fast and accurate perceptions of social information about groups of diverse people from even subtle facial cues. We find that observers visually perceive diversity (variance of race and gender: Studies 1 and 2) and hierarchy (variance of social dominance: Study 3) from minimal information, such as slight variation in face morphology displayed for a mere two seconds. Further, we find evidence that these group perceptions are robust and supported by ensemble-coding as the underlying visual mechanism (Studies 4a–7b).

To successfully perceive groups of people, observers must rapidly identify meaningful dimensions (e.g., gender, dominance) along which faces vary, compare faces within a group along that dimension to extract group mean and variance (e.g., diversity, hierarchy), and do so despite face diversity across other dimensions. Overall, these results suggest that ensemble-coding mechanisms facilitate rapid and accurate *people perception,* supporting visual perceptions of emergent, uniquely group-level properties, for complex, diverse, social stimuli.

Mechanisms of People Perception

Ensemble-coding. Humans evolved as a social species interacting frequently with groups of other people (Dunbar, 1993). Indeed, recent reviews have suggested human visual systems may be especially equipped to summarize properties unique to other people (Haberman & Whitney, 2012), and even groups of people, rather than mere low-level stimuli alone. Our work adds to this perspective, suggesting that ensemble-coding is useful for forming impressions of social groups. For instance, our results are the first to suggest that ensemble-coding enables people to accurately perceive interpersonal characteristics or *traits* (e.g., dominance), in addition to changeable features (e.g., emotion expression; Haberman & Whitney, 2009) or definitional features/states (e.g., face morphology, facial identity, de Fockert & Wolfenstein, 2009; see Todorov & Duchaine, 2008 for discussion of feature distinctions). In this way, the visual system supports perception of the bases of social diversity (facial morphology, skin-tone) as well as social hierarchy (facial dominance). Further, the visual system enables perception of these features for individuals, for group central tendency, and for group variance. In general, the evolution of people perception, and relatedly its applications in everyday life, deserve additional study. For instance, future work should test the role of top-down factors and/or serial-processing, such as motivational goals, social contexts, beliefs, and attention, that likely also impact perception of real-world groups (Phillips et al., 2014).

Ensemble-coding may also support social-comparative processes that provide observers with critical information as they evaluate, interact, and make group-relevant decisions. For instance, whereas prior work demonstrates that people can quickly and accurately perceive individual dominance (Rule, Adams, Ambady, & Freeman, 2012), we extend this to groups, including mean (supporting *inter*group perception; comparison across groups), and variance (supporting *intra*group perception; comparison within groups). Such perceptions may help observers function adaptively both within and across groups. Further, there are many uniquely group-level characteristics that affect group experiences and outcomes. Group cohesion and collective affect, for example, influence group performance, cooperation, and intergroup competition (Barsade, 2002; Evans & Dion, 1991). Ensemble-coding may support perception of these characteristics of groups as well (see Homan, Van Kleef, & Sanchez-Burks, 2016; Magee & Tiedens, 2006).

Variance extraction. Recent theory suggests that visual impressions of groups, or people perception, may occur across three critical phases: selection, extraction, and application (the SEA model; Phillips et al., 2014). Observers must first consider targets to be part of a group, visually extract information about the group as a whole, and finally apply this information to social decisions and behavior. Our results are consistent with theory-based predictions, providing some of the first evidence that observers extract uniquely group-level information (based on variance across social dimensions). A fruitful avenue for future research will be to continue to unpack the specific visual and social-cognitive mechanisms that support the extraction (and application) of variance information. We outline several possible mechanisms below.

First, as previously discussed, recent work suggests that ensemble-coding discounts or ignores outlier stimuli when extracting averages (Haberman & Whitney, 2010). This raises the possibility of boundaries to the accurate extraction of variance information from social groups. For example, when variance is extreme, a single group may no longer be inferred. Instead, observers may not consider outliers as part of the group, or observers may infer the presence of multiple groups. Both the current and previous ensemble-coding work generally present participants with clearly defined groups, precluding participants from making their own *selection* decisions as to who is part of the group and who is not. For instance, we use tightly clustered grids of faces and label each grid as a “group” in our instructions to participants. Thus, basic grouping principles of proximity and similarity may increase assumptions of display groups’ entitativity (for additional discussion, see Phillips et al., 2014). Future work should consider how outliers influence the extraction of group variance information, and how variance may influence selection of members into the group itself (as well as downstream applications, such as judgments of entitativity and cohesiveness; Brewer, Hong, & Li, 2004; Castano, Yzerbyt, & Bourguignon, 2003; Ip, Chiu, & Wan, 2006; La Macchia, Louis, Hornsey, & Leonardelli, 2016; Lickel et al., 2000; Magee & Tiedens, 2006).

Second, previous research suggests that ensemble-coding does not rely on endpoints or range to extract averages; in fact, observers do not explicitly attend to nor remember faces at the ends of a range in a distribution (Haberman & Whitney, 2010). Thus, it seems unlikely that people are relying on these anchors to create their impressions of variance. Rather, perceivers may be extracting some sort of average morphological face representation, and referencing back to it to extract the appropriate average and variance information. When they require variance information, people may consult how “fuzzy” this visual representation is in mind. One possibility for future research is to examine observers’ confidence in how they perceive the group as a whole. Groups with larger variance may be associated with a fuzzier mental representation than groups with less variance (cf. processing fluency: Reber, Schwarz, & Winkielman, 2004; see also Haberman et al., 2015; Haberman, Lee, & Whitney, 2015; Tong, Ji, Chen, & Fu, 2015).

Third, it is possible that when an observer's goal is to perceive variance, then she may focus more on endpoints. Future work should consider what happens when people do not know which judgment they will be making, and how much time can pass between seeing the group and making the judgment itself. Impressions may be updated as groups move and change dynamically across multiple channels and in real time (see also Haberman, Harp, & Whitney, 2009). Previous work suggests visual perceptions of individuals in dynamic, real-life situations—as opposed to static images on computer screens—can influence behavior (Barrett, Mesquita, & Gendron, 2011; Freeman & Ambady, 2011). How visual impressions of *groups* affect social decisions and behavior in dynamic, real-life situations deserves continued exploration (see Barsade, 2002; Harrison, Price, Gavin, & Florey, 2002; Sanchez-Burks & Huy, 2009).

Finally, given observers' accuracy at detecting not only averages but also variance, it may be that face-space dimensions exist for groups, as they do for individuals (Valentine, 1991). For instance, to successfully perceive groups of people, observers must not only detect characteristics in individuals, but also summarize these characteristics across diverse individuals to characterize the group as a whole. Such perception, especially variance perception, also requires comparing faces within a group to one another. Thus, observers must select dimensions along which faces vary—ostensibly, dimensions within face-space—and then summarize that dimension without being influenced by other dimensions. If such group face-space exists, then people may have an idea of normative group face-space, based on their interactions with groups in the world. In turn, this normative sense may influence people's expectations and decisions regarding groups.

People Perception Influences Social Decisions and Behavior

The interaction of multiple pieces of visual information (e.g., mean *and* variance) may be especially important as observers make complex decisions about groups. Here, we find that both variance and mean can, but do not always, affect perceivers' group impressions (e.g., Studies 7a and 7b). Different visual percepts (e.g., mean, variance) may be accessible and considered simultaneously (Haberman et al., 2015), and then interact to influence later social behavior (application stage in SEA Model; Phillips et al., 2014). For instance, visual cues of both diversity and average race may combine to help people generate feelings of whether they belong (e.g., in STEM classrooms). Or percepts of both diversity and hierarchy may combine, suggesting to perceivers in what role they belong (e.g., as student, faculty, or staff).

The specific dynamics of how and when percepts are applied to a range of social behaviors should be explored in future research. Just as person perception processes can bias perceivers (Macrae & Quadflieg, 2010; Olivola, Funk, & Todorov, 2014), people perception may also bias social judgments, making the study of visual group impression formation all the more important. In particular, observers' initial visual impressions may anchor their social judgments of a group, thus affecting subsequent behavior (cf., Rosenthal, 1994). For instance, many believe diverse groups experience more conflict than homogenous groups do (Lount, Sheldon, Rink, & Phillips, 2015; Mannix & Neale, 2005). If observers *perceive* a group as highly diverse, then they may expect conflict and interact as if conflict is likely. In turn, group conflict may become more likely. Similarly, visual impressions of variation in facial dominance may influence the development of hierarchy, which normally emerges rapidly in new groups and teams (Gruenfeld & Tiedens, 2010). In this case, formally designating group evaluation criteria and group member roles may help overcome any biasing effects of early visual impressions.

Accurate Visual Perceptions Versus Accurate Social Judgments

The current results demonstrate that perceivers' visual judgments of diversity and hierarchy, after mere 2 s (and even 200 ms) glances at a group, achieve accuracy that is consistently above chance. However, previous work has shown that people's reported judgments of hierarchy and diversity in teams are highly variable and often inaccurate, biased by both motivation and cognition (Daniels, Greer, & Neale, 2017; Shemla et al., 2014; Unzueta et al., 2012). For example, Whites are more likely to consider simple percentages when judging the diversity of an organization, while minorities are more likely to consider the distribution of leadership positions (Unzueta et al., 2012). Recent work has also demonstrated that "perceived" (subjectively judged) group diversity may predict team outcomes more than "objective" (demographically measured) group diversity (Shemla et al., 2014). What accounts for the gaps between visually perceived and subjectively judged group characteristics? How might these compare to consensus-based and objective measures of group characteristics, as well as predictive power for group outcomes? Our findings of accuracy in visual impressions imply a perhaps even stronger role of motivational and cognitive biasing factors, such as beliefs about the value of diversity and hierarchy, in creating social judgments.

If ensemble-coding mechanisms do indeed generate a single morphological face representation of the group, observers may be generating an accurate representation but then misinterpreting this representation due to motivational and/or cognitive biases. For instance, a racially diverse group may generate a single face representation that is racially ambiguous or nonprototypical; the same representation could also be generated from a nondiverse group if it is made of four, equally nonprototypical faces. Although the representation is accurate in both cases, as the representation feeds-forward to judgments or decisions, it may be skewed by observers' beliefs about racial prototypicality and diversity. Thus, people may not visually misperceive diversity, or other group characteristics, but rather they may experience initial perceptual accuracy, followed by distortion at the application phase.

Another possibility suggested by prior work is that top-down mechanisms may influence and interact with bottom-up perceptions (for reviews, see Dunning & Balcetis, 2013; Freeman & Ambady, 2011; Hughes & Zaki, 2015). Indeed, the SEA model of people perception suggests top-down influences (via different paths) may moderate the selection, extraction, or application of visual group information, particularly in extended or repeated interactions (Phillips et al., 2014; see also Brady & Alvarez, 2011; de Fockert & Marchant, 2008; Kteily, Sheehy-Skeffington, and Ho, 2016). For example, people may pay less attention to female group members than they do to male group members in certain contexts. As a result, the "group" from which observers are extracting diversity information may attentionally be limited to a subset of group members. Future work might also consider motivational influences of observers who are group members themselves versus outsiders to the group.

Further, it is likely that other factors will also inform judgments of group diversity. For instance, observers' explicit judgments about a group's diversity may be swayed by extraneous and/or nonvisual information (e.g., beliefs about the value of diversity; Homan et al., 2010), especially if they are not motivated to extract visual variance. Given our goal to link person perception research to ensemble coding, we specifically operationalize diversity as statistical variance of faces within a group, and we explicitly instructed participants to consider this variance. Outside of our controlled laboratory setting, it is likely that perceivers rely on preexisting cognitive associations when conceiving of diversity (Freeman & Ambady, 2011), such as the ratio or number of minority or women group members (Plaut et al., 2011; Unzueta et al., 2012). Such preexisting cognitive associations may be suppressing overall accuracy of variance perception in our findings, since participants may interpret groups that have low variance (e.g., all Black faces) as high diversity groups due to preexisting cognitive associations. An important shortcoming of our approach is that we cannot assess participants' own definitions of diversity, and how these likely map onto conventional notions of what diversity looks like; thus, we do not capture the role of such preexisting cognitive associations in shaping impressions of group diversity. Future work should explore the important interactions between top-down beliefs and bottom-up visual processes, to garner a fuller understanding of participants' naturalistic impressions of groups.

Relatedly, social judgments of group characteristics may depend on multiple summary statistics extracted and held simultaneously, or may even be informed by ensemble-coded information gleaned from other channels (e.g., voice, body: Piazza, Sweeny, Wessel, Silver, & Whitney, 2013; Sweeny et al., 2013). For instance, judgments of both mean and variance may feed into judgments of whether a group is judged as conceptually diverse. If having many women is cognitively associated with diversity, then a middling gender variance group with a high femininity mean may ultimately be judged as more diverse than a group with the same variance, but more masculine mean. Supporting this possibility, we find that group mean did sometimes affect judgments of variance (see pooled analysis and Studies 7a–b), albeit less so than did group variance itself.

In order to perceive a group's diversity or hierarchy, the bases or cues to diversity or hierarchy must be present. Thus, the strength of those cues may also affect gaps between perception and judgments. For instance, in our work, participants visually judged how "hierarchical" or how "diverse" different groups were; although participants were better than chance at both tasks, and these results replicated across multiple studies, we also found that diversity (race, gender) effect sizes were larger than hierarchy effect sizes. This may be because facial dominance is less directly tied to hierarchy than gender or race is tied to diversity. Diversity is often *defined* by variance in race or gender, and so our cues to diversity are highly valid. Hierarchy, on the other hand, is variance in status or power. We use facial dominance as a subtle proxy for status or power, but this is more indirect than in the case of our race and gender studies. Indeed, facial dominance is likely only linked to interpersonal dominance in certain contexts (e.g., Goetz et al., 2013; Hehman et al., 2015; Hehman, Sutherland, Flake, & Slepian, 2017; Todorov et al., 2015). Hierarchy may thus be an especially conservative test of ensemble-coding mechanisms' ability to extract variance from groups.

While theoretical work has distinguished hierarchy and diversity as critical and distinct properties of groups, our results suggest that these properties may not be so distinct visually. Thus, people may be left with room for their social interpretations of these characteristics to influence their later judgments or impressions. Socially, diversity is the degree of demographic difference, or variance in the horizontal dimension, while hierarchy is the degree of power difference, or variance in the vertical dimension (Gruenfeld & Tiedens, 2010). Visually, however, impressions of hierarchy and diversity both involve detecting variance in facial characteristics. People likely layer social meaning on to these perceptions, thus generating distinctions between vertical and horizontal types of variance. This layering of social meaning may also contribute to our test of hierarchy being more conservative relative to our test of diversity; hierarchy evokes a wide range of components and subjective understandings, while diversity often more straightforwardly involves demography. And, this likely affects whether means *and* variance affect ultimate judgments (see Studies 7a–b).

Subjective meaning may play a role during the application stage of people perception, despite accurate extractions, which might also explain how ensemble-coding supports perceptions of categorical versus continuous traits and features. Both the current and previous work suggest that ensemble-coding supports extraction of summary statistics regarding face dimensions that vary linearly along a continuum (de Fockert & Wolfenstein, 2009; Haberman & Whitney, 2009; Walker & Vul, 2014). However, previous work in person perception suggests that some features (e.g., gender) are categorically judged, despite variation existing along a continuum (Campanella, Chrysochoos, & Bruyer, 2001; but see Freeman & Ambady, 2011; Livingston & Brewer, 2002). Neuroimaging studies have suggested that brain areas supporting visual perception do indeed reflect continuous linear variation in features of individual faces, but this information becomes distorted into categories as higher-level cognitive processing occurs (Freeman, Rule et al., 2010; see also Freeman, Pauker et al., 2010). This suggests that interactions of percepts may be especially influential on the interpretation of cues as continuous versus categorical at the application phase. Thus, the separation of the extraction and application phases of people perception may help account for the interaction of visual, cognitive, and social processes involved in forming impressions of groups (see SEA model, Phillips et al., 2014; see also dynamic constructivism, Freeman & Ambady, 2011).

Conclusion

Person perception has been an important topic in visual, cognitive, social, and organizational literatures (Ambady & Skowronski, 2008). Despite the centrality of groups to social life, scant work has explored visual *people* perception—how humans visually perceive and form impressions of groups of people. Visual perceptions of groups are particularly important to social functioning: accurate perceptions of groups are needed to interact successfully within and across groups. We offer new insight into how social–cognitive and visual-perceptual processes interact to give humans "social vision," even for properties that uniquely emerge in groups, as opposed to individuals. The current studies suggest that observers—supported by ensemble-coding processes—rapidly and accurately visually perceive variance (diversity, hierarchy) of group characteristics from even minimal social cues. The present work expands current models of social vision beyond *person* perception to include critical processes that support *people* perception. Just as person perception outcomes predict individual social behavior, people perception outcomes are likely to predict group-based social behavior.

References

Abelson, R. P., Dasgupta, N., Park, J., & Banaji, M. R. (1998). Perceptions of the collective other. *Personality and Social Psychology Review, 2,* 243–250. http://dx.doi.org/10.1207/s15327957pspr0204_2

Adams, R. B., Ambady, N., Nakayama, K., & Shimojo, S. (2011). *The science of social vision*. New York, NY: Oxford University Press.

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

Ambady, N. E., & Skowronski, J. J. (2008). *First impressions*. New York, NY: Guilford Press Publications.

Anderson, C., Ames, D. R., & Gosling, S. D. (2008). Punishing hubris: The perils of overestimating one's status in a group. *Personality and Social Psychology Bulletin, 34,* 90–101. http://dx.doi.org/10.1177/0146167207307489

Anderson, C., & Brown, C. E. (2010). The functions and dysfunctions of hierarchy. *Research in Organizational Behavior, 30,* 55–89. http://dx.doi.org/10.1016/j.riob.2010.08.002

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12,* 157–162. http://dx.doi.org/10.1111/1467-9280.00327

Ariely, D. (2008). Better than average? When can we say that subsampling of items is better than statistical summary representations? *Perception & Psychophysics, 70,* 1325–1326. http://dx.doi.org/10.3758/PP.70.7.1325

Auster, C. J., & Ohm, S. C. (2000). Masculinity and femininity in contemporary American society: A reevaluation using the Bem Sex-Role Inventory. *Sex Roles, 43,* 499–528. http://dx.doi.org/10.1023/A:1007119516728

Balcetis, E., & Lassiter, G. D. (2010). *Social psychology of visual perception*. New York, NY: Psychology Press.

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20,* 286–290. http://dx.doi.org/10.1177/0963721411422522

Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly, 47,* 644–675. http://dx.doi.org/10.2307/3094912

Barsade, S. G., & Knight, A. P. (2015). Group affect. *Annual Review of Organizational Psychology and Organizational Behavior, 2,* 21–46.

Blair, I. V., Judd, C. M., & Fallman, J. L. (2004). The automaticity of race and Afrocentric facial features in social judgments. *Journal of Personality and Social Psychology, 87,* 763–778. http://dx.doi.org/10.1037/0022-3514.87.6.763

Boehm, C. (1999). *Hierarchy in the forest: Egalitarianism and the evolution of human altruism*. Cambridge, MA: Harvard University Press.

Brady, T. F., & Alvarez, G. A. (2011). Hierarchical encoding in visual working memory: Ensemble statistics bias memory for individual items. *Psychological Science, 22,* 384–392. http://dx.doi.org/10.1177/0956797610397956

Brauer, M., & Er-Rafiy, A. (2011). Increasing perceived variability reduces prejudice and discrimination. *Journal of Experimental Social Psychology, 47,* 871–881. http://dx.doi.org/10.1016/j.jesp.2011.03.003

Brewer, M. B., Hong, Y. Y., & Li, Q. (2004). Dynamic entitativity: Perceiving groups as actors. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 25–38). New York, NY: Psychology Press.

Campanella, S., Chrysochoos, A., & Bruyer, R. (2001). Categorical perception of facial gender information: Behavioural evidence and the face-space metaphor. *Visual Cognition, 8,* 237–262. http://dx.doi.org/10.1080/13506280042000072

Carré, J. M., & McCormick, C. M. (2008). In your face: Facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society of London, Series B: Biological Sciences, 275,* 2651–2656. http://dx.doi.org/10.1098/rspb.2008.0873

Castano, E., Yzerbyt, V., & Bourguignon, D. (2003). We are one and I like it: The impact of ingroup entitativity on ingroup identification. *European Journal of Social Psychology, 33,* 735–754. http://dx.doi.org/10.1002/ejsp.175

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology, 83,* 234–246. http://dx.doi.org/10.1037/0021-9010.83.2.234

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43,* 393–404. http://dx.doi.org/10.1016/S0042-6989(02)00596-5

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences, 20,* 324–335. http://dx.doi.org/10.1016/j.tics.2016.03.006

Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology, 40,* 61–149. http://dx.doi.org/10.1016/S0065-2601(07)00002-0

Dakin, S. C., Mareschal, I., & Bex, P. J. (2005). Local and global limitations on direction integration assessed using equivalent noise analysis. *Vision Research, 45,* 3027–3049. http://dx.doi.org/10.1016/j.visres.2005.07.037

Daniels, D. P., Neale, M. A., & Greer, L. L. (2017). Spillover bias in diversity judgments. *Organizational Behavior and Human Decision Processes, 139,* 92–105. http://dx.doi.org/10.1016/j.obhdp.2016.12.005

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

de Fockert, J. W., & Marchant, A. P. (2008). Attention modulates set representation by statistical properties. *Perception & Psychophysics, 70,* 789–794. http://dx.doi.org/10.3758/PP.70.5.789

de Fockert, J., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 62,* 1716–1722. http://dx.doi.org/10.1080/17470210902811249

DeRue, D. S. (2011). Adaptive leadership theory: Leading and following as a complex adaptive process. *Research in Organizational Behavior, 31,* 125–150. http://dx.doi.org/10.1016/j.riob.2011.09.007

Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size, and language in humans. *Behavioral and Brain Sciences, 16,* 681–735. http://dx.doi.org/10.1017/S0140525X00032325

Dunning, D., & Balcetis, E. (2013). Wishful seeing how preferences shape visual perception. *Current Directions in Psychological Science, 22,* 33–37. http://dx.doi.org/10.1177/0963721412463693

Evans, C. R., & Dion, K. L. (1991). Group cohesion and performance a meta-analysis. *Small Group Research, 22,* 175–186. http://dx.doi.org/10.1177/1046496491222002

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*. Atlanta, GA: Sage.

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118,* 247–279. http://dx.doi.org/10.1037/a0022327

Freeman, J. B., Pauker, K., Apfelbaum, E. P., & Ambady, N. (2010). Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology, 46,* 179–185. http://dx.doi.org/10.1016/j.jesp.2009.10.002

Freeman, J. B., Rule, N. O., Adams, R. B., Jr., & Ambady, N. (2010). The neural basis of categorical face perception: Graded representations of face gender in fusiform and orbitofrontal cortices. *Cerebral Cortex, 20,* 1314–1322. http://dx.doi.org/10.1093/cercor/bhp195

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego, CA: Academic Press.

Goetz, S. M. M., Shattuck, K. S., Miller, R. M., Campbell, J. A., Lozoya, E., Weisfeld, G. E., & Carré, J. M. (2013). Social status moderates the relationship between facial structure and aggression. *Psychological Science, 24,* 2329–2334. http://dx.doi.org/10.1177/0956797613493294

Gruenfeld, D. H., & Tiedens, L. Z. (2010). Organizational preferences and their consequences. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 1252–1287). New York, NY: Wiley. http://dx.doi.org/10.1002/9780470561119.socpsy002033

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., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision, 9,* 1–13. http://dx.doi.org/10.1167/9.11.1

Haberman, J., Lee, P., & Whitney, D. (2015). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *Journal of Vision, 15,* 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

Haberman, J., & Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Attention, Perception & Psychophysics, 72,* 1825–1838. http://dx.doi.org/10.3758/APP.72.7.1825

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. Wolfe & L. Robertson (Eds.), *From perception to consciousness: Searching with Anne Treisman* (pp. 339–349). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:osobl/9780199734337.003.0030

Hackman, J. R., & Katz, N. (2010). Group behavior and performance. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 1208–1251). New York, NY: Wiley. http://dx.doi.org/10.1002/9780470561119.socpsy002032

Halevy, N., Chou, E. Y., Galinsky, A. D., & Murnighan, J. K. (2012). When hierarchy wins evidence from the national basketball association. *Social Psychological & Personality Science, 3,* 398–406. http://dx.doi.org/10.1177/1948550611424225

Hall, J. A., & Friedman, G. B. (1999). Status, gender, and nonverbal behavior: A study of structured interactions between employees of a company. *Personality and Social Psychology Bulletin, 25,* 1082–1091. http://dx.doi.org/10.1177/01461672992512002

Hamilton, D. L. (2005). *Social cognition: Key readings*. New York, NY: Psychology Press.

Hamilton, D. L. (2007). Understanding the complexities of group perception: Broadening the domain. *European Journal of Social Psychology, 37,* 1077–1101. http://dx.doi.org/10.1002/ejsp.436

Hamilton, D. L., Sherman, S. J., Way, N., & Percy, E. J. (2015). Convergence and divergence in perceptions of persons and groups. In M. Mikulincer, P. R. Shaver, J. F. Dovidio, J. A. Simpson, (Eds.), *APA handbook of personality and social psychology, Vol. 2: Group processes*. Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/14342-009

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface-and deep-level diversity on group functioning. *Academy of Management Journal, 45,* 1029–1045. http://dx.doi.org/10.2307/3069328

Hehman, E., Flake, J. K., & Freeman, J. B. (2015). Static and dynamic facial cues differentially impact the consistency of social evaluations. *Personality and Social Psychology Bulletin, 41,* 1123–1134. http://dx.doi.org/10.1177/0146167215591495

Hehman, E., Sutherland, C. A., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology, 113,* 513.

Homan, A. C., Greer, L. L., Jehn, K. A., & Koning, L. (2010). Believing shapes seeing: The impact of diversity beliefs on the construal of group composition. *Group Processes & Intergroup Relations, 13,* 477–493. http://dx.doi.org/10.1177/1368430209350747

Homan, A. C., Van Kleef, G. A., & Sanchez-Burks, J. (2016). Team members' emotional displays as indicators of team functioning. *Cognition and Emotion, 30,* 134–149.

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117,* 1168–1187. http://dx.doi.org/10.1037/a0020463

Hughes, B. L., & Zaki, J. (2015). The neuroscience of motivated cognition. *Trends in Cognitive Sciences, 19,* 62–64. http://dx.doi.org/10.1016/j.tics.2014.12.006

Hunter, M. (2007). The persistent problem of colorism: Skin tone, status, and inequality. *Sociology Compass, 1,* 237–254.

Ip, G. W. M., Chiu, C. Y., & Wan, C. (2006). Birds of a feather and birds flocking together: Physical versus behavioral cues may lead to trait-versus goal-based group perception. *Journal of Personality and Social Psychology, 90,* 368–381. http://dx.doi.org/10.1037/0022-3514.90.3.368

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85,* 616–626. http://dx.doi.org/10.1037/0022-3514.85.4.616

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

Kteily, N. S., Sheehy-Skeffington, J., & Ho, A. K. (2016). Hierarchy in the eye of the beholder: (Anti)-egalitarianism shapes perceived levels of social inequality. *Journal of Personality and Social Psychology, 112,* 136–159.

Kurzban, R., & Leary, M. R. (2001). Evolutionary origins of stigmatization: The functions of social exclusion. *Psychological Bulletin, 127,* 187–208. http://dx.doi.org/10.1037/0033-2909.127.2.187

La Macchia, S. T., Louis, W. R., Hornsey, M. J., & Leonardelli, G. J. (2016). In small we trust: Lay theories about small and large groups. *Personality and Social Psychology Bulletin, 42,* 1321–1334. http://dx.doi.org/10.1177/0146167216657360

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

Linville, P. W. (1998). The heterogeneity of homogeneity. In Darley and Cooper (Eds.), *Attribution and social interaction: The Legacy of Edward E. Jones* (pp. 423–487). http://dx.doi.org/10.1037/10286-008

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

Lount, R. B., Jr., Sheldon, O., Rink, F., & Phillips, K. W. (2015). How much relationship conflict really exists?: Biased perceptions of racially diverse teams and their consequences for resource support. *Organization Science, 26,* 1351–1364. http://dx.doi.org/10.1287/orsc.2015.0994

Macrae, C. N., Alnwick, K. A., Milne, A. B., & Schloerscheidt, A. M. (2002). Person perception across the menstrual cycle: Hormonal influences on social-cognitive functioning. *Psychological Science, 13,* 532–536. http://dx.doi.org/10.1111/1467-9280.00493

Macrae, C. N., & Quadflieg, S. (2010). Perceiving people. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 428–463). New York, NY: Wiley. http://dx.doi.org/10.1002/9780470561119.socpsy001012

Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals, 2,* 351–398.

Magee, J. C., & Tiedens, L. Z. (2006). Emotional ties that bind: The roles of valence and consistency of group emotion in inferences of cohesiveness and common fate. *Personality and Social Psychology Bulletin, 32,* 1703–1715. http://dx.doi.org/10.1177/0146167206292094

Mannix, E., & Neale, M. A. (2005). What differences make a difference? The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest, 6,* 31–55. http://dx.doi.org/10.1111/j.1529-1006.2005.00022.x

Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology, 94,* 365–381. http://dx.doi.org/10.1037/0022-3514.94.3.365

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

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America, 109,* 16474–16479. http://dx.doi.org/10.1073/pnas.1211286109

Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science, 18,* 879–885. http://dx.doi.org/10.1111/j.1467-9280.2007.01995.x

Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception & Psychophysics, 70,* 772–788. http://dx.doi.org/10.3758/PP.70.5.772

Neuberg, S. L., Kenrick, D. T., & Schaller, M. (2010). Evolutionary social psychology. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 761–796). New York, NY: Wiley. http://dx.doi.org/10.1002/9780470561119.socpsy002021

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences, 18,* 566–570. http://dx.doi.org/10.1016/j.tics.2014.09.007

Omi, M., & Winant, H. (2014). *Racial formation in the United States*. New York, NY: Routledge.

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

Palmer, S. E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Park, B., & Judd, C. M. (1990). Measures and models of perceived group variability. *Journal of Personality and Social Psychology, 59,* 173–191. http://dx.doi.org/10.1037/0022-3514.59.2.173

Pauker, K., Weisbuch, M., Ambady, N., Sommers, S. R., Adams, R. B., Jr., & Ivcevic, Z. (2009). Not so black and white: Memory for ambiguous group members. *Journal of Personality and Social Psychology, 96,* 795–810. http://dx.doi.org/10.1037/a0013265

Phillips, L. T., Weisbuch, M., & Ambady, N. (2014). People perception: Social vision of groups and consequences for organizing and interacting. *Research in Organizational Behavior, 34,* 101–127. http://dx.doi.org/10.1016/j.riob.2014.10.001

Piazza, E. A., Sweeny, T. D., Wessel, D., Silver, M. A., & Whitney, D. (2013). Humans use summary statistics to perceive auditory sequences. *Psychological Science, 24,* 1389–1397. http://dx.doi.org/10.1177/0956797612473759

Plaut, V. C., Garnett, F. G., Buffardi, L. E., & Sanchez-Burks, J. (2011). "What about me?" Perceptions of exclusion and whites' reactions to multiculturalism. *Journal of Personality and Social Psychology, 101,* 337–353. http://dx.doi.org/10.1037/a0022832

Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Ditlmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology, 94,* 615–630. http://dx.doi.org/10.1037/0022-3514.94.4.615

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8,* 364–382. http://dx.doi.org/10.1207/s15327957pspr0804_3

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science, 3,* 176–179. http://dx.doi.org/10.1111/1467-8721.ep10770698

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues, 57,* 743–762. http://dx.doi.org/10.1111/0022-4537.00239

Rule, N. O., Adams, R. B., Jr., Ambady, N., & Freeman, J. B. (2012). Perceptions of dominance following glimpses of faces and bodies. *Perception, 41,* 687–706. http://dx.doi.org/10.1068/p7023

Sanchez-Burks, J., Bartel, C. A., Rees, L., & Huy, Q. (2015). Assessing collective affect recognition via the Emotional Aperture Measure. *Cognition and Emotion, 30,* 117–133.

Sanchez-Burks, J., & Huy, Q. N. (2009). Emotional aperture and strategic change: The accurate recognition of collective emotions. *Organization Science, 20,* 22–34. http://dx.doi.org/10.1287/orsc.1070.0347

Shemla, M., Meyer, B., Greer, L., & Jehn, K. A. (2014). A review of perceived diversity in teams: Does how members perceive their team's composition affect team processes and outcomes? *Journal of Organizational Behavior, 37,* 86–109.

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York, NY: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781139175043

Singular Inversions. (2006). *FaceGen 3.1 Ful l SDK Documentation*. Retrieved from http://facegen.com

Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist, 60,* 16.

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62,* 207–218. http://dx.doi.org/10.1037/0022-3514.62.2.207

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

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381,* 520–522.

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology, 84,* 558–568. http://dx.doi.org/10.1037/0022-3514.84.3.558

Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology, 25,* 395–410. http://dx.doi.org/10.1080/02643290802044996

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

Tong, K., Ji, L., Chen, W., & Fu, X. (2015). Unstable mean context causes sensitivity loss and biased estimation of variability. *Journal of Vision, 15,* 15. http://dx.doi.org/10.1167/15.4.15

Unzueta, M. M., Knowles, E. D., & Ho, G. C. (2012). Diversity is what you want it to be: How social dominance motives affect diversity construals. *Psychological Science, 23,* 303–309. http://dx.doi.org/10.1177/0956797611426727

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43,* 161–204. http://dx.doi.org/10.1080/14640749108400966

Walker, D., & Vul, E. (2014). Hierarchical encoding makes individuals in a group seem more attractive. *Psychological Science, 25,* 230–235. http://dx.doi.org/10.1177/0956797613497969

Watkins, C. D., Jones, B. C., & DeBruine, L. M. (2010). Individual differences in dominance perception: Dominant men are less sensitive to facial cues of male dominance. *Personality and Individual Differences, 49,* 967–971. http://dx.doi.org/10.1016/j.paid.2010.08.006

Whitney, D., Haberman, J., & Sweeny, T. D. (2014). From textures to crowds: Multiple levels of summary statistical perception. *The New Visual Neurosciences,* 695–710.

Williams, M. J., & Eberhardt, J. L. (2008). Biological conceptions of race and the motivation to cross racial boundaries. *Journal of Personality and Social Psychology, 94,* 1033–1047. http://dx.doi.org/10.1037/0022-3514.94.6.1033

Won, B. Y., & Jiang, Y. V. (2013). Redundancy effects in the processing of emotional faces. *Vision Research, 78,* 6–13. http://dx.doi.org/10.1016/j.visres.2012.11.013

Zitek, E. M., & Tiedens, L. Z. (2012). The fluency of social hierarchy: The ease with which hierarchical relationships are seen, remembered, learned, and liked. *Journal of Personality and Social Psychology, 102,* 98–115. http://dx.doi.org/10.1037/a0025345

Received October 13, 2016
Revision received September 29, 2017
Accepted October 3, 2017 ■