AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

0096-3445/18/$12.00

2018, Vol. 147, No. 11, 1660–1676
http://dx.doi.org/10.1037/xge0000450

Groups at a Glance: Perceivers Infer Social Belonging in a Group Based on Perceptual Summaries of Sex Ratio

Brianna M. Goodale and Nicholas P. Alt
University of California, Los Angeles

David J. Lick
Facebook, New York, New York

Kerri L. Johnson
University of California, Los Angeles

Human observers extract perceptual summaries for sets of items after brief visual exposure, accurately judging the average size of geometric shapes (Ariely, 2001), walking direction of a crowd (Sweeny, Haroz, & Whitney, 2013), and the eye gaze of groups of faces (Sweeny & Whitney, 2014). In addition to such actuarial summaries, we hypothesize that observers also extract social information about groups that may influence downstream judgments and behavior. In four studies, we first show that humans quickly and accurately perceive the sex ratio of a group after only 500 ms of visual exposure. We then test whether these percepts bias judgments about the group's social attitudes and affect the perceiver's sense of belonging. As the ratio of men to women increased, both male and female perceivers judged the group to harbor more sexist norms, and judgments of belonging changed concomitantly, albeit in opposite directions for men and women. Thus, observers judge a group's sex ratio from a mere glimpse and use it to infer social attitudes and interpersonal affordances. We discuss the implication of these findings for a heretofore overlooked hurtle facing women in male-dominated fields (e.g., science, technology, engineering, or mathematics): how the ratio of men to women provides an early visible cue that signals an individual's potential fit.

Keywords: ensemble coding, sex ratio, belonging, group norms, social vision

> I think the worst time was when I was all alone, after Sandra left. The public perception saw eight men, and then there was this little woman, hardly to be seen. But now, because I'm so senior, I sit toward the middle. I have Justice Kagan on my left, Justice Sotomayor on my right . . . So the public will see that women are all over the bench. They are very much part of the colloquy.
>
> —United States Supreme Court Justice Ruth Bader Ginsburg, 2015

Reflecting on her appointment to the United States Supreme Court in an interview at Georgetown University, Justice Ginsburg discussed the balance of men and women on the bench. At the time of her appointment, Justice Ginsburg was the second woman ever confirmed to the Supreme Court and has, at various points, served as the sole female voice. With the presence of women on the Supreme Court growing to three under President Obama's administration, Justice Ginsburg's comments gestured toward a profound perspective underlying the pursuit of equality: that the ratio of men to women on the Supreme Court was not only salient to the public, but that it also carried important social implications for society's perceived fit of women as arbiters of law.

Justice Ginsburg's reflections may characterize percepts that extend well beyond the judiciary. In a similar way, the ratio of men to women has yet to reach parity in other traditionally masculine fields. In science, technology, engineering, and mathematics (STEM), for example, women comprise only 24% of the American workforce (Beede et al., 2011), a statistic that indicates that observers of and participants in those fields are likely to encounter three times as many men as women. Here we test how one's first glimpse of a group's sex ratio shapes perceivers' mental representations of a group and their feelings of fit therein. We bring to bear existing research and methods from the vison and cognitive sciences to probe how the spontaneous perception of a group's sex ratio impacts perceivers' judgments of the group's social attitudes and interpersonal affordances (i.e., judgments about whether a group offers the individual opportunities for belonging). We contend that these group percepts and the inferences that they evoke

Brianna M. Goodale and Nicholas P. Alt, Department of Psychology, University of California, Los Angeles; David J. Lick, Facebook, New York, New York; Kerri L. Johnson, Department of Communication, University of California, Los Angeles.

Preliminary data and findings were presented by Kerri L. Johnson at the 2015 annual meeting of the Society for Experimental Social Psychology (SESP) Social Cognition Preconference as part of a larger talk on ensemble coding. Brianna M. Goodale presented select methods and results from Studies 1b, 2, and 3 at the 2016 annual meeting of the Society for Personality and Social Psychology Nonverbal Preconference in San Diego, California. Finally, Kerri L. Johnson presented select findings from Studies 1a, 1b, and Study 3 at the 2017 annual SESP meeting in Boston, Massachusetts. Brianna M. Goodale and David J. Lick each acknowledge support from the National Science Foundation Graduate Research Fellowship, Nicholas P. Alt acknowledges support from the National Defense Science and Engineering Graduate Fellowship, and Kerri L. Johnson is partially supported by the National Institute of Health Grant R01-HD082844.

Correspondence concerning this article should be addressed to Brianna M. Goodale, Department of Psychology, University of California, Los Angeles, 2455 Franz Hall, Box 951563, Los Angeles, CA 90095-1563. E-mail: bgoodale@ucla.edu

have broad implications for individuals in a range of fields, including the courtroom, and extending to tech start-ups and beyond.

The notion that merely perceiving a group might influence judgments of social attitudes and interpersonal affordances first requires that observers can (and do) achieve some degree of accuracy from minimal visual information, more generally. Indeed, people readily and rapidly form impressions about others, often achieving a remarkable level of accuracy (Hugenberg & Wilson, 2013; Macrae & Quadflieg, 2010). From merely glimpsing a face, people can judge a person's demographic identities (sex, age, race) and social traits (trustworthiness, dominance). These early percepts compel perceivers to approach or avoid others based upon appearance alone (Johnson, Iida, & Tassinary, 2012). Thus, the causes and consequences of social perception involving isolated faces are well understood. In daily life, however, people are often encountered in groups. Here we test both the accuracy and interpersonal implications of visually perceiving a group of unknown others. As has been observed in research probing the perception of individuals, we hypothesized that perceivers deduce similar characteristics of a group of people. Specifically, we predicted (and found) that observers' actuarial summaries about a group occur quickly and accurately, and that these summaries have important implications for subsequent social inferences involving a group's social attitudes and the interpersonal affordances the group provides.

Interpersonal Affordances: The Centrality of Belonging

Feelings of belonging are fundamental to the human experience (Baumeister & Leary, 1995). Indeed, belonging is so important that individuals will engage in goal-directed activity to satisfy the need to belong, and if their need goes unfulfilled, people actively seek out new groups or different relationships (Baumeister & Leary, 1995). Furthermore, persistent lack of belonging portends negative mental and physical health outcomes. For instance, lack of belonging fosters feelings of loneliness, anxiety, and depression (Baumeister & Tice, 1990; Conte, Weiner, & Plutchik, 1982; Hagerty & Williams, 1999; Lofland, 1982). Similarly, social isolation has been linked to increased cortisol activity (Kiecolt-Glaser et al., 1984), higher blood pressure (Cacioppo & Hawkley, 2003; Cacioppo et al., 2002; Hawkley & Cacioppo, 2003; Uchino, Cacioppo, & Kiecolt-Glaser, 1996), poorer immune functioning and healing ability (Cacioppo & Hawkley, 2003; Kiecolt-Glaser et al., 1984, 1987; Kiecolt-Glaser, Glaser, Cacioppo, & Malarkey, 1998), greater occurrence of psychosomatic health problems (DeLongis, Folkman, & Lazarus, 1988), and decreased sleep efficacy (Cacioppo & Hawkley, 2003; Cacioppo et al., 2002). Efforts to enhance social belonging appear to be effective, showing better self-reported health and less frequent health care visits over a 3-year period (Walton & Cohen, 2011). Belonging interventions also enhance motivation (Walton, Cohen, Cwir, & Spencer, 2012) and academic achievement (Cohen, Garcia, Apfel, & Master, 2006; Shnabel, Purdie-Vaughns, Cook, Garcia, & Cohen, 2013; Walton & Cohen, 2011). Thus, felt belonging confers numerous emotional and physical benefits.

The determinants of an individual's feelings of belonging or fit with a group are multifaceted. In some instances, the group's own self-characterization provides potent outward facing signals. For instance, an explicit statement of a company's philosophy in a job-fair brochure formed the basis of African American job applicants' trust in the company (Purdie-Vaughns, Steele, Davies, Ditlmann, & Crosby, 2008). More subtle cues to exclusion yield similar effects. In one study, observers responded to job postings and mock interviews that used either gender-exclusive (e.g., referring exclusively to "he" or "him") or gender-inclusive (e.g., "he or she" or "his or hers") language (Stout & Dasgupta, 2011). Among women, gender-exclusive language decreased felt belonging and motivation to pursue the position, but increased anticipated workplace sexism. The opposite was true among men; gender-exclusive language enhanced their motivation to pursue the position. Thus, both explicit and subtle language signaled distinct workplace environments to potential applicants, thereby affecting their likelihood to engage with the group. Moreover, once immersed within a group, individuals tend to distance themselves from traits that appear to be incongruent with the group's identity (Pronin, Steele, & Ross, 2004). Specifically, Pronin et al. (2004) found that women who enrolled in a high percentage of male-dominated courses (e.g., mathematics) eschewed feminine characteristics (e.g., wearing make-up) in an effort to enhance their fit within their social environment.

Although most research to date has tested perceivers' response to social cues about the group, some evidence indicates that visual cues can change individuals' identification with a group or within a certain context. Dubbed "ambient belonging," this sense of social fit arises based solely on the presence of inanimate objects placed within the environment. In a computer classroom, simply replacing Star Trek posters with nature posters boosted women's interest in pursuing computer science as a major (Cheryan, Plaut, Davies, & Steele, 2009). Similarly, the presence of more minority individuals in a job pamphlet led African American participants to report feeling more organizational trust and to anticipate being more comfortable in that work space (Purdie-Vaughns et al., 2008). Thus, in addition to subtler linguistic indicators, visible cues can serve to either enhance or curtail feelings of belonging.

Collectively, therefore, feelings of belonging show widespread importance, and they are informed by multiple factors ranging from subtle linguistic cues to incidental visual cues. For an observer attempting to discern whether a group affords the possibility of belonging, the earliest available information might be obtained visually from merely a glimpse, a possibility to which we now turn.

The Social Vision of Groups

Observers' visual perceptions of social spaces and the groups of people that inhabit them are likely to inform a range of subsequent judgments and behaviors. Although some evidence indicates that perceivers attend to visual cues when making judgments about a social space, the process by which perceivers aggregate across groups of objects and/or people to draw inferences remains poorly understood. Speaking to this point, Phillips, Weisbuch, and Ambady (2014) called for more empirical research into *people* perception to augment the sizable literature in *person* perception, arguing that people perception occurs much more frequently and has important implications for organizational behavior that rely on individuals' assessments of a social space (e.g., organization acculturation and team leadership). The current research is among the first to empirically test these processes, by examining how *visual* percepts (cf. verbal descriptions; Hamilton et al., 2015) about a group inform downstream social judgments of the group's likely social attitudes and interpersonal affordances.

Visually extracting the "gist." Early evidence from vision science provides support for our hypothesis that inferences about groups may have roots in objects processed as a collective set. In one study designed to probe the mechanisms of scene perception, participants briefly viewed pictures of everyday locations (e.g., streets, kitchens, and store counters) and indicated whether or not they had seen a given object (Biederman, Glass, & Stacy, 1973). Participants' judgments were more efficient when an object was logically improbable for a given scene (e.g., a car in the kitchen) than when it was commonplace (e.g., a teacup in the kitchen). The researchers concluded that perceivers may rely on a schema or script for a scene's appearance, allowing them to quickly rule out objects that do not appear to belong there. Subsequent research demonstrated that this process could occur in as little as 100 ms (Biederman, Rabinowitz, Glass, & Stacy, 1974) and that specific details about an object's shape or identity are not necessary for accurate scene categorization (Oliva & Schyns, 2000; Oliva & Torralba, 2001; Torralba, 2003; Torralba & Oliva, 2003). Furthermore, perceivers are more accurate at summarizing spatial details when extracting the overall scene gist than when recalling the locale of individual objects within it (Alvarez & Oliva, 2008, 2007). Thus, prior research implies that perceivers can rapidly extract the gist of a scene, making holistic judgments that incorporate prior knowledge quickly and efficiently at some cost to local detail (for a more thorough review of prior research on scene gist extraction, see Oliva & Torralba, 2007).

Whether and how perceivers similarly extract the gist of a group of objects/people has become a focus of more recent research. Given the vast amount of visual information available to perceivers on a moment-to-moment basis, perceptual mechanisms evolved to ease the processing burden (Attneave, 1954; Barlow, 1961; Geisler, 2008). In one such manifestation, the visual system tends to aggregate when it encounters group of similar objects to yield a perceptual summary, or *ensemble representation* (Alvarez, 2011). This process, *ensemble coding,* facilitates quick and accurate summary percepts for object groups. This process appears to be extensive, informing a range of aggregate judgments for a group of objects, including their size (Ariely, 2001; Chong & Treisman, 2003, 2005), the ratio of specific shapes (Burr & Ross, 2008; Ross & Burr, 2010), and their physical orientation (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). Importantly, similar ensemble representations occur for social percepts. From brief exposures, perceivers accurately summarize a group's average facial emotion (de Fockert & Wolfenstein, 2009; Haberman & Whitney, 2007), eye gaze direction (Sweeny & Whitney, 2014), gender (Haberman & Whitney, 2007), race (Thornton, Srismith, Oxner, & Hayward, 2014), and walking direction (Sweeny et al., 2013).

Interpersonal consequences of ensemble coding. Ensemble coding appears to guide visual representations for both descriptive features of inanimate objects and overarching group characteristics. Still, the interpersonal consequences of these descriptive summaries remain untested. In the same way that ensemble coding eases one's perceptual burden by producing actuarial summaries (Olshausen, 2003), it might also ease one's social burden by providing a heuristic for estimating a group's social affordances. In particular, we hypothesized that ensemble coding may provide a means of judging whether a group affords personal belonging—a fundamental social need (Baumeister & Leary, 1995).

Thus, research has extended the concept of ensemble coding from summaries of simple visual objects to summaries of complex person characteristics, although the interpersonal consequences of these summaries remain unexplored. Theoretically, perceivers might perceptually group targets to increase the efficiency of decision making (Olshausen, 2003). Summary representations of a group may therefore guide broader social judgments about a group in a heuristic fashion, enabling perceivers to estimate the behaviors, attitudes, and beliefs of individual group members given little more than the composition of the group itself. We hypothesized that this heuristic inference process may be especially important in the domain of personal belonging. Deciding in advance whether or not one belongs in a group presents a difficult challenge. Ensemble perception may provide an efficient solution to the challenge: upon seeing a group, perceivers may extract a summary representation of the "average" group member and use that representation to guide assumptions about the group overall.

The Present Studies

We tested these possibilities by integrating established methods from social, cognitive, and vision sciences. In Studies 1a and 1b, we first examined whether groups differing in the ratio of men to women were accurately perceived and then tested whether they elicited different perceptions of belonging from men and women. We predicted that perceivers would: (a) accurately perceive the group's sex ratio following brief presentation; (b) draw inferences about a groups attitudes based on sex ratio (i.e., the endorsement of sexist norms); and, (c) use this information to gauge their own feelings of belonging within the group. Study 2 replicated the observed effects using a design that affords inferences about the temporal process. Study 3 used a different methodology to extend our observations by testing perceivers' mental representations of the average group member as a function of a group's sex ratio.

Study 1a

Method

Design. Study 1a aimed to test how observers extract actuarial summary statistics and draw social inferences from brief exposure to social ensembles. A key goal in our approach was to test how the ratio of men to women in an ensemble affected judgments of social attitudes (i.e., harboring sexist attitudes) and interpersonal affordances (i.e., a sense of personal belonging). Study 1a used a 2 (Perceiver Sex: male vs. female) × 5 (Actual Sex Ratio: 0/8 men, 2/8 men, 4/8 men, 6/8 men, 8/8 men) mixed-model, quasi-experimental design in which each participant provided multiple judgments for each stimulus. As such, data were nested within stimuli and participant. Along with all subsequent studies described here, Study 1a received research ethics committee approval prior to the collection of data.

Participants. Ninety-one U.S.-based Internet users (57% women) completed an online study via Amazon's Mechanical Turk (MTurk). Sample size estimates and power analyses for nested data require knowing numerous parameters including, but not limited to, the number of Level 1 groups (e.g., how many trials seen per participant), the size of the effect, each random effect's variance, covariance estimates for random effects, regression coefficients, and the design effect (Aguinis, Gottfredson, & Culpepper, 2013; Snijders, 2005; Westfall, Kenny, & Judd, 2014). Given the novelty of our theory and the adaptation of vision science techniques, we had no prior effect sizes and variance estimates on which to base an a priori power analysis. We therefore recruited what we judged to be a conservatively high number of participants, given other samples used in social vision research. Consistent with scientific norms in vision science, traditional work in ensemble coding used a psychophysics approach and relied on repeated measures experimental designs; furthermore, the within subject design and high trial numbers (often exceeding 10,000 trials) in these studies allows for sufficient power even with much smaller samples than we typically see in social psychological research (e.g., ranging from 2 to 55 observers at most; see Albrecht & Scholl, 2010; Alvarez & Oliva, 2007; Ariely, 2001; Bai, Leib, Puri, Whitney, & Peng, 2015; Haberman, Brady, & Alvarez, 2015; Haberman & Whitney, 2010). By such standards, we sought to recruit a large sample of participants for a completely within-subjects design, given the cross-classified nature of our hypotheses. As a starting point for moving forward, we conducted post hoc power analyses using the parameter estimates derived from our model. As shown in Table 1, post hoc power analyses revealed we were fully powered to find our hypothesized results (achieved power for main effects and their interaction ranged from 0.872 to 1.0).[1]

Stimuli. Stimuli included 20 social ensembles depicting eight White faces. Ensembles varied in the ratio of men:women (0:8, 2:6, 4:4, 6:2, 8:0; see Figure 1). Target faces were drawn randomly from the Chicago Face Database (Ma, Correll, & Wittenbrink, 2015) and arranged at random into a 2 × 4 array. This procedure was repeated four times for each condition, yielding 20 unique ensembles, presented randomly within each counterbalanced block. Each ensemble image was saved as 2,386 pixel × 1,791 pixel JPG image to ensure standardization across trials and participants.

Procedure. Participants learned that they would provide judgments about groups of people. Upon entering the survey hosted by Qualtrics, participants provided informed consent.

Participants provided judgments in four counterbalanced blocks. Within each block, ensembles were presented individually and in random order for 500 ms prior to a judgment prompt. Prior research in vision science has consistently relied on 500-ms display time to allow for the extraction of ensemble summaries while preventing the serial processing of individual targets (Bai et al., 2015; Haberman et al., 2015; Haberman & Whitney, 2009; Sweeny & Whitney, 2014).

Two blocks probed participants' perceptual acuity in ensemble coding. In one block, participants estimated the Perceived Sex Ratio in each ensemble using a visual scale that depicted schematic men and women that varied in the ratio of men:women, from 0:10 to 10:0, increasing by 1 for each scale item (see Figure 2, Panel A). In another block, participants judged the gendered appearance of the average face in an ensemble. This Perceived Facial Masculinity scale depicted 21 computer-generated faces that varied anthropometrically in face appearance from highly feminine to highly masculine (see Figure 2, Panel B).

Two additional blocks probed participants' inferences about each ensemble's social attitudes and affordances. In one block, participants provided judgments about the norms within the group, defined to participants as "a group's spoken or understood rules about member behavior; they explain how members ought to act," with an example of raising one's hand in class. For each ensemble, participants rated the perceived importance of eight norms related to sexist ideology, presented in random order, using a 7-point scale anchored by "*Extremely Unimportant to the Group*" and "*Extremely Important to the Group*" (see the Appendix for complete scale). In a fourth block, participants indicated the extent to which they thought they would personally "fit in" and "belong" in each ensemble using 9-point scales anchored by "*Not at All*" and "*Extremely Well*" and "*I Wouldn't Belong at All*" and "*I Would Definitely Belong Here*", respectively. Because participants provided repeated responses within each of these trial blocks, we measured each scale's internal reliability by proportioning participants' response variance into its subcomponents and estimating how much variance was accounted for by the repeated-measures design, the actual stimuli themselves, and the items within each scale (for more details on how to estimate these variances, see Cranford et al., 2006). Observed reliability of change (R_C) was 0.62 and 0.75 for judgments of norms and belonging, respectively. As such, we averaged each scale to yield two composite indices—Perceived Sexist Norms and Belonging.

Upon completion, participants provided basic demographic information. They were then debriefed and compensated.

Statistical analysis. Because responses were nested within perceiver and target, we used cross-classified random coefficient models incorporating random intercepts and slopes in all analyses. Across all studies, perceiver sex moderated effects only where noted. In an effort to provide some context for the magnitude of our significant effects, we report both the intraclass correlation coefficient (ICC) for the null model, as well as the pseudo R^2 statistic for each significant effect. The ICC provides an indication of how much variance in the dependent variable occurs between participants (Hox, 2002; Kreft & de Leeuw, 1998; Lahuis, Hartman, Hakoyama, & Clark, 2014; Raudenbush & Bryk, 2002); thus, we would expect this to be generally small, and we hypothesize that most of the differences in perceiver judgments will depend on changes in the participant's response to a group (e.g., a different response to seeing zero men in an ensemble vs. seeing eight men). The ICC helps us understand how much variance in a model is explained by between-subjects as compared with within-subject effects. To provide additional context, we also report the R^2 derived from multilevel variance partitioning (R^2_{MVP}). Although some debate surrounds the best way to represent effect sizes in multilevel models given the complexity of the data (Aguinis et al., 2013; Lahuis et al., 2014; Peugh, 2010; Ugille, Moeyaert, Beretvas, Ferron, & Van den Noortgate, 2014), the R^2_{MVP} has been shown to be the least biased and most efficient effect estimator in cross-level, random-slope models (Lahuis et al., 2014); thus, it should be interpreted as the amount of overall variance in the dependent variable explained by the model and will accordingly only be reported once per model.

[1] We used a custom R script to run 10,000 Monte Carlo simulations for each hypothesized model, separately looking at each predicted main effect and interaction. Additional follow-up post hoc analyses revealed that, given our effect sizes and experimental design, we would have had adequate power to find our predicted effects with only 8 participants (four men and four women, power = 0.96). The R script for our post-hoc power analyses can be accessed via our GitHub repository at: https://github.com/UCLASoComLab/Publications

Table 1
Post-Hoc Power Analyses of Sample Size from Studies 1a and 1b

Study	Tested effect	Achieved power
1a	Main effect of Actual Sex Ratio on Perceived Facial Masculinity	1.0
1a	Main effect of Actual Sex Ratio on Belonging	.872
1a	Main effect of Actual Sex Ratio on Perceived Sexist Norms	1.0
1a	Main effect of Actual Sex Ratio on Perceived Sex Ratio	1.0
1a	Interaction effect of gender by Actual Sex Ratio on Belonging	1.0
1b	Main effect of Actual Sex Ratio on Perceived Facial Masculinity	.9982
1b	Main effect of Actual Sex Ratio on Belonging	.9958
1b	Main effect of Actual Sex Ratio on Perceived Sexist Norms	1.0
1b	Main effect of Actual Sex Ratio on Perceived Sex Ratio	1.0
1b	Interaction effect of gender by Actual Sex Ratio on Belonging	1.0

Note. All power analyses indicate the proportion of times the modeled effect was significant ($p < .05$) when tested on 10,000 randomly simulated datasets.

We used the "lme" function with residual maximum likelihood estimation (REML) in the open-source R packages "lme4" and "lmerTest" to model our four outcomes of interest, as all our criterion were continuous (Bates, Maechler, Bolker, & Walker, 2015; Kuznetsova, Brockhoff, & Christensen, 2015). Perceiver Sex was dummy coded in all models, with 0 representing participants who self-identified as male and 1 representing participants who self-identified as female. The Actual Sex Ratio variable was coded to reflect the number of men in each ensemble: 0, 2, 4, 6, and 8. In each statistical model for Study 1a, we first tested the main effects of Actual Sex Ratio and Perceiver Sex before adding their interaction term in a stepwise fashion. Our reported results include the unstandardized regression coefficients for each effect as well as its corresponding significance tests. All data and statistical analyses scripts described in Study 1a as well as in subsequent studies are publicly available at: https://github.com/UCLASoComLab/Publications

Results and Discussion

Accuracy in encoding actuarial summaries. To test the perceptual acuity of ensemble representations, we regressed Perceived Sex Ratio and Perceived Facial Masculinity (separately) onto Perceiver Sex, the Actual Sex Ratio, and their interaction. The ICC for the null model predicting Perceived Sex Ratio and Perceived Facial Masculinity was 0.01 and 0.32, respectively; most of the observed variance in actuarial summaries occurs due to within participant differences. Testing our hypotheses, we found that as the ensemble's ratio of men to women increased, Perceived Sex Ratio also increased ($B = 1.03$, $SE = 0.06$, $t = 18.16$, $p < .001$, $R^2_{MVP} = 0.78$), and judgments of the average face became decidedly more masculine ($B = 0.73$, $SE = 0.11$, $t = 6.71$, $p < .001$, $R^2_{MVP} = 0.37$; see Figure 3, Panels A and C). Thus, both of these measures show a high level of sensitivity to the sex ratio in an ensemble, indicating that observers extract such information readily and rapidly, on the basis of 500 ms of visual exposure.

Figure 1. Sample stimulus from Study 1a in which Actual Sex Ratio varied within each ensemble. This example depicts an Actual Sex Ratio of six men and two women. See the online article for the color version of this figure.

Social attitudes and affordances. Next, we tested our novel prediction that perceivers infer social attitudes and affordances from ensemble representations by regressing Perceived Sexist Norms and Belonging (separately) onto Actual Sex Ratio, Perceiver Sex, and their interaction. The ICC for the null models of Perceived Sexist Norms and Belonging were 0.39 and 0.32, respectively. As the ensemble's ratio of men to women increased, groups were judged to harbor more sexist attitudes ($B = 0.15$, $SE = 0.03$, $t = 5.89$, $p < .001$, $R^2_{MVP} = 0.32$), and perceivers' sense of belonging increased ($B = 0.25$, $SE = 0.06$, $t = 4.50$, $p < .001$, $R^2_{MVP} = 0.39$). The effect on Belonging varied by Perceiver Sex, however (interaction $B = -0.71$, $SE = 0.07$, $t = -10.65$, $p < .001$; see Figure 4, Panels A and C). As the ensemble's ratio of men to women increased, Belonging increased among men, but decreased among women (Bs $= 0.25$ and -0.46, SEs $= 0.05$ and 0.06, ts $= 4.68$ and -8.10, ps$<.001$). Thus, observers formed accurate representations of a group's sex ratio and drew inferences about its attitudes from brief exposures. Male-dominated groups were judged to be especially sexist, and perceptions of personal belonging increased as representation of one's own sex increased.

Study 1b

Study 1b sought to replicate the discoveries from Study 1a, employing a more rigorous design and using larger 12-person groups with variable ratios of men:women. In particular, we developed a fully randomized stimuli creation paradigm to remove (rather than control for) potential variance that could be explained by using the same target stimuli across participants. Furthermore,

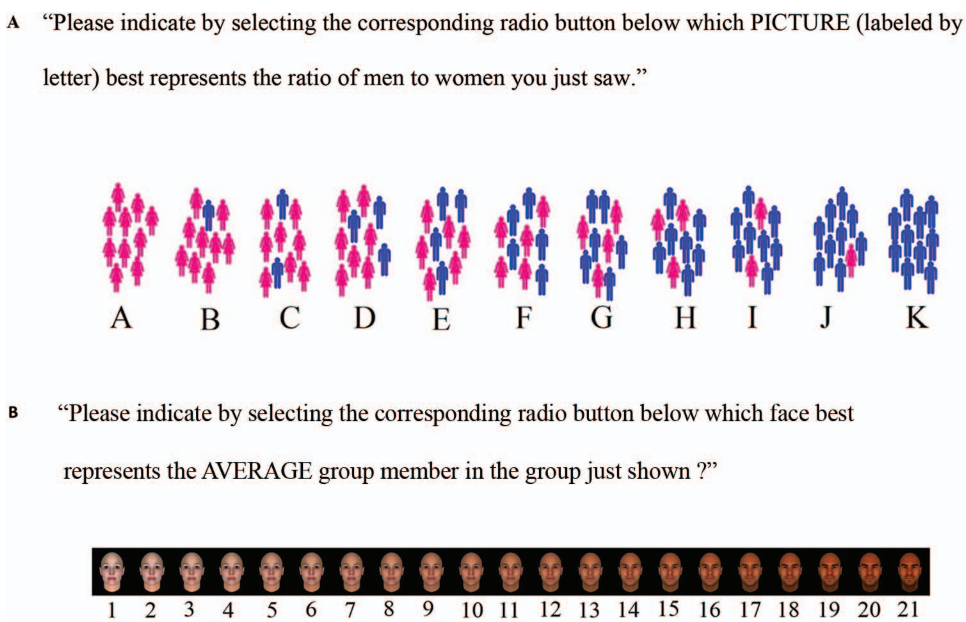

Figure 2. Dependent measures of actuarial summaries assessing Perceived Sex Ratio (Panel A) and Perceived Facial Masculinity (Panel B) used in Studies 1a, 1b, and 2. See the online article for the color version of this figure.

Study 1b tested whether ensemble coding can occur with more complex stimuli sets.

Method

Design. Study 1b sought to replicate findings from Study 1a using novel methodology and employed a 2 (Perceiver Sex: male vs. female) × 5 (Actual Sex Ratio: 0/12 men, 3/12 men, 6/12 men, 9/12 men, 12/12 men) quasi-experimental mixed-model design.

Participants. Seventy-six undergraduates from a large, public West Coast university participated in exchange for course credit. One participant who indicated their sex to be "non-binary/genderqueer" was excluded from our analyses, leaving a final sample of 75 participants (75% women). As with Study 1a, a priori power analyses were not estimable given the theoretical novelty of our study design. Study 1b sought to improve upon Study 1a by fully randomizing each ensemble, removing the need for specifying cross-classification in the multilevel model. Because power analyses for multilevel models rely on the structure of nested data (Snijders, 2005; Westfall et al., 2014), we lacked reasonable prior estimates for effect sizes and variance estimates of data *without* cross-classification. For this reason, we again recruited what we judged to be a conservative sample size and calculated post hoc power analyses via a customized R script. As shown in Table 1, our post hoc analyses revealed we were adequately powered to find our hypothesized effect, with achieved power for all main effects and interactions equal to or greater than 0.995.

Stimuli. In Study 1a, we manually created 20 social ensembles using a random number generator. In Study 1b, we created a custom Python script to automate the process. Individual faces were drawn randomly from the White targets within the Chicago Face Database (Ma et al., 2015) and placed into a 3 × 4 array to form 12-person social ensembles. As in Study 1a, the ratio of men:women varied: 0:12, 3:9, 6:6, 9:3, 12:0. The fully randomized stimulus generation in Study 1b occurred in real time; the custom Python script simultaneously generated and tracked the cumulative number of each actual sex ratio condition each participant had seen. It drew separately from the population of men and women in the overall stimulus set to generate each unique ensemble (sampling without replacement within each trial, but with replacement between trials). Furthermore, the Python script ensured uniform presentation of individual target stimuli such that each target face within the overall ensemble was 1,280 pixels × 960 pixels high, subtending visual angles of 29.62° × 23.09°. Participants were free to move their heads during the experiment. Within each block, participants judged 10 unique ensembles for each Actual Sex Ratio, presented randomly, resulting in 50 unique stimuli.

Procedure. Participants reported in-person to the lab, where they provided informed consent before beginning the study on one of three iMac desktop computers. The computers used in Study 1b and all subsequent in-lab studies reported here had a 27" Retina display with a 5,120-pixel × 2,880-pixel resolution. Participants learned that they would provide judgments about groups of people presented briefly on the computer screen. Each trial consisted of a centered fixation cross for 500 ms, a randomly generated ensemble for 500 ms, a blank screen for 500 ms, and finally, a judgment prompt. Before beginning judgment trials, participants completed three preview trials to familiarize themselves with the procedure. Preview trials depicted individual cartoon faces from a famous, recognizable 1970s cartoon series rather than pictures of actual people and were intended to alert participants to the pace of presentation.

As in Study 1a, participants completed four blocks in counterbalanced order. In each block, participants judged a total of 50 ensembles (10 per sex ratio). Judgment items were identical to those in Study 1a, with one exception. For Perceived Sex Ratio, we updated the visual scale to depict 12 schematic men:women to

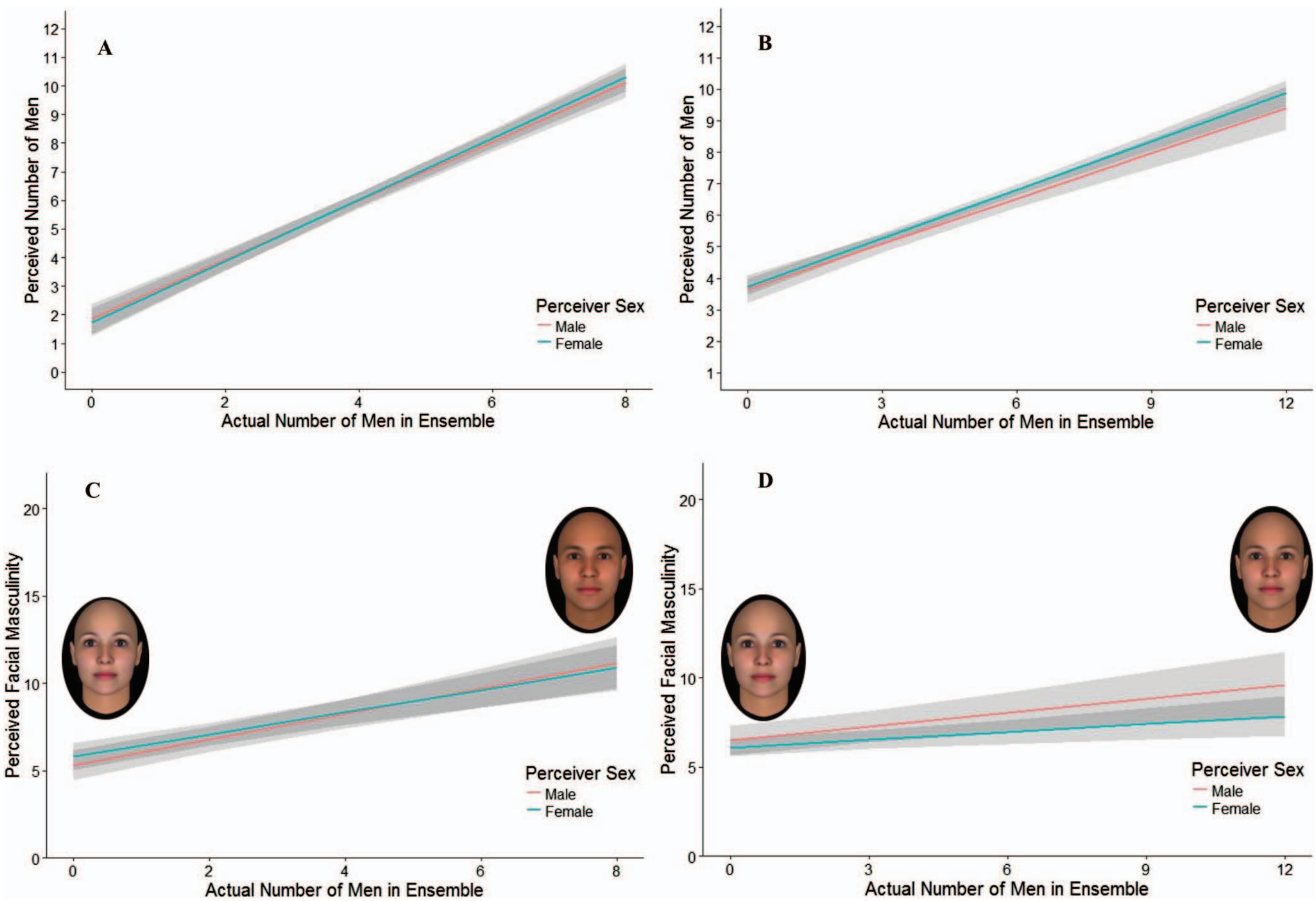

Figure 3. After 500-ms exposure, perceivers accurately report a group's sex ratio in Studies 1a (Panel A) and 1b (Panel B). Perceivers chose a more masculine-gendered face to represent the average group member in Studies 1a (Panel C) and 1b (Panel D) as the number of men in the group increased. The median face corresponding to participants' ratings for ensembles comprised of all-women or all-men are shown above their respective conditions. See the online article for the color version of this figure.

reflect our increase in ensemble size. Once again, participants provided basic demographic information and were debriefed.

Statistical analysis. We replicated our previous findings using the same analytic strategy described in Study 1a. Given our full-randomization of target stimuli, we no longer needed to account for target cross-classification so the Study 1b models below specified only an observation's nesting within participant.

Results and Discussion

Accuracy in encoding actuarial summaries. The ICC, calculated based on variance components of the null model with random intercepts for Perceived Sex Ratio, was 0.00. This indicates that all the variance in responses was within participant; that is to say, all participants responded nearly identically to the measure. For Perceived Facial Masculinity, the ICC equaled 0.34. Consistent with findings from Study 1a, as an ensemble's ratio of men to women increased, Perceived Sex Ratio also increased ($B = 0.48$, $SE = 0.04$, $t = 11.66$, $p < .001$, $R^2_{\mathrm{MVP}} = 0.39$), and judgments of the average face became more masculine ($B = 0.26$, $SE = 0.07$, $t = 3.54$, $p < .001$, $R^2_{\mathrm{MVP}} = 0.06$; see Figure 3, Panels B and D).

Social attitudes and affordances. The ICC for Perceived Sexist Norms and Belonging in the null models equaled 0.24 and 0.28, respectively. We again replicated our findings from Study 1a; as the ensemble's ratio of men to women increased, groups were judged to harbor more sexist attitudes ($B = 0.09$, $SE = 0.02$, $t = 5.24$, $p < .001$, $R^2_{\mathrm{MVP}} = 0.16$; see Figure 4, Panel B). Once again, the effect of Actual Sex Ratio on Belonging varied for men and women, interaction ($B = -0.26$, $SE = 0.04$, $t = -6.76$, $p < .001$, $R^2_{\mathrm{MVP}} = 0.08$). As an ensemble's ratio of men to women increased, Belonging increased among men, but decreased among women (*B*s = 0.10 and −0.16, *SE*s = 0.04 and 0.02, *t*s = 2.60 and −8.10, *p*s = .018 and < .001; see Figure 4, Panel D).

Collectively, Studies 1a and 1b found that ensemble representations accurately reflect the sex ratio of groups and that variations in sex ratio inform inferences about a group's sexist attitudes and personal belonging. That said, these studies could not test causal links between these judgments because of the randomization of blocks, a temporal factor that precludes tests of mediation (Baron & Kenny, 1986). We set out to test our hypothesis that actuarial

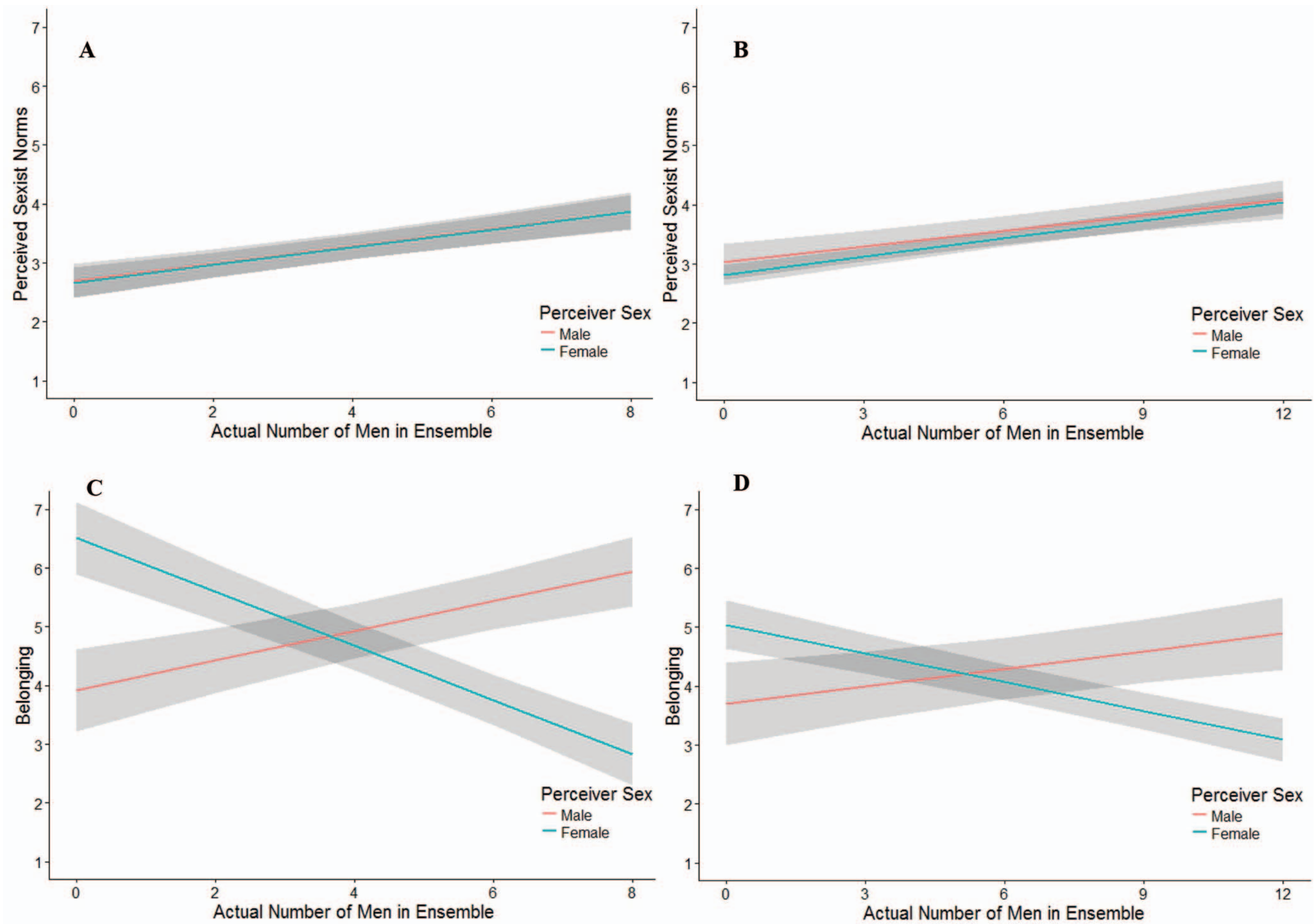

Figure 4. Participants view male-dominated groups as having more sexist norms in Studies 1a (Panel A) and 1b (Panel B). Their feelings of Belonging depended on the number of same-sex others in the group in Studies 1a (Panel C) and 1b (Panel D). See the online article for the color version of this figure.

summaries of a group drive perceivers' feelings of belonging in Study 2.

Study 2

In Studies 1a and 1b, we demonstrated that perceivers accurately extract actuarial summaries from a group of people. Furthermore, we showed that the group's sex ratio significantly impacts feelings of belonging and perceptions of sexist norms. Although offering convergent evidence that ensemble representations arise from instantaneous judgments about a group, Studies 1a and 1b could not offer insight into the underlying mechanisms; in Study 1a, while targets were identical across blocks, the data's cross-classification structure prevented us from running a mediation model testing our hypothesis. Study 1b no longer used a cross-classified structure, however, participants did not see the same targets across blocks because each stimulus was unique, preventing us from yoking participants' responses to a given image. A more direct test of our hypothesis required a test of moderated mediation, so we designed Study 2 to avoid cross-classification and allow participants to see the same ensembles across blocks.

Consistent with Studies 1a and 1b, we expected both male and female perceivers would accurately perceive increases in the ensemble sex ratio as the actual number of men in the ensemble increased. We predicted that these actuarial summaries would differentially mediate the effect of the Actual Sex Ratio on Belonging.

Method

Design. Study 2 aimed to test the mechanism by which differences in an ensemble's sex ratio impacts one's sense of belonging within the group. Study 2 used a 2 (Perceiver Sex: male v. female) $\times$ 5 (Actual Sex Ratio: 0/12 men, 3/12 men, 6/12 men, 9/12 men, 12/12 men) quasi-experimental mixed-model design in which participants provided multiple judgments for each stimulus.

Participants. Estimating desired sample size a priori for multilevel mediation analyses is a notoriously difficult task; while custom scripts are recommended as best practice to estimate sample size for a single multilevel regression equation (Muthén & Muthén, 2002), mediation analysis requires knowing threefold the amount of information a priori (Krull & MacKinnon, 1999). Additionally, mediation can occur at the individual and/or group level (e.g., for a given observation as well as for a participant across trials; Kenny, Kashy, & Bolger, 1998; Kenny, Korchmaros, & Bolger, 2003), and variance matrices will, as a result, differ across each component of the mediation analysis (Krull & MacKinnon, 1999). At the time of data collection, there were no readily available scripts or practical guidelines for estimating a priori sample size for multilevel mediation analysis. Solving this problem was beyond the scope and purpose of this paper. Thus, we recruited what we considered a conservative sample size, based off recommendations from Krull and MacKinnon (1999). According to their simulations, larger sample sizes (e.g., 50 participants or more) and greater than 30 trials per person decreased bias in estimating the coefficients in the mediator model. On the basis of their simulation findings, we recruited 58 college undergraduates (81% women) who completed Study 2 in a laboratory setting in exchange for course credit. Participants were only eligible if they had not participated in Study 1b.

Stimuli. Testing our proposed mediation required us to eliminate the stimulus cross-classification in our design while still ensuring each participant saw the same unique-to-them set of ensembles. We therefore created a custom Python script that randomly generated and stored 100 unique ensembles (20 per condition) as described in Study 1b. Here, however, the exact ensembles that were generated in the first test block were repeated in subsequent blocks. Thus, each participant viewed a randomly generated set of ensembles that was repeated within participant throughout subsequent blocks. This continuity in the design was essential for testing our hypotheses involving multilevel mediation without stimulus cross-classification. As in Study 1b, each stimulus was 1,280 pixels × 960 pixels, subtending visual angles of 29.62° × 23.09°.

Procedure. After providing informed consent, participants completed a series of preview trials identical to those used in Study 1b. Following the preview trials, participants completed two blocks of trials in which they judged a total of 100 ensembles (20 per sex ratio) for Perceived Sex Ratio and Belonging. The presentation order of dependent variables remained the same across participants. All other aspects of stimulus presentation were identical to Study 1b.

Statistical analysis. We used the statistical R package "mediation" (Tingley, Yamamoto, Hirose, Keele, & Imai, 2014) to test how an ensemble's sex ratio compels a sense of belonging. We recoded both Actual Sex Ratio and Perceived Sex Ratio to reflect the number of same-sex faces in each ensemble, relative to Perceiver Sex (e.g., ensembles of three men to nine women were coded as 3 for male participants and 9 for female participants), hereafter Actual and Perceived Same-Sex Others, respectively. This allowed us to collapse across Perceiver Sex to test a common causal process for both men and women. Our mediation model tested whether the relation between Actual Same-Sex Others and Belonging was mediated by Perceived Same-Sex Others. This model therefore tests whether one's sense of belonging stems, at least in part, from the first glimpse of a group of people.

The "mediation" analysis package in R requires the user specify a control group and a comparison group. We compared the proposed mediation model across four separate contrasts: 0 vs. 3 Same-Sex Others, 0 vs. 6 Same-Sex Others, 0 vs. 9 Same-Sex Others, and 0 vs. 12 Same-Sex Others.

Results and Discussion

We analyzed judgments using a multilevel mediation model. Using a quasi-Bayesian Monte Carlo approximation with 10,000 simulations, we estimated average direct effects and the average casual mediation effects of Actual and Perceived Same-Sex Others respectively on Belonging (Tingley et al., 2014). As predicted, the effect of Actual Same-Sex Others on Belonging was mediated by Perceived Same-Sex Others in the ensemble. Across all four contrast comparisons, Perceived Same-Sex Others partially mediated the effect of Actual Same-Sex Others on Belonging (see Table 2). The proportion mediated and relative effects were similarly sized across contrasts, suggesting a steady influence of actuarial statistics on judgments of social affordance. Of note, the perception of just three Same-Sex Others in a 12-person ensemble group (proportionately only 25%) was sufficient to increase perceivers' sense of belonging (see Figure 5). As the number of Actual Same-Sex Others in the group increased, so did judgments of Perceived Same-Sex Others, again indicating sensitivity to the sex ratio of ensembles. These percepts accounted for 14% of the variance between Actual Same-Sex Others and Belonging.

Study 3

Having established that perceivers accurately represent social ensembles and use these percepts to draw inferences about groups' social attitudes, we sought to better characterize the mental representations resulting from ensemble coding. In Study 3, we used reverse correlation, a method that provides a visual approximation of participants' mental representations (e.g., an ethnic in-group member; Dotsch & Todorov, 2012; Dotsch, Wigboldus, Langner,

Table 2
Mediation Models Comparing the Effect of Perceived Same-Sex Others on Actual Same-Sex Other's Influence on Belonging

Number of Actual Same-Sex Others	Direct effect	Indirect effect	95% CI for indirect effect	Proportion mediated
3	.57**	.09**	[.05, .14]	.14**
6	1.14**	.18**	[.09, .27]	.14**
9	1.71**	.27**	[.13, .41]	.14**
12	2.28**	.36**	[.18, .54]	.14**

Note. All mediation models compare the Actual number of Same-Sex Others in the ensemble to the control condition of 0 Same-Sex Others in the ensemble.
** $p \leq .01$.

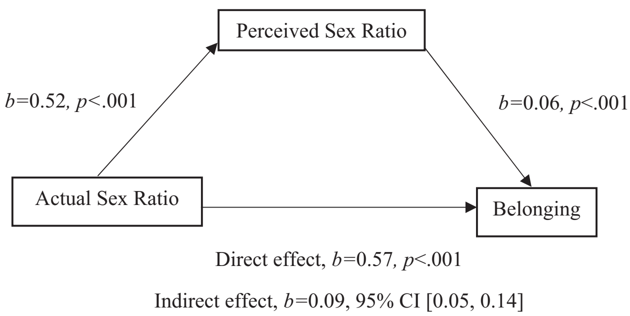

Figure 5. In Study 2, Perceived Sex Ratio mediated the effect of Actual Sex Ratio on feelings of Belonging.

& Van Knippenberg, 2008; Ratner, Dotsch, Wigboldus, van Knippenberg, & Amodio, 2014), to test how sex ratio affects one's mental representation of the average group member's face. Predominantly used by researchers in social cognition, reverse correlation methodology has also been successfully used to test questions pertaining to perceptual judgments in vision science (e.g., configural face processing; Sekuler, Gaspar, Gold, & Bennett, 2004). The images used in reverse correlation derive from the same base face, with random patterns of sinusoidal noise added or subtracted to them; when collapsing across participants' trials, the random noise cancels out and leaves an image that has been systemically, albeit unconsciously, selected by the participant. Although Studies 1 and 2 allowed us to examine explicit belonging based on actuarial summaries, Study 3 allowed us to visualize perceivers' mental representations of groups varying in their sex ratio.

Stage I Materials: Generating Mental Representations of the Average Group Member

Design. Study 3 used a between-subjects design in which we manipulated Condition; participants saw only majority-male (9 men, 3 women) or only majority-female (3 men, 9 women) 12-person ensembles.

Participants. Prior studies employing reverse correlation have ranged broadly in sample size (e.g., from 28 to 176 participants; see Dotsch & Todorov, 2012; Dotsch et al., 2008; Ratner et al., 2014). Although the first study to apply reverse correlation to group targets, we based our sample size off the approximate average sample size from prior reverse correlation studies. We thus recruited 110 college students, who received course credit for completing the study. We excluded data from 16 participants (1 was exluded due to computer failure, 3 did not follow instructions, and 12 had recently completed a near-identical task with full debriefing), leaving a final sample size of 95 (63% women).

Stimuli. Participants were randomly assigned to view ensembles that consisted of a majority men or majority women. A custom Python program fully randomized the selection of faces and their placement within the 12-person ensemble group. White male and female photographs were drawn without replacement within trials, but with replacement between trials, from the Chicago Face Database (Ma et al., 2015). As in Studies 1b and 2, each stimulus was 1,280 × 960 pixels, subtending visual angles of 29.62° × 23.09°.

Reverse correlation image classification task. Using the R package "rcicr" (Dotsch, 2016), we generated 700 pairs of greyscale images that depicted variations of an androgynous base face. These image pairs served as the forced-choice alternatives for our measure. This technique imposes visual noise over the base image, thus occluding it in systematic ways. Each image pair was constructed by both adding and subtracting a randomly generated visual noise pattern to/from the base image. Consistent with prior methods (Dotsch & Todorov, 2012; Dotsch et al., 2008; Ratner et al., 2014), we presented these image pairs simultaneously.

Procedure. In each of 700 trials, an ensemble appeared for 500 ms followed by a pair of reverse correlation test images. Participants indicated which of the two images best represented the average face in the group (see Figure 6, Panel A). Using the R package "rcicr" (Dotsch, 2016), we aggregated the noise patterns first across trials within a participant and then across participants to obtain a composite image that visualized the mental representation of the average group member for majority-male and majority-female ensembles (see Figure 6, Panels B and C).

Stage II materials: Evaluating Mental Representations of the Average Group Member from Majority-Male and Majority-Female Ensembles

Design. We compared the mental representations from majority-male and majority-female ensembles generated in Stage I using a within-subject design. An independent group of participants categorized each composite image as male or female and rated them on interpersonal characteristics (e.g., anger, warmth, approachability). We predicted that judgments obtained for composite images in Stage II would vary according to the sex ratios of ensembles presented in Stage I.

Participants. We recruited 142 U.S.-based Internet users to complete our study from MTurk. Our sample size is consistent with Stage II sample sizes from prior reverse correlation studies (see Dotsch & Todorov, 2012; Dotsch et al., 2008; Ratner et al., 2014). We excluded data from 22 participants (18 failed the manipulation check, 3 failed to indicate their gender, and 1 had an identical IP address as an earlier respondent). This left a final sample of 120 participants (45% women).

Procedure. In Stage II, participants judged the two composite images from Stage I in three counterbalanced blocks. In the Perceived Sex Category block, participants categorized the two mental representations as male or female. In the trait judgments block, participants judged how angry, friendly, warm, approachable, intelligent, competent, happy, and attractive each mental representation appeared using a series of 7-point Likert scales (1 = "*not at all*", 7 = "*very*"). Finally, in the Perceived Facial Masculinity block, participants indicated how masculine or feminine each mental representation appeared using a 7-point Likert scale (1 = "*extremely feminine*", 7 = "*extremely masculine*").

Statistical analysis. To test whether gender-linked judgments reflected the Actual Sex Ratio of ensembles in Stage I, we regressed Perceived Sex Category (using conditional logistic regression) onto Condition. We used paired samples *t*-tests to independently compare trait judgments and Perceived Facial Masculinity of the two mental representations. We used a Bonferroni correction to ensure the family-wise alpha level did not rise above 0.05; to do this, we divided the

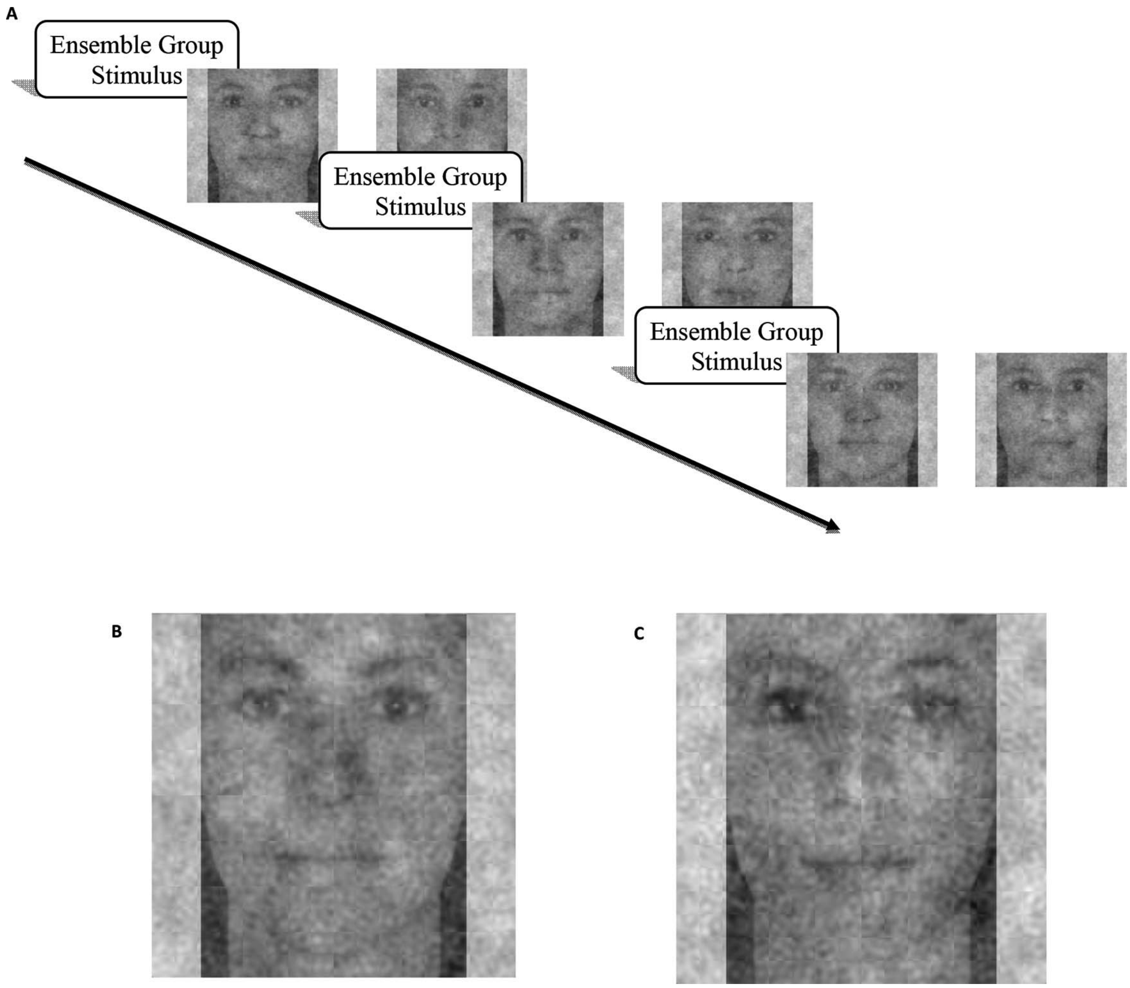

Figure 6. In Study 3, Stage I, participants completed 700 trials in which they saw a majority-male or majority-female ensemble. They indicated after each trial which of two faces most resembled the average group member (Panel A). Their choices were aggregated into two composite images derived from the majority-male (Panel B) and the majority-female (Panel C) conditions.

desired family-wise alpha-level of 0.05 by the number of paired samples *t*-tests ($n = 9$) we were running to arrive at the revised significance level of less than or equal to 0.0056. We used this corrected alpha level when deciding whether a given trait judgment varied by Condition in Stage II. Furthermore, perceiver sex did not moderate the effect of condition on any of the nine trait judgments.

Results and Discussion

First, we regressed Perceived Sex (using conditional logistic regression) and Perceived Facial Masculinity (paired samples *t* test) onto Condition. Relative to judgments of the composite image derived from majority-female ensembles, the composite derived from majority-male ensembles was 4.67 times more likely to be categorized as male ($B = 1.54$, $z = 2.42$, $p = .0155$) and was rated as significantly more masculine (see Table 3). Similar differences were obtained for gender-stereotyped judgments using paired samples *t*-tests with Bonferroni correction, such that composite images derived from majority-male, relative to majority-female, ensembles were rated as angrier, less warm, less approachable, less friendly, less happy, and less attractive. We observed no significant difference in perceived competence or intelligence between mental representations from majority-male and majority-female ensembles (see Table 3).

These findings reveal that, in addition to compelling accurate representations of a groups' sex ratio, conveying information about a group's attitudes, and informing one's sense of belonging within a group, ensemble coding also alters more implicit mental representations. As such, the mental representation of an average group member is imbued with gendered-linked appearance cues, including stereotypic cues to warmth, approachability, and anger.

General Discussion

Taken together, our findings indicate perceivers accurately discern the sex composition of a group of faces after brief visual exposure, forming mental representations of the average group member that embody gendered features and stereotypes (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972; Wood &

Table 3
Majority-Female Ensembles Rated More Favorably Than the Mental Representation of Majority-Male Ensembles

Traits	Mental representation of majority-male ensembles *M*	*SD*	Mental representation of majority-female ensembles *M*	*SD*	*t*-test
Angry	2.60	1.47	2.15	1.26	4.46**
Approachable	4.30	1.36	4.85	1.19	−5.05**
Attractive	4.16	1.32	4.83	1.18	−6.25**
Competent	4.46	1.21	4.55	1.12	−0.94
Friendly	4.19	1.28	4.78	1.12	−5.46**
Happy	3.92	1.34	4.68	1.20	−6.25**
Intelligent	4.42	1.19	4.55	1.22	−1.28
Masculine	4.00	1.28	3.09	1.30	6.71**
Warm	3.99	1.38	4.65	1.31	−5.25**

** $p < .0001$, with Bonferroni correction.

Eagly, 2010). Studies 1a and 1b served first as successful conceptual replications of prior work in vision science, demonstrating perceivers accurately encode high-level information about a group's composition. Furthermore, we also found that these split-second percepts gave rise to more abstract social evaluative judgments. Perceivers felt a greater sense of belonging as members of their own sex increased numerically, and both men and women inferred that male-dominated groups were likely to harbor sexist attitudes. Thus, we discovered that perceivers deduce group norms even in the absence of interaction with other group members; the ratio of men to women in an ensemble sends a rapid, strong nonverbal signal about whether future gender discrimination is likely to occur. Study 2 demonstrated that perceptions of same-sex others in an ensemble partially drive observers' feelings of fit and belonging; our social attitudes and affordances shift in response to actuarial summaries of groups, even from brief visual exposures. Finally, in Study 3, reverse correlation techniques revealed perceivers' mental representations of the average group member and found that representations of majority-female ensembles tend to be perceived as more welcoming and approachable than the representations of majority-male ensembles. Collectively, findings from our four studies illuminate the efficiency with which human perceivers extract visual information about a group and use it as a foundation on which to base important social judgments.

Implications for Vision Science

These findings have important implications across multiple domains. Methodologically, this research stands to enrich the experimental repertoire available to both vision scientists and social psychologists alike. Drawing inspiration from early studies on ensemble coding, we developed new experimental protocols that help push the boundaries of vision science methodology. Using novel methods that we adapted from prior work utilizing a psychophysics paradigm (e.g., de Fockert & Wolfenstein, 2009), we found that perceivers can accurately achieve summary percepts for a groups of individuals. Although perhaps unsurprising, this provides an important conceptual replication of prior work using a more widely accessible alternate method that affords testing of larger sample sizes. Additionally, the range of measures explored throughout our work expands the scope of inquiry available to vision scientists that heretofore have focused exclusively on narrowly focused aspects of perceptual accuracy. The development of these methods corroborates the robustness of the ensemble coding phenomena, thereby complementing and expanding upon earlier research.

Theoretically, this research also expands on prior work in vision science by demonstrating that, in addition to supporting accurate visual percepts, a mere glimpse of a group is also sufficient for perceivers to infer a broader array of social information. These overarching social impressions appear to occur just as rapidly and readily as the visual estimates, sparked by only a half-second visual exposure to an image. Thus, in addition to encoding the average distribution of various social categories within an ensemble (e.g., sex, Haberman & Whitney, 2007; or race, Thornton et al., 2014), observers also extract social meaning. Perceivers in our studies accurately extracted the *ratio* of men to women in an ensemble, and they used that knowledge to infer the groups' attitudes and estimate their own fit within the group. Their ability to do so was not immediately evident from prior research in vision science; we present novel evidence of the downstream social consequences of ensemble coding, with our findings suggesting that complex social inferences occur alongside lower level processing of groups from the briefest of exposures. Future researchers may seek to tease apart the precise threshold at which social attitudes shift as a result of changes in the number of same-sex others in the group. While we found adding three same-sex others to an ensemble significantly increased felt belonging and perceived discriminatory norms, it may be that in much larger groups many more same-sex others are needed to change perceivers' social affordances. It seems unlikely that seeing only three same-sex others in a 100-person ensemble will elicit similar feelings of belonging. It may be, instead, that a particular ratio of same-sex to dissimilar others must be achieved before perceivers feel like they belong.

Although our research draws on methodology and theory promoted by prior ensemble coding researchers, an alternative hypothesis from vision science could partially explain our results. Researchers have begun to tease apart the differences between ensemble coding and *numerosity*, defined as the ability to infer the relative number of objects in a set. Developing the theory of numerosity, Burr and Ross (2008) concluded that, like color or size, humans have a general sense of "numberness"; perceivers can guess roughly how many objects comprise a set, although environmental factors influence their accuracy. Additionally, numerosity appears to be a robust primary visual process coming online developmentally early (Burr & Ross, 2008; Ross & Burr, 2010; Xu, Spelke, & Goddard, 2005). Some initial evidence suggests that ensemble coding occurs over and above the effects of numerosity; while the two processes appear correlated, changes in the *number of* objects in a set does not seem to hamper perceivers' ability to extract summary information *about* the set (Sweeny, Wurnitsch, Gopnik, & Whitney, 2015).

According to the alterative numerosity hypothesis, perceivers in Studies 1 and 2 could have based their responses to the dependent variables on a general sense of the number of men to women in each group, rather than extracting an actuarial summary of the ratio or average facial masculinity. However, most prior research on numerosity has asked participants to make comparative judgments between sets of objects (e.g., which of two images contains more circles, Kirjakovski & Matsumoto, 2016). In contrast, across our four studies, perceivers extracted information about the *type* of object in each set, as well as the relative ratio of men to women. If the numerosity hypothesis were to underlie our effects, perceivers would have had to first determine whether each face in an ensemble was male or female before tallying the number per gender. While theoretically possible, these judgments would have had to occur quickly given the 500-ms presentation speed; perceivers would have had fewer than 42 ms on average to sort each of the 12 faces in Studies 1b, 2, and 3. Thus, it seems unlikely that numerosity alone can explain our findings. Rather, considering the rapid presentation of stimuli, it appears more likely that perceivers extracted holistic, summary information about the sex ratio of each group and used it to inform downstream social attitudes and affordances.

Furthermore, our work differs from prior research in numerosity by using reverse correlation. It is unlikely that a numerical sense of the number of men and women in the ensemble could lead to the systemic choice preferences in average group member seen in Study 3. Instead, it seems more plausible that perceivers draw on a mental representation extracted from the composite target images when choosing from the test pair images. Indeed, this explanation is consistent with the theory underlying reverse correlation as a methodology, which posits that visual categorization relies on a holistic extraction of features rather than the summary of individual parts (Mangini & Biederman, 2004). It is a limitation of the current work—albeit one reflected more broadly in the field—that we cannot definitively determine whether numerosity or ensemble coding processes underlie perceivers' response to groups varying in the ratio of men to women. While we believe our findings provide further evidence for ensemble coding, the social consequences of seeing similar or dissimilar others in a group remain the same: namely, increases in the ratio of men to women in a group affect feelings of fit and lead all perceivers to rate the group as more likely to endorse sexist norms. Although outside the scope of the current research, differentiating between numerosity and ensemble coding process effects on social judgments constitutes a very interesting and important avenue for future research.

Implications for Social Cognition

The current research also provides a viable means by which social cognitive scholars can probe the perceptual underpinnings of group judgments. Whereas it is widely recognized that the composition of a group is likely to impact social behaviors (Hackman, 1992), a majority of those observations have utilized immersive experience (e.g., Seger, Smith, Kinias, & Mackie, 2009). Here, we demonstrated that seeing a group, devoid of any other information, can be sufficient to influence perceivers' social judgments, using both explicit and indirect measures (e.g., reverse correlation). Additionally, from a theoretical perspective, our research highlights that inferences about a group's social attitudes and affordances may arise earlier than we previously knew. A perceiver need not be immersed in an environment, real or imagined, to feel the weight of being outnumbered; merely a glimpse of a group triggered consequential social percepts, including a broad sense of social belonging and inferences about gender-based attitudes. As such, this work complements prior observations that prolonged exposure to a seemingly hostile environment (e.g., women taking a math test in a room with geek paraphernalia; Cheryan et al., 2009) can lead to performance decrements and changes in gender attitudes (Pronin et al., 2004) by providing evidence for immediate shifts in social cognitions upon first espying a group. Our visual environment, including the people we see, rapidly and strongly impacts our perceptions of whether we belong and what behaviors to expect.

These findings form a foundation for answering more nuanced questions about group perception. To be sure, our manipulation of the sex ratio of ensembles provides an important first step toward understanding the broader social implications of group perception, but it is far from exhaustive. With these insights in hand, we are now poised to probe a range of factors that differ among individuals that inhabit a group and how they uniquely impinge on percepts. For instance, given that impressions of *individuals* exhibit a tight perceptual tethering between gender and race categories (Johnson, Freeman, & Pauker, 2012; Carpinella, Chen, Hamilton, & Johnson, 2015), percepts of ensembles that vary systematically along both dimensions simultaneously are likely to vary accordingly. The gendering of race may therefore bias both the accuracy of actuarial measures and the extremity of social inferences, depending on the specific racial composition of the group. Testing intersectional ensemble coding would provide a unique opportunity for future researchers to understand the relationship between race and gender in visual processing.

Additionally, our findings also provide a foundation for expanding the scope of inquiry to probe behavioral approach and avoidance tendencies when perceiving personally relevant or contextually situated groups. For example, based merely on a glimpse of a male-dominated STEM class, a woman may draw meaningful inferences that could impact her educational and professional trajectory: first, in her perception of the sex ratio; followed by her impression of group norms and stereotypes; and finally, manifesting in inferences about her own fit in the group. Similarly, percepts of real-world groups with established norms and/or visual similarities (e.g., a math club or sports team) might yield more extreme summary impressions than those we have observed based on a random selection of faces, in much the same way that described similarity enhances judgments of entitavity (Hamilton et al., 2015). If correct, this implies that our findings may even underestimate the potency of a spontaneous impression of a group. Thus, the visual heuristics that ease the task of social perception may fluctuate in strength as a function of personal relevance of social context, thus holding the potential to incur interpersonal costs well before observers interact with group members. This approach provides a viable way for testing such implications.

We wish to clarify, furthermore, that we do not believe this phenomenon unique to women. As our findings demonstrate, men also feel decreased fit in majority-female groups from only a mere glimpse. Although mental representations of majority-female groups may be perceived as less threatening overall (Alt, Goodale, Lick, & Johnson, 2017), nevertheless gender disparity may cause men to avoid female-dominated fields. In the 2014–2015 academic year, for example, only 12% of the Bachelor's degrees in registered nursing were awarded to men (U.S. Department of

Education, 2017). Other fields stereotyped as more feminine in nature, like education or psychology, show a similar disparity in educational achievement (Forsman & Barth, 2017; Shen-Miller & Smiler, 2015; U.S. Department of Education, 2017). Although researchers in the last decade have focused primarily on increasing women's participation in historically male-dominated fields (Shen-Miller & Smiler, 2015), our findings shed light on a broader mechanism by which men in female-dominated fields may feel similarly disadvantaged and out of place. Increasing gender parity across fields may positively boost both men and women's participation, with further research needed to probe the long-term effects of seeing similar or dissimilar others on vocational achievement.

Concluding Remarks

Our research provides the first empirical evidence confirming not only that the ratio of men to women in a group is readily and accurately perceived, but also that it drives feelings of fit and belonging within the group and broader inferences about the group. Given these findings, perhaps it is unsurprising that Judge Ginsburg, as one woman surrounded by eight men, felt "all alone." Indeed, such feelings are likely to be shared by naïve observers of any group exhibiting such an extreme ratio of men to women. Our findings also imply an important remedy. Each and every step toward equal representation on the Supreme Court (or any group for that matter) produces a commensurate shift in observers' actuarial representations of and interpersonal inferences about the group. In this same interview, Justice Ginsburg opined, "People ask me . . . When will there be enough women on the [Supreme] Court?" Her answer underscored the potency of the relationship between sex ratio and belonging: "When there are *nine*."

Context

The conceptualization of these study ideas grew organically out of individual and lab-wide discussions involving all authors. After reading recent vision science research on ensemble coding of shapes and people, the authors hypothesized that ensemble coding of people could also drive downstream social judgments, broader implications that had not been tested. Each of the first three authors pursued separate, but related, questions that were inherently social psychological in nature, leading to a multiproject collaboration. The studies described here will contribute to the first author's doctoral dissertation, with follow-up research already underway. All studies were conducted in the lab of Dr. Kerri L. Johnson and drew on her expertise in methodologies from cognitive psychology, vision science, and social psychology. The authors conferred with Dr. J. Krull and the Institute for Digital Research and Education for guidance on multilevel mediation analyses.

References

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management, 39,* 1490–1528. http://dx.doi.org/10.1177/0149206313478188

Albrecht, A. R., & Scholl, B. J. (2010). Perceptually averaging in a continuous visual world: Extracting statistical summary representations over time. *Psychological Science, 21,* 560–567. http://dx.doi.org/10.1177/0956797610363543

Alt, N. P., Goodale, B. M., Lick, D. J., & Johnson, K. L. (2017). Threat in the company of men: Ensemble perception and threat evaluations of groups varying in sex ratio. *Social Psychological and Personality Science*. Advance online publication. http://dx.doi.org/10.1177/1948550617731498

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

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19,* 392–398. http://dx.doi.org/10.1111/j.1467-9280.2008.02098.x

Alvarez, G. A., & Oliva, A. (2007). The representation of ensemble visual features outside the focus of attention. *Journal of Vision, 7,* 129. http://dx.doi.org/10.1167/7.9.129

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

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review, 61,* 183–193. http://dx.doi.org/10.1037/h0054663

Bai, Y., Leib, A. Y., Puri, A. M., Whitney, D., & Peng, K. (2015). Gender differences in crowd perception. *Frontiers in Psychology, 6,* 1300.

Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182. http://dx.doi.org/10.1037/0022-3514.51.6.1173

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117,* 497–529. http://dx.doi.org/10.1037/0033-2909.117.3.497

Baumeister, R., & Tice, D. (1990). Point-counterpoints: Anxiety and social exclusion. *Journal of Social and Clinical Psychology, 9,* 165–195. http://dx.doi.org/10.1521/jscp.1990.9.2.165

Beede, D., Julian, T., Langdon, D., McKittrick, G., Khan, B., & Doms, M. E. (2011, August). *Women in STEM: A gender gap to innovation* (Issue Brief No. 04-11). Rochester, NY: U.S. Department of Commerce, Economics and Statistics Administration. http://dx.doi.org/10.2139/ssrn.1964782

Biederman, I., Glass, A. L., & Stacy, E. W., Jr. (1973). Searching for objects in real-world scenes. *Journal of Experimental Psychology, 97,* 22–27. http://dx.doi.org/10.1037/h0033776

Biederman, I., Rabinowitz, J. C., Glass, A. L., & Stacy, E. W., Jr. (1974). On the information extracted from a glance at a scene. *Journal of Experimental Psychology, 103,* 597–600. http://dx.doi.org/10.1037/h0037158

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues, 28,* 59–79. http://dx.doi.org/10.1111/j.1540-4560.1972.tb00018.x

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology, 18,* 425–428. http://dx.doi.org/10.1016/j.cub.2008.02.052

Cacioppo, J. T., & Hawkley, L. C. (2003). Social isolation and health, with an emphasis on underlying mechanisms. *Perspectives in Biology and Medicine, 46*(Suppl. 3), S39–S52. http://dx.doi.org/10.1353/pbm.2003.0049

Cacioppo, J. T., Hawkley, L. C., Crawford, L. E., Ernst, J. M., Burleson, M. H., Kowalewski, R. B., . . . Berntson, G. G. (2002). Loneliness and health: Potential mechanisms. *Psychosomatic Medicine, 64,* 407–417. http://dx.doi.org/10.1097/00006842-200205000-00005

Carpinella, C. M., Chen, J. M., Hamilton, D. L., & Johnson, K. L. (2015). Gendered facial cues influence race categorizations. *Personality and Social Psychology Bulletin, 41,* 405–419. http://dx.doi.org/10.1177/0146167214567153

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology, 97,* 1045–1060. http://dx.doi.org/10.1037/a0016239

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

Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics, 67,* 1–13. http://dx.doi.org/10.3758/BF03195009

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313,* 1307–1310. http://dx.doi.org/10.1126/science.1128317

Conte, H. R., Weiner, M. B., & Plutchik, R. (1982). Measuring death anxiety: Conceptual, psychometric, and factor-analytic aspects. *Journal of Personality and Social Psychology, 43,* 775–785. http://dx.doi.org/10.1037/0022-3514.43.4.775

Cranford, J. A., Shrout, P. E., Iida, M., Rafaeli, E., Yip, T., & Bolger, N. (2006). A procedure for evaluating sensitivity to within-person change: Can mood measures in diary studies detect change reliably? *Personality and Social Psychology Bulletin, 32,* 917–929. http://dx.doi.org/10.1177/0146167206287721

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

DeLongis, A., Folkman, S., & Lazarus, R. S. (1988). The impact of daily stress on health and mood: Psychological and social resources as mediators. *Journal of Personality and Social Psychology, 54,* 486–495. http://dx.doi.org/10.1037/0022-3514.54.3.486

Dotsch, R. (2016). rcicr: Reverse correlation classification toolbox. (R package version 0.3.4.1) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/rcicr/index.html

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

Dotsch, R., Wigboldus, D. H. J., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science, 19,* 978–980. http://dx.doi.org/10.1111/j.1467-9280.2008.02186.x

Forsman, J. A., & Barth, J. M. (2017). The effect of occupational gender stereotypes on men's interest in female-dominated occupations. *Sex Roles, 76,* 460–472. http://dx.doi.org/10.1007/s11199-016-0673-3

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167–192. http://dx.doi.org/10.1146/annurev.psych.58.110405.085632

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., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17*(17), R751–R753. http://dx.doi.org/10.1016/j.cub.2007.06.039

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

Hackman, J. R. (1992). Group influences on individuals in organizations. In M. D. D. L. M. Hough (Ed.), *Handbook of industrial and organizational psychology* (2nd ed., Vol. 3, pp. 199–267). Palo Alto, CA: Consulting Psychologists Press.

Hagerty, B. M., & Williams, R. A. (1999). The effects of sense of belonging, social support, conflict, and loneliness on depression. *Nursing Research, 48,* 215–219. http://dx.doi.org/10.1097/00006199-199907000-00004

Hamilton, D. L., Chen, J. M., Ko, D. M., Winczewski, L., Banerji, I., & Thurston, J. A. (2015). Sowing the seeds of stereotypes: Spontaneous inferences about groups. *Journal of Personality and Social Psychology, 109,* 569–588. http://dx.doi.org/10.1037/pspa0000034

Hawkley, L. C., & Cacioppo, J. T. (2003). Loneliness and pathways to disease. *Brain, Behavior, and Immunity, 17,* S98–S105. Retrieved from http://www.demenzemedicinagenerale.net/images/pdf/Loneliness_and_pathways_to_disease.pdf

Hoeksma, J. B., & Knol, D. L. (2001). Testing predictive developmental hypotheses. *Multivariate Behavioral Research, 36,* 227–248. http://dx.doi.org/10.1207/S15327906MBR3602_05

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications. Quantitative methodology series*. New York, NY: Routledge.

Hugenberg, K., & Wilson, J. P. (2013). Faces are central to social cognition. In D. E. Carlston (Ed.), *The Oxford handbook of social cognition* (pp. 167–193). New York, NY: Oxford University Press.

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

Johnson, K. L., Iida, M., & Tassinary, L. G. (2012). Person (mis)perception: Functionally biased sex categorization of bodies. *Proceedings of the Royal Society B Biological Sciences, 279,* 4982–4989. http://dx.doi.org/10.1098/rspb.2012.2060

Kenny, D. A., Kashy, D., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of Social Psychology* (4th ed., pp. 233–265). New York, NY: McGraw-Hill. Retrieved from http://www.columbia.edu/~nb2229/docs/KennyKashyBolger1998-Data_analysis.pdf

Kenny, D. A., Korchmaros, J. D., & Bolger, N. (2003). Lower level mediation in multilevel models. *Psychological Methods, 8,* 115–128. http://dx.doi.org/10.1037/1082-989X.8.2.115

Kiecolt-Glaser, J. K., Fisher, L. D., Ogrocki, P., Stout, J. C., Speicher, C. E., & Glaser, R. (1987). Marital quality, marital disruption, and immune function. *Psychosomatic Medicine, 49,* 13–34. http://dx.doi.org/10.1097/00006842-198701000-00002

Kiecolt-Glaser, J. K., Glaser, R., Cacioppo, J. T., & Malarkey, W. B. (1998). Marital stress: Immunologic, neuroendocrine, and autonomic correlates. *Annals of the New York Academy of Sciences, 840,* 656–663. http://dx.doi.org/10.1111/j.1749-6632.1998.tb09604.x

Kiecolt-Glaser, J. K., Ricker, D., George, J., Messick, G., Speicher, C. E., Garner, W., & Glaser, R. (1984). Urinary cortisol levels, cellular immunocompetency, and loneliness in psychiatric inpatients. *Psychosomatic Medicine, 46,* 15–23. http://dx.doi.org/10.1097/00006842-198401000-00004

Kirjakovski, A., & Matsumoto, E. (2016). Numerosity underestimation in sets with illusory contours. *Vision Research, 122,* 34–42. http://dx.doi.org/10.1016/j.visres.2016.03.005

Kreft, I., & de Leeuw, J. (1998). *Introducing multivlevel modeling*. London, UK: SAGE. http://dx.doi.org/10.4135/9781849209366

Krull, J. L., & MacKinnon, D. P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review, 23,* 418–444. http://dx.doi.org/10.1177/0193841X9902300404

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *lmerTest: Tests in linear mixed effects models*. Retrieved from https://cran.r-project.org/package=lmerTest

Lahuis, D. M., Hartman, M. J., Hakoyama, S., & Clark, P. C. (2014). Explained variance measures for multilevel models. *Organizational Research Methods, 17,* 433–451. http://dx.doi.org/10.1177/1094428114541701

Lofland, L. H. (1982). Loss and human connection: An exploration into the nature of the social bond. In W. Ickes & E. S. Knowles (Eds.), *Personality, roles, and social behavior* (pp. 219–242). New York, NY: Springer. http://dx.doi.org/10.1007/978-1-4613-9469-3_8

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47,* 1122–1135. http://dx.doi.org/10.3758/s13428-014-0532-5

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

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

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9,* 599–620. http://dx.doi.org/10.1207/S15328007SEM0904_8

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41,* 176–210. http://dx.doi.org/10.1006/cogp.1999.0728

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42,* 145–175. http://dx.doi.org/10.1023/A:1011139631724

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *Trends in Cognitive Sciences, 11,* 520–527. http://dx.doi.org/10.1016/j.tics.2007.09.009

Olshausen, B. A. (2003). Principles of image representation in visual cortex. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1603–1615). Cambridge, MA: MIT Press. Retrieved from https://groups.oist.jp/sites/default/files/img/ocnc/2004/Olshausen.pdf

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4,* 739–744. http://dx.doi.org/10.1038/89532

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology, 48,* 85–112. http://dx.doi.org/10.1016/j.jsp.2009.09.002

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

Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology, 40,* 152–168. http://dx.doi.org/10.1016/S0022-1031(03)00088-X

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

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

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and data analysis methods: Applications and data analysis methods* (2nd ed., Vol. 1). Thousand Oaks, CA: SAGE.

Ross, J., & Burr, D. (2010). Vision senses number directly. *Journal of Vision, 10,* 10.1–8. http://dx.doi.org/10.1167/10.2.10

Seger, C. R., Smith, E. R., Kinias, Z., & Mackie, D. M. (2009). Knowing how they feel: Perceiving emotions felt by outgroups. *Journal of Experimental Social Psychology, 45,* 80–89. http://dx.doi.org/10.1016/j.jesp.2008.08.019

Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology, 14,* 391–396. http://dx.doi.org/10.1016/j.cub.2004.02.028

Shen-Miller, D., & Smiler, A. P. (2015). Men in female-dominated vocations: A rationale for academic study and introduction to the special issue. *Sex Roles, 72*(7–8), 269–276. http://dx.doi.org/10.1007/s11199-015-0471-3

Shnabel, N., Purdie-Vaughns, V., Cook, J. E., Garcia, J., & Cohen, G. L. (2013). Demystifying values-affirmation interventions: Writing about social belonging is a key to buffering against identity threat. *Personality and Social Psychology Bulletin, 39,* 663–676. http://dx.doi.org/10.1177/0146167213480816

Snijders, T. B. (2005). Power and sample size in multilevel linear models. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 3, pp.1570–1573), New York, NY: Wiley. http://dx.doi.org/10.1002/0470013192.bsa492

Stout, J. G., & Dasgupta, N. (2011). When he doesn't mean you: Gender-exclusive language as ostracism. *Personality and Social Psychology Bulletin, 37,* 757–769. http://dx.doi.org/10.1177/0146167211406434

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

Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: Ensemble perception of a crowd's gaze. *Psychological Science, 25,* 1903–1913. http://dx.doi.org/10.1177/0956797614544510

Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2015). Ensemble perception of size in 4- to 5-year-old children. *Developmental Science, 18,* 556–568. http://dx.doi.org/10.1111/desc.12239

Thornton, I. M., Srismith, D., Oxner, M., & Hayward, W. G. (2014). *Estimating the racial composition of groups of faces: An ensemble other-race effect*. Retrieved from https://www.openstarts.units.it/bitstream/10077/10518/2/T15.pdf

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software, 59,* 1–38. http://dx.doi.org/10.18637/jss.v059.i05

Torralba, A. (2003). Contextual priming for object detection. *International Journal of Computer Vision, 53,* 169–191. http://dx.doi.org/10.1023/A:1023052124951

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computation in Neural Systems, 14,* 391–412. http://dx.doi.org/10.1088/0954-898X_14_3_302

Uchino, B. N., Cacioppo, J. T., & Kiecolt-Glaser, J. K. (1996). The relationship between social support and physiological processes: A review with emphasis on underlying mechanisms and implications for health. *Psychological Bulletin, 119,* 488–531. http://dx.doi.org/10.1037/0033-2909.119.3.488

Ugille, M., Moeyaert, M., Beretvas, S. N., Ferron, J. M., & Van den Noortgate, W. (2014). Bias correction for standardized effect size estimates used with single-subject experimental designs. *Journal of Experimental Education, 82,* 358–374. http://dx.doi.org/10.1080/00220973.2013.813366

U.S. Department of Education. (2017). *Bachelor's, master's, and doctor's degrees conferred by postsecondary institutions, by sex of student and discipline division: 2014–2015*. Washington, DC: National Center for Education Statistics. Retrieved from https://nces.ed.gov/programs/digest/d16/tables/dt16_318.30.asp

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331,* 1447–1451. http://dx.doi.org/10.1126/science.1198364

Walton, G. M., Cohen, G. L., Cwir, D., & Spencer, S. J. (2012). Mere belonging: The power of social connections. *Journal of Personality and Social Psychology, 102,* 513–532. http://dx.doi.org/10.1037/a0025731

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143,* 2020–2045. http://dx.doi.org/10.1037/xge0000014

Wood, W., & Eagly, A. H. (2010). Gender. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed., pp. 629–667). New York, NY: Wiley. http://dx.doi.org/10.1002/9780470561119.socpsy001017

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science, 8,* 88–101. http://dx.doi.org/10.1111/j.1467-7687.2005.00395.x

Appendix

Perceived Sexist Norms Scale

Norms are a group's spoken or understood rules about member behavior; they explain how members ought to act. For example, in many classrooms, students understand the norm that they must raise their hands before they speak. Please answer the following questions about the group's potential norms.

How important, if at all, are the following norms to the majority of the group?

Norm	Extremely Unimportant to the Group		Unimportant to the Group		Important to the Group		Extremely Important to the Group
Members should treat women like they're good at math.*	1	2	3	4	5	6	7
Members should treat men differently than women.	1	2	3	4	5	6	7
It's acceptable to exclude some people.	1	2	3	4	5	6	7
Women have a lot to contribute to the group.*	1	2	3	4	5	6	7
No one should be made fun of.*	1	2	3	4	5	6	7
Group members should be friendly towards each other.*	1	2	3	4	5	6	7
Women should defer to men.	1	2	3	4	5	6	7
It is okay to ask for help.*	1	2	3	4	5	6	7

Note. Reverse coded items are marked with an asterisk.

Received June 15, 2017
Revision received March 9, 2018
Accepted March 22, 2018 ■