



# Gender differences in crowd perception

Yang Bai<sup>1</sup>, Allison Y. Leib<sup>1</sup>, Amrita M. Puri<sup>2</sup>, David Whitney<sup>1</sup> and Kaiping Peng<sup>3\*</sup>

<sup>1</sup> Department of Psychology, University of California, Berkeley, Berkeley, CA, USA, <sup>2</sup> Department of Psychology, Illinois State University, Normal, IL, USA, <sup>3</sup> Department of Psychology, Tsinghua University, Beijing, China

In this study, we investigated whether the first impression of a crowd of faces—crowd perception—is influenced by social background and cognitive processing. Specifically, we explored whether males and females, two groups that are distinct biologically and socially, differ in their ability to extract ensemble characteristics from crowds of faces that were comprised of different identities. Participants were presented with crowds of similar faces and were instructed to scroll through a morphed continuum of faces until they found a face that was representative of the average identity of each crowd. Consistent with previous research, females were more precise in single face perception. Furthermore, the results showed that females were generally more accurate in estimating the average identity of a crowd. However, the correlation between single face discrimination and crowd averaging differed between males and females. Specifically, male subjects' ensemble integration slightly compensated for their poor single face perception; their performance on the crowd perception task was not as poor as would be expected from their single face discrimination ability. Overall, the results suggest that group perception is not an isolated or uniform cognitive mechanism, but rather one that interacts with biological and social processes.

## OPEN ACCESS

### Edited by:

Nicole Meredith Else-Quest,  
University of Maryland,  
Baltimore County, USA

### Reviewed by:

Alexander Paul Marchant,  
London South Bank University, UK  
Andrea Hildebrandt,  
Ernst-Moritz-Arndt Universität  
Greifswald, Germany

### \*Correspondence:

Kaiping Peng,  
Department of Psychology, Tsinghua  
University, 401 Weiqing Building,  
Beijing, 100084, China  
pengkp@tsinghua.edu.cn

### Specialty section:

This article was submitted to  
Personality and Social Psychology,  
a section of the journal  
Frontiers in Psychology

**Received:** 12 November 2014

**Accepted:** 13 August 2015

**Published:** 02 September 2015

### Citation:

Bai Y, Leib AY, Puri AM, Whitney D  
and Peng K (2015) Gender differences  
in crowd perception.  
Front. Psychol. 6:1300.  
doi: 10.3389/fpsyg.2015.01300

**Keywords:** gender differences, ensemble coding, group perception, social interaction, statistical summary

## Introduction

Humans constantly interact with groups, and many of the known perceptual and social processes reflect this important fact of human life. Our perceptions and impressions of groups of people mediate our social interactions on a daily basis—what we think and perceive about groups shapes our long-term impressions and beliefs about them and guides our moment-to-moment behavior toward them (e.g., Darwin, 1872; Park and Hastie, 1987; Jans et al., 2011). The social psychology literature on group perception has established that people's group perception process is largely affected by social influences (Linville et al., 1989; Fiske and Von Hendy, 1992), while cognitive psychologists consider group perception to be an ability which does not differ across the population of adult observers (Herman and Whitney, 2007, 2009; de Fockert and Wolfenstein, 2009; Herman et al., 2009). Little research, however, has investigated if both social background and cognitive abilities shape differences in people's group perception habits. Given the complexity of group perception, we propose in the current paper that individual differences (both biologically and socially constructed) influence the evaluation of a crowd. In three studies, we aim to explore whether males and females, two groups that differ cognitively and socially, also differ in the perceptual processes that underline the very first seconds of group impression formation.

Cognitive and perceptual psychologists have taken a keen interest in group perception (Haberman and Whitney, 2007; Sweeny et al., 2013). In previous literature, researchers found that when people encounter groups or crowds of people, there is too much information for the visual system to process within a short period of time (Louie et al., 2007; Whitney and Levi, 2011). The visual system overcomes this challenge by taking advantage of redundancies (Ariely, 2001; Parkes et al., 2001; Chong and Treisman, 2003, 2005; Haberman and Whitney, 2007; for reviews, see Alvarez, 2011; Whitney et al., 2014). The process of summarizing similar items or extracting summary statistical information is referred to as *ensemble coding*. Rather than analyzing each item or person individually, the visual system calculates summary statistics of similar items—for example, the average facial expression in a crowd (Haberman and Whitney, 2007, 2009; de Fockert and Wolfenstein, 2009; Haberman et al., 2009). Thus, even when individuals are unable to recall specific faces (Haberman and Whitney, 2007, 2009; Haberman et al., 2009), and unable to notice changes in crowd membership (Haberman and Whitney, 2011), they are still able to calculate the summary statistics of the crowd (i.e., mean emotion or identity) using ensemble coding mechanisms. The common assumption shared by previous cognitive researchers is that group perception is a cognitive ability that is similar across all adult perceivers and is influenced by the object features of the crowd, e.g., the size of the crowd (Haberman and Whitney, 2009), the duration of the displaying faces (Haberman et al., 2009), the orientation of the crowd (Haberman and Whitney, 2007, 2009; Haberman et al., 2009), etc.

Groups are clearly at the heart of many social psychological phenomena (Hamilton and Gifford, 1976). Therefore, research on group impressions is ubiquitous: for example, group entitativity (e.g., Campbell, 1958; Lickel et al., 2000), homogeneity (e.g., Rothbart and Park, 2004; Hamilton, 2007), stereotyping (Schaller and O'Brien, 1992; Hamilton and Sherman, 1994), as well as important self-relevant processes such as self-identity (Turner, 1985; Spears et al., 2001), out-group threat (Fiske, 2002; Yuki and Yokota, 2009), conformity signals (Asch, 1956; Cialdini and Trost, 1998), or even urgent social information (Helbing et al., 2000) are all based on the outcomes of group perception. Differing from cognitive psychologists, social psychologists assume that group perception is a high-level process which is diverse across individuals—emphasizing social influences on people's evaluation, inference, and attribution processes (Linville et al., 1989). For instance, a substantial amount of previous research has found that people tend to perceive the group they belong to (their in-group) with more variance (e.g., Park and Rothbart, 1982; Linville et al., 1989; Ostrom and Sedikides, 1992). In addition, one's power status (Guinote et al., 2002), cultural orientation (Spencer-Rodgers et al., 2007), and ethnic background (Chen and Hamilton, 2012; Chen et al., 2014) were also found to influence one's group perception habits.

Although these methods and analyses in both perceptual and social psychology are valuable, their assumptions are not complete. On one hand, group perception is not solely a high level process—low level visual discrimination of facial identity, ethnicity, and emotion shapes the impressions we form of

other people, and social communication depends on these impressions as well (Asch, 1946; Ekman, 1973; Todorov, 2012). On the other hand, crowd perception necessarily fluctuates among groups of individuals due to their different social and cognitive backgrounds (Park and Rothbart, 1982; Guinote et al., 2002; Spencer-Rodgers et al., 2007; Leib et al., 2012). As a first step toward testing how both social and cognitive factors influence our group perception process, we chose gender as a test bed since it has long been regarded as a group factor which predicts various individual differences within the fields of both social psychology and cognitive psychology (Fiske, 2010). The origin and scope of psychological gender differences are some of the most challenging and fascinating questions for both social and cognitive psychologists, and require interdisciplinary investigations that go beyond simple contrasts between the genders (Eagly and Wood, 1999; Wood and Eagly, 2002). For cognitive psychologists, the innate biological differences such as steroidal gonadal hormone difference (e.g., Hines, 1982; Collaer and Hines, 1995), genetic disparity (e.g., Heath et al., 1999), and cognitive ability gap (e.g., Hyde and Linn, 1988; Miller and Halpern, 2014) reportedly contribute to the differences found between females and males. Social psychologists, on the other hand, tend to focus on social construction explanations, where the importance of societal role differentiations for females and males (e.g., Bohan, 1993) provides an explanation for observed discrepancies between the two genders.

Previous research has indicated that, in general, females perform better than males on single face perception tasks—including tasks using affective and non-affective faces (e.g., Rehnman and Herlitz, 2006; McBain et al., 2009; Heisz et al., 2013). This difference can be explained from both the cognitive and social perspectives. Cognitively, females' superior performance on single face perception tasks could be attributed to their innate perceptual advantages, such as low spatial frequency detection or better holistic processing (McBain et al., 2009). Socially, historical gender differences, such as males participating in larger social networks and females participating in smaller social networks, may also influence the way males and females view groups vs. individual faces (for review, see Belle, 1989; Benenson, 1990). However, no study to date has tested gender differences in group identity perception. On the one hand, it is reasonable to hypothesize that superior single face perception will contribute to performance on crowd averaging, as the crowd average is necessarily comprised of individual face assessments. On the other hand, however, from a social or cognitive perspective, group perception is also not simply a sum of single face judgments (Alvarez, 2011; Whitney et al., 2014). Previous cognitive research has found that ensemble percepts may slightly compensate for single item perception (Leib et al., 2012). In the present experiment, we investigated whether males and females differ in ensemble face perception and whether or not this difference interacts with their abilities to discriminate single faces. We hypothesize that consistent with previous findings, females will outperform males on single face perception which will largely contribute to their crowd perception performance. We also hypothesize that, compared to females, males will exhibit their own advantage in

group perception. This is consistent with previous findings in the cognitive literature that ensemble perception performance among prosopagnosics is far better than would be expected from their single face perception ability (Leib et al., 2012), and social observations that males' social network size is larger than females' on average (Benenson, 1990). We hypothesize that males' ensemble coding performance will be better than is predicted by their single face discrimination and even exceed females' group integration performance. In Experiment 1, we tested gender differences in homogeneous and heterogeneous crowd perception using an identity averaging task (Haberma and Whitney, 2007, 2009; Haberman et al., 2009). In Experiments 2a and 2b, we tested gender differences in single face perception (Chong et al., 2008; Haberman et al., 2009) and in ensemble coding for both redundant (crowds with repeated face sets and same variance) and non-redundant (crowds with non-repeated faces and changing variance) crowds.

This project was approved by both the University of California Berkeley Committee for the Protection of Human Subjects and Tsinghua University Committee for the Protection of Human Subjects. Written consent from all individuals for their voluntary participation was obtained before beginning of the study.

## Experiment 1

### Participants

Twenty-three undergraduate students (12 male;  $M = 20.17$  years,  $SD = 1.15$  years) at Tsinghua University in China participated in the experiment in exchange for optional course credit. Participants gave written informed consent according to the guidelines of the local research ethics committee.

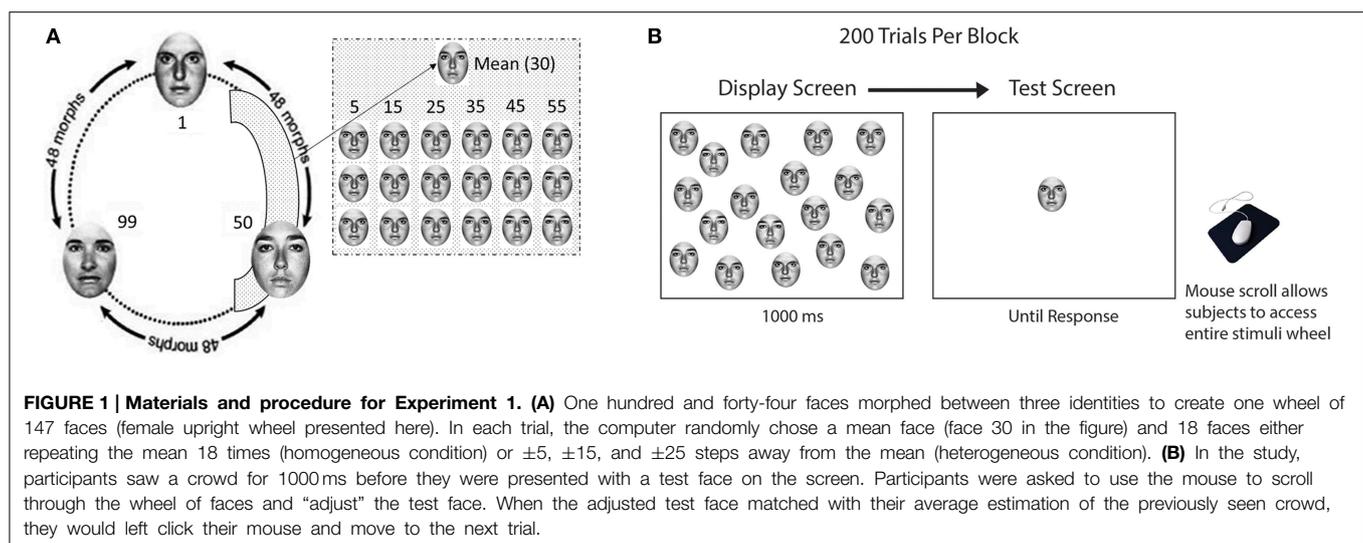
### Materials

In Experiment 1, our stimuli consisted of two arrays of faces. Each array contained 147 faces, generated by morphing together photographs of three different individuals (A, B, and C), each with a neutral expression (Figure 1). One array

was morphed from three grayscale female photos from the Ekman gallery (Ekman and Friesen, 1976), while the other array was morphed from three grayscale male photos from the same gallery. We used the program Morph 2.5 (Gryphon Software, San Diego, CA) to linearly interpolate each of the three pairs of original faces (A and B, B and C, C and A) to create 48 morphs per pair, yielding a "circle" of 147 neutral faces with no end or beginning (144 morphed pictures + 3 original = 147 faces total). We repeated this procedure to create an additional array of male faces (see Figure 1). During the experiment, participants viewed 18 faces drawn from either the male or female array. The faces were shown in a grid pattern on the computer screen. In two blocks, we displayed faces upright, while in two other blocks we displayed faces inverted counterbalanced across subjects. Subjects performed the experiment in a darkened room, seated at a distance of 80 cm from the monitor. For a more detailed description of stimulus creation and viewing conditions, please see the Supplementary Material.

### Procedure

For each trial, participants were presented with a crowd of 18 faces for 1000 ms. The 18 faces were drawn from the stimulus set and centered around a randomly chosen face (see Figure 1). After the crowd of faces disappeared, a single test face, which was randomly chosen from the entire wheel, was presented centrally on the screen (see Figure 1). Participants were instructed to "adjust" the test face until it looked exactly like the average identity of the previous 18-face crowd display. By moving the mouse left or right, participants viewed a face that seamlessly transformed into different identities (see Supplementary Video 1). Participants could click (using the left mouse button) on the face that matched their precise estimation of the mean. This allows us to objectively measure participants' ability to mentally represent the average identity of the crowd. Participants adept at mentally representing the average identity would select a morphed face that was quantitatively close to



the mean, whereas participants with a diminished capacity to represent the average identity would select a face that was quantitatively further from the mean of the crowd. The next trial began 200 ms after the button was pressed. Sets with consistent face gender (female or male) and orientation (upright or inverted) were presented in randomized blocks of 200 trials, in different orders for each participant.

Half of the displays in each block were heterogeneous crowds comprised of six identities, each repeated three times, for a total of 18 faces in the crowd. The six identities in each crowd were  $\pm 5$ ,  $\pm 15$ , and  $\pm 25$  steps away from the mean face in morph units. The mean face was randomly chosen from the 147 face array, but was not included in the display. The other half of the trials in a block contained homogeneous displays in which all 18 faces were the same. This homogeneous condition tested participants' recognition of a single identity (all the faces displayed were identical). In other words, the homogeneous condition is similar to the heterogeneous condition—except it does not require participants to average faces in order to respond correctly.

## Analysis

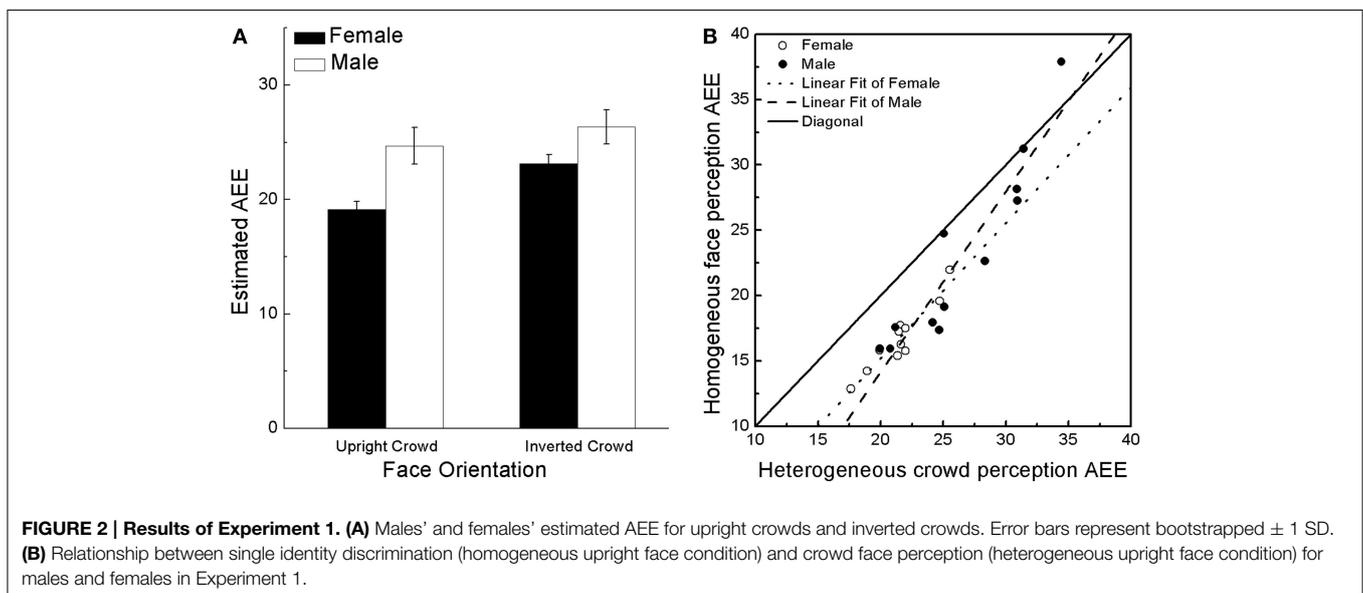
We analyzed participants' accuracy for each trial, using the following equation: Error = Actual Mean of Display (in morph units) - Participant's Response (in morph units). The units used in all equations are morph units. Importantly, we took the mean of the absolute error in each condition and obtained the Average Estimation Error (AEE).

## Results and Discussion

We conducted a 2 (subject gender: female vs. male)  $\times$  2 (picture orientation: upright vs. inverted)  $\times$  2 (display condition: heterogeneous vs. homogeneous)  $\times$  2 (picture gender: female vs. male) ANOVA. Subject gender was a between-subject variable. Consistent with previous research (Haberman and Whitney, 2009; Leib et al., 2012), the ANOVA revealed a significant main

effect of picture orientation: participants were significantly better in estimating the mean identity for upright faces compared to inverted faces [ $F_{(1, 21)} = 23.56, p < 0.001, \eta_p^2 = 0.53$ ]. The variance also had a significant effect on subjects' results; subjects performed much better when the displays were homogeneous [ $F_{(1, 21)} = 82.37, p < 0.001, \eta_p^2 = 0.80$ ]. Furthermore, as we predicted, we found a significant effect of subject gender [ $F_{(1, 21)} = 6.27, p < 0.05, \eta_p^2 = 0.23$ ]; females' AEE for mean identity ( $M = 21.14, SE = 1.26$ ) was smaller (more precise) than male subjects' AEE ( $M = 25.52, SE = 1.20$ ). The display variance  $\times$  participant's gender interaction was not significant [ $F_{(1, 21)} = 1.21, p = 0.28 > 0.05$ ], suggesting that females were overall more accurate in perceiving single face and ensemble coding group of faces. The orientation  $\times$  participants' gender interaction was marginally significant [**Figure 2A**;  $F_{(1, 21)} = 4.07, p = 0.057, \eta_p^2 = 0.16$ ]. Simple effect analysis of this interaction revealed that when the display faces were upright, females performed significantly better than males [ $F_{(1, 21)} = 8.82, p < 0.01, \eta_p^2 = 0.30$ ]; when the display faces were inverted, the two genders' performance was only marginally different [ $F_{(1, 21)} = 3.12, p = 0.092, \eta_p^2 = 0.13$ ]. This result suggests that females are not simply better at identifying simple shape or color contrasts in the crowd of faces (the shapes and color contrasts are still available when the crowd is inverted). Instead, females are actually better at ensemble coding a crowd of faces as they would be presented in typical social settings (upright). In addition, the interaction between participants' gender and stimulus gender was not significant [ $F_{(1, 21)} = 0.40, p = 0.532, \eta_p^2 = 0.02$ ], excluding the possible own-gender effect which might lead to the observed gender difference.

The ANOVA result revealed that females performed better than males in the homogeneous upright crowd condition and in the ensemble group perception condition. However, the exact nature of females' perceptual advantage remained unclear. To further examine the relationship between single face



perception ability and ensemble coding ability for both genders, we examined the correlation between the heterogeneous and homogeneous conditions (**Figure 2B**). We fitted all subjects' AEE with a linear function  $Y = bX + c$ , in which each subject's AEE for the homogeneous upright condition was  $Y$  and AEE for the heterogeneous upright condition was  $X$ . The slope parameter ( $b$ ) of the fitted line reveals the relationship between sensitivity to the heterogeneous upright face condition (mixture of identities) and sensitivity to the homogeneous upright face condition (one identity). A steeper slope indicates that individuals tend to perform better in the ensemble coding condition relative to their homogeneous identity discrimination. The result showed that males had a steeper slope compared to females ( $b_{\text{male}} = 1.38$ ;  $b_{\text{female}} = 1.04$ ). Although a permutation test<sup>1</sup> showed that the difference between the slopes for the two genders was not significantly different ( $p = 0.149$ ), the males' slope coefficient was greater than that of the females, suggesting that despite their relatively poor sensitivity to single identities, male subjects' sensitivity to crowd identity might be slightly increased.

Consistent with previous research showing that females were better at discriminating single faces (e.g., Rehnman and Herlitz, 2006; McBain et al., 2009), these data suggest that females are relatively better at discriminating a single identity (homogeneous condition in this case). Furthermore, these results are the first to show a novel gender difference in group perception. Our most robust finding is that females perceive the mean identity of a crowd (heterogeneous condition) more accurately than males.

## Experiment 2a

In Experiment 1, it was unclear whether females' relatively better performance was due to their precision in single face discrimination or due to an advantage in integrating the faces in the display (i.e., ensemble coding). One limitation of Experiment 1 was that it did not contain a true single face discrimination task. In the baseline homogeneous condition of Experiment 1, participants viewed a crowd of faces (albeit the same face repeating 18 times). Another limitation is the variability of the crowds displayed in Experiment 1. Previous research has found that high variation in the heterogeneous display negatively impacts ensemble coding performance (Haberman and Whitney, 2009). It is therefore possible that the variance between identities in the heterogeneous crowd of faces in Experiment 1 was higher than ideal for ensemble coding experiments. Additionally, the sample size (23) may be too small to draw definitive conclusions about the differences in performance between males and females. Therefore, by reducing the variance of the displayed faces and

creating a true "single face" condition, Experiment 2a aimed to replicate the ensemble coding result of Experiment 1, and to determine whether females' advantage was due to superior single face perception or superior integration abilities.

## Participants

Sixty undergraduate students at Tsinghua University in China participated in exchange for an optional course credit. Two subjects did not complete all conditions; their data was excluded from the analysis. The remaining 58 were comprised of 28 females and 30 males ( $M = 18.81$  years,  $SD = 1.01$ ).

## Materials

The stimuli were identical to the stimuli used in Experiment 1, except that we varied the number of faces displayed. For each trial, participants were presented with 1, 4, 8, or 12 faces. Trials with varying set sizes were randomly interleaved. Each face displayed was 3 cm by 3.7 cm, subtending 2.86° of the visual angle. All stimuli were viewed on a 19 inch monitor with resolution of 1280 × 1024 pixels, with a 75 Hz refresh rate. Subjects performed the experiment in a darkened room, seated at a distance of 80 cm from the monitor.

## Procedure

For Experiment 2a, face gender (female or male) and orientation (upright or inverted) was presented in blocks of 200 trials each. In each trial, the display included 1, 4, 8, or 12 faces for 500 ms. For the real single face condition, a random face was chosen and displayed on the screen for each trial. For the crowd condition of 12 faces, just as before, the mean face was randomly chosen from the array of 147 faces and never shown to participants; the faces displayed were three and nine steps away from the mean face in either direction, with a consistent variance. This design yields a crowd of 12 faces with four distinct identities, each repeated three times. Crowds consisting of four or eight faces contained repeated identities one or two times, respectively. As in the first experiment, participants used the method of adjustment to report the mean identity of the crowd. The next trial began 200 ms after the participant registered their response.

## Results and Discussion

Participants' AEE under different conditions was measured as described in Experiment 1. We conducted a 2 (subject gender: female vs. male) × 2 (picture orientation: upright vs. inverted) × 4 (display condition: 1, 4, 8, or 12 displayed faces) × 2 (picture gender: female vs. male) ANOVA. Only subject gender was a between-subject variable. The results were similar to Experiment 1. The ANOVA revealed a significant main effect of picture orientation; participants were significantly better in estimating the mean identity for upright faces compared to inverted faces [ $F_{(1, 56)} = 90.30$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.62$ ]. The number of display face(s) also significantly influenced people's sensitivity to crowd identity [ $F_{(3, 56)} = 8.57$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.13$ ]. Bonferroni *post-hoc* tests suggested that participants perceived the single face more accurately than they ensemble coded a group of faces, regardless of group size ( $M_{1\text{face}} = 23.44$ ,  $SE = 0.48$ ;  $M_{4\text{faces}} = 24.79$ ,  $SE = 0.50$ ;  $M_{8\text{faces}} = 24.52$ ,  $SE =$

<sup>1</sup>We used a permutation test commonly used in statistics (Good, 2000; Chau et al., 2004). Specifically, we created two permuted groups by randomly shuffling the male and female labels and rearranging paired single face AEE and multi-face AEE. For each permutation, we calculated the slope of each permuted subgroup (permuted "male" and "female") and subtracted the permuted "male" group's slope from the permuted "female" group's slope. This procedure was repeated 1000 times, yielding a null distribution of slope differences. By comparing the original difference between real male and female slopes against the null distribution, we can determine whether the difference between male and female slopes is significant by calculating the proportion of sampled permutations whose slope difference was greater or equal to the observed gender slope difference.

0.56;  $M_{12\text{faces}} = 24.63$ ,  $SE = 0.58$ ). Consistent with previous research (Haberman et al., 2009), the interaction between display face number and face orientation was significant [ $F_{(3, 168)} = 4.24$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.07$ ]. The number of faces in the display had a significant impact on participants' estimates in the upright face condition [ $F_{(3, 171)} = 14.18$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.20$ ], but had no effect when the faces were inverted [ $F_{(3, 171)} = 1.73$ ,  $p = 0.163$ ,  $\eta_p^2 = 0.03$ ].

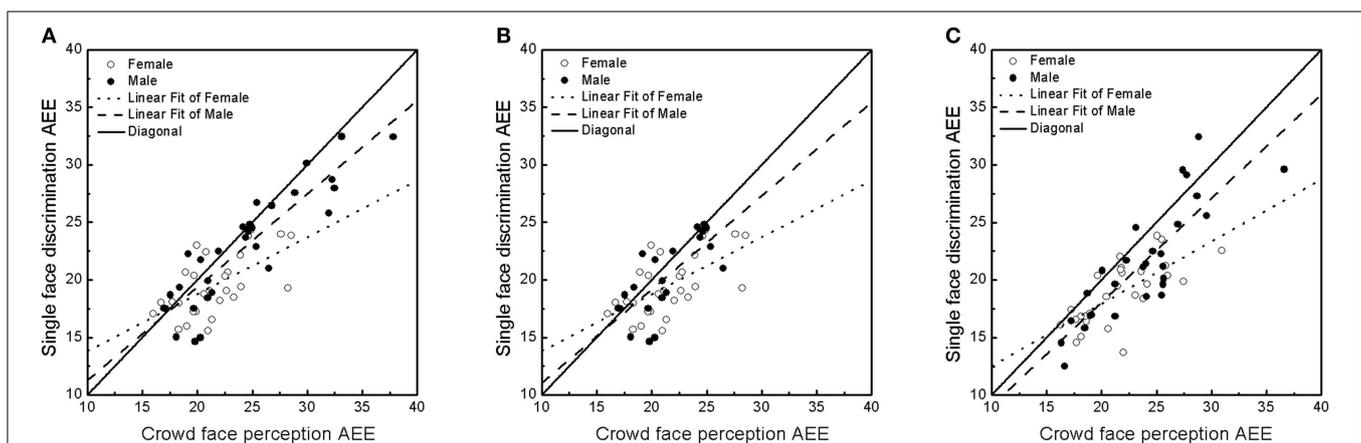
Consistent with the results of Experiment 1, we found a significant main effect of subject gender [ $F_{(1, 56)} = 4.53$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.08$ ]. Female participants ( $M = 23.29$ ,  $SE = 0.72$ ) were more accurate at estimating the mean face compared to male subjects ( $M = 25.41$ ,  $SE = 0.69$ ), and there was no significant interaction between subject gender and the number of display faces [ $F_{(3, 168)} = 0.98$ ,  $p = 0.40 > 0.05$ ]. Although the interaction between face orientation and participants' gender did not reach significance [ $F_{(1, 56)} = 2.66$ ,  $p = 0.109$ ,  $\eta_p^2 = 0.05$ ], a simple effect analysis revealed that females outperform males only when the display faces were upright [ $F_{(1, 56)} = 6.03$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.10$ ], while males' and females' performance did not differ when the display faces were inverted [ $F_{(1, 56)} = 2.10$ ,  $p = 0.152$ ,  $\eta_p^2 = 0.04$ ]. Again, participants' gender did not significantly interact with the stimulus gender [ $F_{(1, 56)} = 0.12$ ,  $p = 0.732$ ,  $\eta_p^2 = 0.002$ ].

In this experiment, in contrast to the single face condition, the other three conditions required participants to ensemble code the crowd in order to provide a more precise average estimation. In order to test the relationship between single face detection and ensemble coding of crowds, we collapsed the 4, 8, and 12 upright face conditions together and performed the same linear regression described in Experiment 1 (Figure 3A) using each subject's AEE for the single upright face condition as Y and AEE for the multi upright face condition (collapsed across 4, 8, and 12 upright faces conditions) as X. The results again showed that males had a steeper slope compared to females ( $b_{\text{male}} =$

0.811;  $b_{\text{female}} = 0.495$ ), and a permutation test showed that this difference was significant ( $p < 0.001$ ). This indicates that, despite their relatively poor sensitivity to single faces, male subjects' sensitivity to crowd identity was relatively high.

In a complementary analysis, we examined the proportion of subjects that had a smaller AEE in the crowd perception condition compared to their performance in the single face discrimination condition. More precision in heterogeneous crowd identity perception than single face identity perception indicates that more than one face has been integrated into analysis when perceiving the crowd (the only way to perceive heterogeneous crowds more precisely than a single face is by averaging, which cancels out noise). We found that 37% of male subjects had a smaller estimation error in the crowd perception condition compared to their performance in the single face discrimination condition, whereas only 25% of females had a smaller estimation error in the crowd perception condition compared to their performance in the single face discrimination condition (male vs. female difference:  $p = 0.096$ , permutation test). A relatively smaller (better) estimation error in the crowd vs. single face condition suggests that subject performance in the single face condition is not the absolute limit of subject performance in the crowd condition. That is, some subjects perceived crowds *more precisely* than they perceived single faces, and this held more widely across the male participants than across the female participants. Figure 3A depicts the larger proportion of males that performed adequately during ensemble coding tasks, despite relatively poor single face perception performance.

In order to test whether the observed gender differences (the slope difference in Figure 3A) was caused by a difference in the two genders' abilities in the single face condition, we re-conducted the analysis using only data from males who showed similar single face perception performance to females. That is, we used only male subjects whose AEE was in the range of



**FIGURE 3 | Results of Experiments 2a and 2b. (A)** Relationship between single face discrimination and perception of crowds of redundant upright faces (crowds contained repeated pictures of four identities on each trial) for males and females in Experiment 2a. Crowd face perception AEE was the average AEE for the 4, 8, and 12 face conditions. **(B)** Relationship between single face discrimination and redundant upright crowd face perception for all females and the subset of males who had an equivalent single face perception as females in Experiment 2a. **(C)** Relationship between single face discrimination and non-redundant upright crowd face perception for males and females in Experiment 2b.

the females' AEE for the upright single face condition, ensuring equivalent single face estimation performance between female and selected male participants. The analysis confirmed that the slope of male subjects was steeper than that of female subjects (**Figure 3B**;  $b_{\text{male}} = 0.813$ ;  $b_{\text{female}} = 0.495$ ). A permutation test showed that those males who performed as well as females in the single face perception condition still tended to have a larger slope than female subjects ( $p = 0.06$ ), suggesting that the observed gender difference was not driven by a subset of males who performed poorly.

Experiment 2a provides several important insights. First of all, we replicated our main finding in Experiment 1 that females perform better than males in both the single face discrimination task and the ensemble coding task. Additionally, the proportion test and slope test above both provide evidence showing that although males' *absolute* sensitivity is relatively lower than females' sensitivity (males' overall estimation error was larger than females'), males still achieve adequate ensemble coding performance. Essentially, male subjects' performance in the crowd condition was not as poor as would be expected from their single face discrimination. Even when controlling for the male subjects' abilities in the single face discrimination condition, we still observed a steeper slope in male subjects' performance. These results suggest that while females robustly perceive crowds with higher precision, males might have their own perceptual advantage in the ensemble coding process.

## Experiment 2b

In Experiment 2a, the different set sizes shared the same variance. Thus, the information conveyed by the 4, 8, and 12 set size displays was identical—except that the 8 and 12 set size displays contained redundant faces. Previous research indicated that redundant faces within crowds improved participants' ensemble coding performance (Haberman and Whitney, 2010). In order to minimize this redundancy effect, we repeated the experiment using crowds of faces with randomly chosen variance (see **Supplementary Figure 1**). Thus, Experiment 2b aimed to replicate the results of Experiment 2a, while allowing participants to view faces with a fluctuating variance. In other words, in some trials participants viewed crowds with a variance of 0 while in other trials participants viewed crowds with a variance of 12.73, and many values of variance in between. The variance changed according to the randomly chosen faces<sup>2</sup> (see **Supplementary Figure 1**).

## Participants

Fifty eight undergraduate students at Tsinghua University in China participated in exchange for an optional course credit. Among all the subjects, two subjects did not complete all conditions and their results were excluded. The remaining 56 were comprised of 28 females and 28 males ( $M = 20.59$  years,  $SD = 1.47$ ).

<sup>2</sup>The randomly chosen faces were still drawn from within a designated range from the mean of the trial.

## Materials and Procedure

The stimuli and settings were identical to the stimuli used in Experiment 2a, with one significant difference. In Experiment 2b, the variance of the display was not restricted. For each trial, the program first chose 12 faces. The 12 faces were comprised of four faces that were either three or nine units away from the randomly selected mean face in both directions, repeated three times (similar to the display depicted in **Figure 1A**). Next, we randomly presented 1, 2, 4, or 8 faces out of the chosen 12-face set. Thus, the variance in the displayed faces was not consistent throughout the experiment as it was in Experiment 2a. In Experiment 2b, the displayed face variance fluctuated according to the faces selected for each trial. This minimized the effects of redundancy. All other procedures were the same as those in Experiment 2a.

## Results and Discussion

Similar to Experiment 2a, we conducted a 2 (subject gender: female vs. male)  $\times$  2 (picture orientation: upright vs. inverted)  $\times$  4 (display condition: 1, 2, 4, or 8 visible faces)  $\times$  2 (picture gender: female vs. male) ANOVA. Only subject gender was a between-subject variable. Similar to previous research and the results of Experiment 1, the ANOVA revealed a significant main effect of picture orientation. Participants were significantly better in estimating the mean face for upright faces compared to inverted faces [ $F_{(1, 54)} = 81.86$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.60$ ]. The number of display face(s) also significantly influenced ensemble coding performance [ $F_{(3, 162)} = 25.23$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.32$ ]. Consistent with the results of Experiment 2a, Bonferroni *post-hoc* tests revealed that participants were more accurate in single face discrimination compared to ensemble coding a group of faces, regardless of the group size ( $M_{1\text{face}} = 22.04$ ,  $SE = 0.53$ ;  $M_{4\text{faces}} = 24.09$ ,  $SE = 0.51$ ;  $M_{8\text{faces}} = 24.13$ ,  $SE = 0.53$ ;  $M_{12\text{faces}} = 24.45$ ,  $SE = 0.52$ ). In addition, we again found a significant effect of subject gender [ $F_{(1, 54)} = 5.32$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.075$ ], which is consistent with our previous results. Female participants ( $M_{\text{female}} = 22.55$ ) were more accurate at estimating the mean face compared to male subjects ( $M_{\text{male}} = 24.80$ ), and there was no significant interaction between subject gender and display face number [ $F_{(3, 162)} = 0.37$ ,  $p = 0.77 > 0.05$ ]. Consistent with results of Experiment 1 and 2a, participants' gender did not interact with the stimulus gender [ $F_{(1, 54)} = 2.68$ ,  $p = 0.107$ ,  $\eta_p^2 = 0.05$ ].

We collapsed the AEE across the 2, 4, and 8 upright face conditions and performed the same slope fitting test again, as in Experiments 1 and 2a. The results again showed that, compared to females, males had a steeper slope (**Figure 3C**;  $b_{\text{male}} = 0.903$ ;  $b_{\text{female}} = 0.541$ ). A permutation test showed that the males' slope was significantly steeper than females' slope ( $p < 0.05$ ). The slope of the subgroup of males who had equivalent single face perception performance to females ( $b_{\text{male}} = 0.675$ ) was compared to the slope of the full set of females ( $b_{\text{female}} = 0.541$ ). Although a permutation test did not find a significant difference between the two slopes ( $p = 0.11$ ), the pattern suggests a trend that males who performed as well as females in the single face perception condition tended to have steeper slopes. Again, a higher proportion of males tended to perform better in the crowd

than in the single face condition (Male = 21.4%; Female = 10.7%,  $p = 0.063$ ).

In sum, Experiment 2b replicated the result of Experiment 2a, even though the faces in each set were not necessarily redundant.

## Simulation Results

Many ensemble coding experiments simulate the number of faces/objects integrated into the ensemble percept (Haberman and Whitney, 2010; Leib et al., 2012; Im and Halberda, 2013). This follow-up analysis allows researchers to investigate specifics of the ensemble coding mechanism when large numbers of items are integrated into the ensemble percept. This is important because while statistical averaging can technically include as little as two faces from the crowd, in natural settings integrating more faces would yield the greatest information about the crowd. Thus, we investigated whether previously observed patterns between males and females remain robust under such simulations. Specifically, how does single face sensitivity correlate with crowd sensitivity when larger numbers of faces are integrated from the display? The simulation contained the following four steps: (1) We simulated the value (relative to the mean) of the faces displayed—in this case (−9 to +9) (see **Figure 4A**). (2) Next, we simulated the number of faces ( $N$ ) sampled from the display. Each face was drawn from a Gaussian distribution centered at the display face value. (3) Next, noise was added to the chosen face values based on the range of the participants' empirical data in the single face condition (**Figure 4B** shows a description of this step). (4)  $N$  face values were averaged in a noise-free integration step. (5) Error was determined by subtracting this average from the true mean of the crowd<sup>3</sup>. We simulated 1000 bootstrapped trials and

<sup>3</sup>This is not a realistic estimate but is a conservative estimate. In actuality, participants are probably much noisier (not centered on the perfect average). If this is the case, the results of the simulation would yield a much higher error, suggesting that participants integrated even more faces.

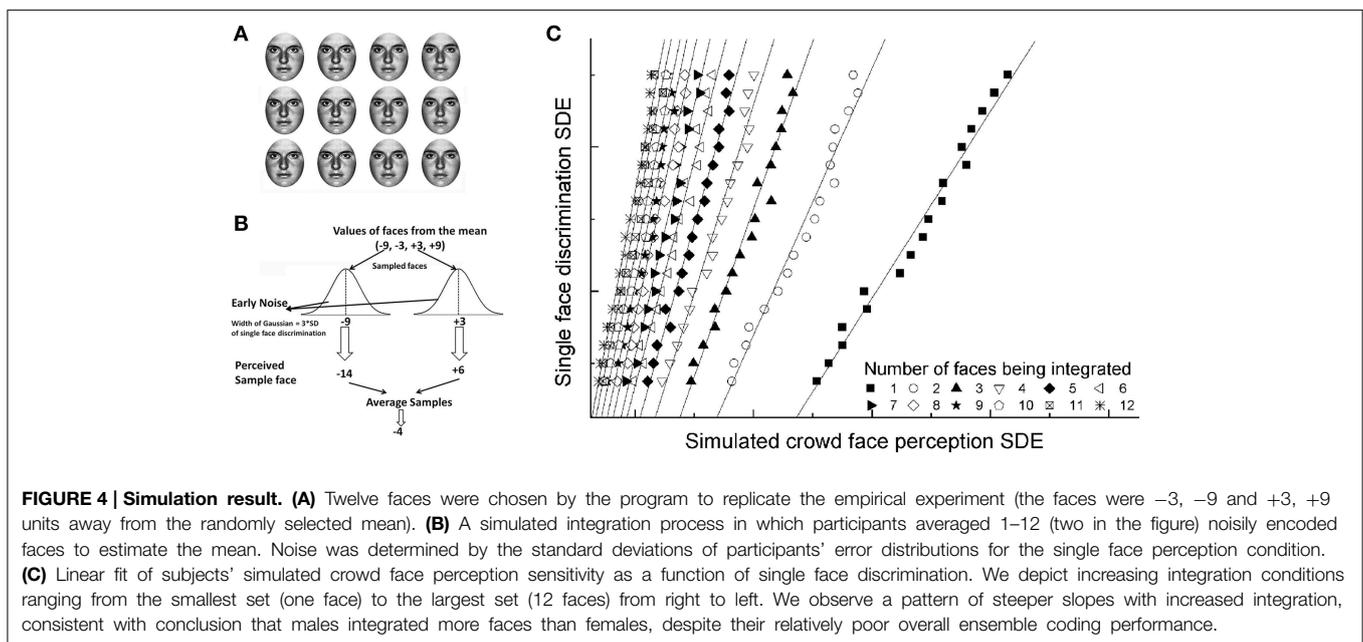
the standard deviation of the error distribution was calculated and regarded as the simulated group perception performance. For each simulation, we iteratively chose a higher  $N$  to simulate performance when participants systematically integrated more and more faces in the display.

We compared the participants' empirical single face discrimination score to the simulated ensemble coding performance using a linear regression test (**Figure 4C**). The analysis indicates that as more faces were integrated into the mean judgment, the slope became steeper. This pattern suggests that it is not only the discrimination of single faces that contributes to performance, but also the integration process that allows participants to successfully evaluate the mean identity of the crowd. Importantly, we observed a significant difference in males' and females' slope values. Males exhibited a steeper slope compared to females as the number of faces integrated increased. The gender slope difference combined with the simulation result indicate that although males' performance was lower overall, they may be integrating a larger number of faces into their ensemble percepts.

The simulation is an exploratory analysis conducted *post-hoc*. Nonetheless, it offers some potential insight into the ensemble coding process of both males and females. Females are superior at ensemble coding, but males are theoretically still able to integrate a large number of faces into their ensemble percept, despite their poor single face performance. This may even suggest that ensemble coding serves as a compensatory mechanism for males.

## General Discussion

The current study is the first to examine how males and females might differ in their initial perceptions of crowds. While previous social psychology studies have explored how both males and females formulate long-term group impressions (e.g., Hoxter and Lester, 1994; Ekehammar et al., 2003), our study examined



whether there are actual *perceptual* differences between males and females when they evaluate a crowd of faces at first glance. Our data show that compared to males, females have a perceptual advantage in determining average crowd characteristics. It is important to note that the main effect of gender was associated with a small effect size. Nevertheless, since we have replicated the main effect across all the three studies, it is nonetheless a robust finding.

This perceptual advantage in crowd discrimination is likely a result of females' relatively superior individual face identification abilities, which has been discovered by previous research (Rehman and Herlitz, 2006; McBain et al., 2009; Heisz et al., 2013; Sommer et al., 2013). In a series of experiments, McBain et al. (2009) manipulated both the orientation of the faces (upright vs. inverted) and the information available in the faces (low vs. high noise). They found that females consistently performed better in upright, high-noise face processing tasks by relying on holistic processing skills. Our results extend these findings. Ensemble coding by definition requires participants to gather information about individual faces first, and our study suggests that females' single face advantages may contribute to their ensemble coding performance. Furthermore, females' superior episodic memory and working memory (Herlitz et al., 1997, 1999; Harness et al., 2008) may also contribute to females' better crowd perception. While the relationship between memory and ensemble coding has not been fully explored, previous authors have suggested that perception inherently involves memory (e.g., Luck and Vogel, 1997). Brady and Alvarez demonstrated that ensemble coding influences working memory. In a working memory task, participants reported the size and color of individual items in a display, and they found participants' reports were biased toward the mean of the set of items (Brady and Alvarez, 2011). It may well be that the relationship is reciprocal. Therefore, it is possible that females' superior memory enhances the accuracy of their ensemble percept<sup>4</sup>.

While females demonstrably performed better in perceiving a single face and crowds of faces, our data consistently showed that males also exhibited unique strengths: although males' crowd perception was relatively poor compared to female, males' crowd perception was better than would be predicted from their single face perception. Compared to females, a greater proportion of males performed better in the crowd than in the single face perception condition. How is this paradoxical result possible? Considering that crowds are perceived as a summary statistic or ensemble (Haberma and Whitney, 2007), the pattern may not be surprising. For instance, prosopagnosics, individuals who have difficulty discriminating individual faces, still achieved an accurate ensemble percept (Leib et al., 2012). Prosopagnosics' precision in ensemble coding may be a result of noise reduction. In any mathematical averaging process, noise is reduced with a greater number of samples, and greater precision is achieved (Alvarez, 2011; Robitaille and Harris, 2011). Our

male subjects' individual face representations were noisy (at least compared to females); however, averaging larger sample sizes (i.e., crowds of faces) may increase precision via noise cancellation. Thus, by integrating more faces into the ensemble percept, males may compensate for relatively poor sensitivity to single faces and improve their perception of the crowd average.

From a typical social psychological perspective, our findings are reasonable. Darwinian consensus is that the acquisition of any mental power or capacity is the result of evolution. Evolutionary pressure may have contributed to the identity sensitivity for females and crowd familiarity for males. For example, in most human societies, males have historically participated in larger social networks compared to females (for review, see Belle, 1989; Benenson, 1990). In contrast, females have been reported to have more intimate confidants and relationships than males (e.g., Williams, 1959; Powers and Bultena, 1976). It is possible that historically differing social networks may impact males' and females' ensemble coding habits, allowing females to precisely scrutinize single faces and allowing males to integrate a greater number of faces, despite their relatively poor single face and group perception as a whole.

This study is a first step toward examining how social and cognitive factors shape people's group perception, and we were primarily interested in the comparison between male and female participants. However, another potentially interesting avenue of research is exploring the own-gender effect (Herlitz and Lovén, 2013). In our particular experiment, we did not observe the own-gender effect. However, this may be because we tested Asian participants with Caucasian stimuli. Future studies should also investigate how participant ethnicity interacts with stimulus ethnicity.

In the current paper, we emphasize that ensemble coding is not solely a cognitive process, but is also influenced by high-level social factors, such as gender constructs. It will be important for future work to test if culture, a well-discussed social factor, shapes participants' immediate group perception. Previous research has documented culture differences in cognitive processes (Nisbett et al., 2001; Norenzayan et al., 2007; Masuda et al., 2008, 2012; Miyamoto, 2013), in which researchers found that while Asians focus on relationships between a focal object and its context, Westerners focus on an object independently from its context (Masuda and Nisbett, 2001; Kitayama et al., 2003). This perceptual habit difference was also found to lead to face perception differences across the two cultures (Miyamoto et al., 2011). For instance, Masuda et al. (2008) asked participants to distinguish the emotion of faces, and found that Caucasians mainly focus on the target face itself, while Japanese participants took contextual information into account. These findings raise interesting questions for future ensemble coding research. Will Asians be more influenced by the context of the crowd compared to Caucasians? Will Asians, while judging a subset of the crowd, be more influenced by faces outside the subset? Examining the interaction between culture and ensemble coding may yield interesting insights into differences between Westerners' and Easterners' perceptions of crowds.

<sup>4</sup>In our particular task, working memory may also play a significant role during the test phase. During this phase, participants must keep in mind the average identity while scrolling through the test faces (147 in all).

Across all three studies, we tested group perception using morphed faces because these faces have been well controlled for subtle variations and thus allow us to concretely measure error in participants' average face judgment (Haberman and Whitney, 2007, 2009; Haberman et al., 2009). However, the complex nature of group perception does place limits on the extent to which real life group perception can be examined experimentally and prospectively in a laboratory setting. Using current research as a basis, future research needs to explore using different stimuli (e.g., crowds of emotional faces, crowds with full-body pictures, etc.) and different groups (e.g., people with different ethnicities, people with different social status, etc.). Furthermore, while the current study limits its focus to examining the differences in ensemble coding performance between males and females, it is intriguing to consider whether perceptual differences in crowd evaluation might affect males' and females' impression formation over time and across various social contexts. A number of social psychology studies report that crowd observations, as well as the subsequent schemas that develop, guide future social interactions (e.g., Devine, 1989; Spencer et al., 1999). We suggest that the relationship between perceptual analysis of the crowd and social interaction may be reciprocal. Variations in visual processing influence social evaluations of groups, and, conversely, diversities in social background influence perception. Considering the possible interactions between perceptual and social processes within crowds and their members, an interesting avenue of future investigation will be to evaluate whether an individual's ensemble coding performance influences or is related to social phenomena such as stereotyping, impression formation,

category accentuation, prejudice, and intergroup conflicts and cooperation. This future work will illuminate how our perception of the world mediates our social interactions, and how our social identities mediate the processes of social perception.

## Acknowledgments

This research was supported by "The Berkeley Program for Advanced Study in Culture and Psychology" to KP, "The Sheldon Zedeck Program for Culture, Behavior and Management Study" to YB, NSF 20123108 and NIH R01EY018216 to DW. We thank Zaijia Liu and Fei Wang for their assistance with data collection and Ge Fang for his help with programming.

## Supplementary Material

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01300>

**Supplementary Figure 1 | Display examples for Experiments 2a and 2b. (A)** Each display contained either a single face or 4 identities, each repeating 1, 2, or 3 times. The four faces in the 4, 8, and 12 face conditions were 3 and 9 steps away from the mean face in either direction, with a consistent variance. **(B)** Each display contained 1, 2, 4, or 8 faces randomly selected from 12 faces. Thus, the variance of the displayed faces was not restricted. Note: all the faces were randomly chosen and randomly displayed on the screen.

**Supplementary Video 1 | Two trials for Experiment 1.** Participants were presented with 18 faces on the screen before one test face was shown on the screen. By scrolling the mouse, participants could search through the whole face wheel to find the estimated average face.

## References

- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends Cogn. Sci.* 15, 122–131. doi: 10.1016/j.tics.2011.01.003
- Ariely, D. (2001). Seeing sets: representation by statistical properties. *Psychol. Sci.* 12, 157–162. doi: 10.1111/1467-9280.00327
- Asch, S. E. (1946). Forming impressions of personality. *J. Abnorm. Soc. Psychol.* 41, 258–290. doi: 10.1037/h0055756
- Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychol. Monogr. Gen. Appl.* 70, 1–70. doi: 10.1037/h0093718
- Belle, D. (1989). "Gender differences in children's social networks and supports" in *Children's Social Networks and Supports*, ed D. Belle (New York, NY: Wiley), 173–188.
- Benenson, J. F. (1990). Gender differences in social networks. *J. Early Adolesc.* 10, 472–495. doi: 10.1177/0272431690104004
- Bohan, J. S. (1993). Regarding gender: essentialism, constructionism, and feminist psychology. *Psychol. Women Q.* 17, 5–21. doi: 10.1111/j.1471-6402.1993.tb00673.x
- Brady, T. F., and Alvarez, G. A. (2011). Hierarchical encoding in visual working memory ensemble statistics bias memory for individual items. *Psychol. Sci.* 22, 384–392. doi: 10.1177/0956797610397956
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behav. Sci.* 3, 14–25.
- Chau, W., McIntosh, A. R., Robinson, S. E., Schulz, M., and Pantev, C. (2004). Improving permutation test power for group analysis of spatially filtered MEG data. *Neuroimage* 23, 983–996. doi: 10.1016/j.neuroimage.2004.07.007
- Chen, J. M., and Hamilton, D. L. (2012). Natural ambiguities: racial categorization of multiracial individuals. *J. Exp. Soc. Psychol.* 48, 152–164. doi: 10.1016/j.jesp.2011.10.005
- Chen, J. M., Moons, W. G., Gaither, S. E., Hamilton, D. L., and Sherman, J. W. (2014). Motivation to control prejudice predicts categorization of multiracials. *Pers. Soc. Psychol. Bull.* 40, 590–603. doi: 10.1177/0146167213520457
- Chong, S. C., Joo, S. J., Emmanouil, T. A., and Treisman, A. (2008). Statistical processing: not so implausible after all. *Percept. Psychophys.* 70, 1327–1334. doi: 10.3758/PP.70.7.1327
- Chong, S. C., and Treisman, A. (2003). Representation of statistical properties. *Vision Res.* 43, 393–404. doi: 10.1016/S0042-6989(02)00596-5
- Chong, S. C., and Treisman, A. (2005). Statistical processing: computing the average size in perceptual groups. *Vision Res.* 45, 891–900. doi: 10.1016/j.visres.2004.10.004
- Cialdini, R. B., and Trost, M. R. (1998). "Social influence: social norms, conformity, and compliance," in *The Handbook of Social Psychology, 4th Edn.*, Vol. 2D, eds T. Gilbert, S. T. Fiske, and G. Lindzey (Boston, MA: McGraw-Hill), 151–192.
- Collaer, M. L., and Hines, M. (1995). Human behavioral sex differences: a role for gonadal hormones during early development? *Psychol. Bull.* 118, 55–107. doi: 10.1037/0033-2909.118.1.55
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals, 3rd Edn.* New York, NY: Oxford University Press. doi: 10.1037/10001-000
- de Fockert, J., and Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. *Q. J. Exp. Psychol.* 62, 1716–1722. doi: 10.1080/17470210902811249
- Devine, P. G. (1989). Stereotypes and prejudice: their automatic and controlled components. *J. Pers. Soc. Psychol.* 56, 5–18. doi: 10.1037/0022-3514.56.1.5

- Eagly, A. H., and Wood, W. (1999). The origins of sex differences in human behavior: evolved dispositions versus social roles. *Am. Psychol.* 54, 408–423. doi: 10.1037/0003-066X.54.6.408
- Ekehammar, B., Akrami, N., and Araya, T. (2003). Gender differences in implicit prejudice. *Pers. Individ. Dif.* 34, 1509–1523. doi: 10.1016/S0191-8869(02)00132-0
- Ekman, P. (1973). “Cross-cultural studies of facial expression,” in *Darwin and Facial Expression: A Century of Research in Review*, ed P. Ekman (New York, NY: Academic Press), 169–222.
- Ekman, P., and Friesen, W. V. (1976). Measuring facial movement. *Environ. Psychol. Nonverbal Behav.* 1, 56–75. doi: 10.1007/BF01115465
- Fiske, S. T. (2002). What we know now about bias and intergroup conflict, the problem of the century. *Curr. Dir. Psychol. Sci.* 11, 123–128. doi: 10.1111/1467-8721.00183
- Fiske, S. T. (2010). Venus and mars or down to earth: stereotypes and realities of gender differences. *Perspect. Psychol. Sci.* 5, 688–692. doi: 10.1177/1745691610388768
- Fiske, S. T., and Von Hendy, H. M. (1992). Personality feedback and situational norms can control stereotyping processes. *J. Pers. Soc. Psychol.* 62, 577–596. doi: 10.1037/0022-3514.62.4.577
- Good, P. (2000). *Permutation Tests, 2nd Edn.* New York, NY: Springer. doi: 10.1007/978-1-4757-3235-1
- Guinote, A., Judd, C. M., and Brauer, M. (2002). Effects of power on perceived and objective group variability: evidence that more powerful groups are more variable. *J. Pers. Soc. Psychol.* 82, 708–721. doi: 10.1037/0022-3514.82.5.708
- Haberman, J., Harp, T., and Whitney, D. (2009). Averaging facial expression over time. *J. Vis.* 9, 1–13. doi: 10.1167/9.11.1
- Haberman, J., and Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Curr. Biol.* 17, R751–R753. doi: 10.1016/j.cub.2007.06.039
- Haberman, J., and Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 718–734. doi: 10.1037/a0013899
- Haberman, J., and Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Atten. Percept. Psychophys.* 72, 1825–1838. doi: 10.3758/APP.72.7.1825
- Haberman, J., and Whitney, D. (2011). Efficient summary statistical representation when change localization fails. *Psychon. Bull. Rev.* 18, 855–859. doi: 10.3758/s13423-011-0125-6
- Hamilton, D. L. (2007). Understanding the complexities of group perception: broadening the domain. *Eur. J. Soc. Psychol.* 37, 1077–1101. doi: 10.1002/ejsp.436
- Hamilton, D. L., and Gifford, R. K. (1976). Illusory correlation in interpersonal perception: a cognitive basis of stereotypic judgments. *J. Exp. Soc. Psychol.* 12, 392–407. doi: 10.1016/S0022-1031(76)80006-6
- Hamilton, D. L., and Sherman, J. W. (1994). “Stereotypes,” in *Handbook of Social Cognition, 2nd Edn.*, Vol. 2, eds R. S. Wyer Jr. and T. K. Srull (San Diego, CA: Academic Press), 1–68.
- Harness, A., Jacot, L., Scherf, S., White, A., and Warnick, J. E. (2008). Sex differences in working memory. *Psychol. Rep.* 103, 214–218. doi: 10.2466/pr0.103.1.214-218
- Heath, A. C., Madden, P. A. F., Buchholz, K. K., Dinwiddie, S. H., Slutske, W. S., Bierut, L. J., et al. (1999). Genetic differences in alcohol sensitivity and the inheritance of alcoholism risk. *Psychol. Med.* 29, 1069–1081. doi: 10.1017/S0033291799008909
- Heisz, J. J., Pottruff, M. M., and Shore, D. I. (2013). Females scan more than males a potential mechanism for sex differences in recognition memory. *Psychol. Sci.* 24, 1157–1163. doi: 10.1177/0956797612468281
- Helbing, D., Farkas, I., and Vicsek, T. (2000). Simulating dynamical features of escape panic. *Nature* 407, 487–490. doi: 10.1038/35035023
- Herlitz, A., Airaksinen, E., and Nordström, E. (1999). Sex differences in episodic memory: the impact of verbal and visuospatial ability. *Neuropsychology* 13, 590–597. doi: 10.1037/0894-4105.13.4.590
- Herlitz, A., and Lovén, J. (2013). Sex differences and the own-gender bias in face recognition: a meta-analytic review. *Vis. Cogn.* 21, 1306–1336. doi: 10.1080/13506285.2013.823140
- Herlitz, A., Nilsson, L. G., and Bäckman, L. (1997). Gender differences in episodic memory. *Mem. Cognit.* 25, 801–811. doi: 10.3758/BF03211324
- Hines, M. (1982). Prenatal gonadal hormones and sex differences in human behavior. *Psychol. Bull.* 92, 56–80. doi: 10.1037/0033-2909.92.1.56
- Hoxter, A. L., and Lester, D. (1994). Gender differences in prejudice. *Percept. Mot. Skills* 79, 1666–1666. doi: 10.2466/pms.1994.79.3f.1666
- Hyde, J. S., and Linn, M. C. (1988). Gender differences in verbal ability: a meta-analysis. *Psychol. Bull.* 104, 53–69. doi: 10.1037/0033-2909.104.1.53
- Im, H. Y., and Halberda, J. (2013). The effects of sampling and internal noise on the representation of ensemble average size. *Atten. Percept. Psychophys.* 75, 278–286. doi: 10.3758/s13414-012-0399-4
- Jans, L., Postmes, T., and Van der Zee, K. I. (2011). The induction of shared identity: the positive role of individual distinctiveness for groups. *Pers. Soc. Psychol. Bull.* 37, 1130–1141. doi: 10.1177/0146167211407342
- Kitayama, S., Duffy, S., Kawamura, T., and Larsen, J. T. (2003). Perceiving an object and its context in different cultures: a cultural look at new look. *Psychol. Sci.* 14, 201–206. doi: 10.1111/1467-9280.02432
- Leib, A. Y., Puri, A. M., Fischer, J., Bentin, S., Whitney, D., and Robertson, L. (2012). Crowd perception in prosopagnosia. *Neuropsychologia* 50, 1698–1707. doi: 10.1016/j.neuropsychologia.2012.03.026
- Lickel, B., Hamilton, D. L., Wierzchowska, G., Lewis, A., Sherman, S. J., and Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *J. Pers. Soc. Psychol.* 78, 223–246. doi: 10.1037/0022-3514.78.2.223
- Linville, P. W., Fischer, G. W., and Salovey, P. (1989). Perceived distributions of the characteristics of in-group and out-group members: empirical evidence and a computer simulation. *J. Pers. Soc. Psychol.* 57, 165–188. doi: 10.1037/0022-3514.57.2.165
- Louie, E. G., Bressler, D. W., and Whitney, D. (2007). Holistic crowding: selective interference between configural representations of faces in crowded scenes. *J. Vis.* 7, 1–11. doi: 10.1167/7.2.24
- Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281. doi: 10.1038/36846
- Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., and Van de Veerdonk, E. (2008). Placing the face in context: cultural differences in the perception of facial emotion. *J. Pers. Soc. Psychol.* 94, 365–381. doi: 10.1037/0022-3514.94.3.365
- Masuda, T., and Nisbett, R. E. (2001). Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *J. Pers. Soc. Psychol.* 81, 922–934. doi: 10.1037/0022-3514.81.5.922
- Masuda, T., Wang, H., Ishii, K., and Ito, K. (2012). Do surrounding figures’ emotions affect the judgment of target figure’s emotion? Comparing the patterns of attention between European-Canadians, Asian-Canadians, Asian international students, and Japanese using eye-trackers. *Front. Integr. Neurosci.* 6:72. doi: 10.3389/fnint.2012.00072
- McBain, R., Norton, D., and Chen, Y. (2009). Females excel at basic face perception. *Acta Psychol.* 130, 168–173. doi: 10.1016/j.actpsy.2008.12.005
- Miller, D. I., and Halpern, D. F. (2014). The new science of cognitive sex differences. *Trends Cogn. Sci.* 18, 37–45. doi: 10.1016/j.tics.2013.10.011
- Miyamoto, Y. (2013). Culture and analytic versus holistic cognition: toward multilevel analyses of cultural influences. *Adv. Exp. Soc. Psychol.* 47, 131–188. doi: 10.1016/B978-0-12-407236-7.00003-6
- Miyamoto, Y., Yoshikawa, S., and Kitayama, S. (2011). Feature and configuration in face processing: Japanese are more configural than Americans. *Cogn. Sci.* 35, 563–574. doi: 10.1111/j.1551-6709.2010.01163.x
- Nisbett, R. E., Peng, K., Choi, I., and Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychol. Rev.* 108, 291–310. doi: 10.1037/0033-295X.108.2.291
- Norenzayan, A., Choi, I., and Peng, K. (2007). “Perception and cognition,” in *Handbook of Cultural Psychology*, eds S. Kitayama and D. Cohen (New York, NY: Guilford), 569–594.
- Ostrom, T. M., and Sedikides, C. (1992). Out-group homogeneity effects in natural and minimal groups. *Psychol. Bull.* 112, 536–536. doi: 10.1037/0033-2909.112.3.536
- Park, B., and Hastie, R. (1987). Perception of variability in category development: instance- versus abstraction-based stereotypes. *J. Pers. Soc. Psychol.* 53, 621–635. doi: 10.1037/0022-3514.53.4.621
- Park, B., and Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: memory for the subordinate attributes of in-group and

- out-group members. *J. Pers. Soc. Psychol.* 42, 1051–1068. doi: 10.1037/0022-3514.42.6.1051
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* 4, 739–744. doi: 10.1038/89532
- Powers, E. A., and Bultena, G. L. (1976). Sex differences in intimate friendships of old age. *J. Marriage Fam.* 38, 739–747. doi: 10.2307/350693
- Rehman, J., and Herlitz, A. (2006). Higher face recognition ability in girls: magnified by own-sex and own-ethnicity bias. *Memory* 14, 289–296. doi: 10.1080/09658210500233581
- Robitaille, N., and Harris, I. M. (2011). When more is less: extraction of summary statistics benefits from larger sets. *J. Vis.* 11, 1–8. doi: 10.1167/11.12.18
- Rothbart, M., and Park, B. (2004). “The mental representation of social categories: category boundaries, entitativity, and stereotype change,” in *The Psychology of Group Perception: Perceived Variability, Entitativity and Essentialism*, eds V. Yzerbyt, C. M. Judd, and O. Corneille (Philadelphia, PA: Psychology Press), 79–100.
- Schaller, M., and O’Brien, M. (1992). “Intuitive analysis of covariance” and group stereotype formation. *Pers. Soc. Psychol. Bull.* 18, 776–785. doi: 10.1177/0146167292186014
- Sommer, W., Hildebrandt, A., Kunina-Habenicht, O., Schacht, A., and Wilhelm, O. (2013). Sex differences in face cognition. *Acta Psychol.* 142, 62–73. doi: 10.1016/j.actpsy.2012.11.001
- Spears, R., Jetten, J., and Doosje, B. (2001). “The (il)legitimacy of ingroup bias: from social reality to social resistance,” in *The Psychology of Legitimacy: Emerging Perspectives on Ideology, Justice, and Intergroup Relations*, eds J. Jost and B. Major (New York, NY: Cambridge University Press), 332–362.
- Spencer, S. J., Steele, C. M., and Quinn, D. M. (1999). Stereotype threat and women’s math performance. *J. Exp. Soc. Psychol.* 35, 4–28. doi: 10.1006/jesp.1998.1373
- Spencer-Rodgers, J., Williams, M. J., Hamilton, D. L., Peng, K., and Wang, L. (2007). Culture and group perception: dispositional and stereotypic inferences about novel and national groups. *J. Pers. Soc. Psychol.* 93, 525–543. doi: 10.1037/0022-3514.93.4.525
- Sweeny, T. D., Haroz, S., and Whitney, D. (2013). Perceiving group behavior: sensitive ensemble coding mechanisms for biological motion of human crowds. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 329–337. doi: 10.1037/a0028712
- Todorov, A. (2012). “The social perception of faces,” in *The SAGE Handbook of Social Cognition*, eds S. T. Fiske and C. N. Macrae (Thousand Oaks, CA: SAGE Publications), 96–112. doi: 10.4135/9781446247631.n6
- Turner, J. C. (1985). “Social categorization and the self-concept: a social cognitive theory of group behaviour,” in *Advances in Group Processes: Theory and Research*, Vol. 2, ed E. J. Lawler (Greenwich, CT: JAI), 77–122.
- Whitney, D., Haberman, J., and Sweeny, T. D. (2014). “From textures to crowds: multiple levels of summary statistical perception,” in *The New Visual Neurosciences*, eds J. S. Werner and L. M. Chalupa (Cambridge, MA: MIT Press), 695–710.
- Whitney, D., and Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn. Sci.* 15, 160–168. doi: 10.1016/j.tics.2011.02.005
- Williams, R. M. (1959). Friendship and social values in a suburban community: an exploratory study. *Pac. Sociol. Rev.* 2, 3–10. doi: 10.2307/1388330
- Wood, W., and Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: implications for the origins of sex differences. *Psychol. Bull.* 128, 699–727. doi: 10.1037/0033-2909.128.5.699
- Yuki, M., and Yokota, K. (2009). The primal warrior: outgroup threat priming enhances intergroup discrimination in men but not women. *J. Exp. Soc. Psychol.* 45, 271–274. doi: 10.1016/j.jesp.2008.08.018

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Bai, Leib, Puri, Whitney and Peng. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.