



Published in final edited form as:

Nat Hum Behav. 2017 ; 1: 828–842. doi:10.1038/s41562-017-0225-z.

Differential hemispheric and visual stream contributions to ensemble coding of crowd emotion

Hee Yeon Im^{1,2}, Daniel N. Albohn³, Troy G. Steiner³, Cody A. Cushing², Reginald B. Adams Jr.³, and Kestutis Kveraga^{1,2}

¹Department of Radiology, Harvard Medical School

²Athinoula A. Martinos Center for Biomedical Imaging, Department Radiology, Massachusetts General Hospital

³Department of Psychology, The Pennsylvania State University

Abstract

In crowds, where scrutinizing individual facial expressions is inefficient, humans can make snap judgments about the prevailing mood by reading “crowd emotion”. We investigated how the brain accomplishes this feat in a set of behavioral and fMRI studies. Participants were asked to either avoid or approach one of two crowds of faces presented in the left and right visual hemifields. Perception of crowd emotion was improved when crowd stimuli contained goal-congruent cues and was highly lateralized to the right hemisphere. The dorsal visual stream was preferentially activated in crowd emotion processing, with activity in the intraparietal sulcus and superior frontal gyrus predicting perceptual accuracy for crowd emotion perception, whereas activity in the fusiform cortex in the ventral stream predicted better perception of individual facial expressions. Our findings thus reveal significant behavioral differences and differential involvement of the hemispheres and the major visual streams in reading crowd versus individual face expressions.

Keywords

crowd emotion; ensemble coding; face perception; facial expression; hemispheric lateralization

Correspondence to: Kestutis Kveraga, A.A. Martinos Center for Biomedical Imaging, Department of Radiology, Harvard Medical School, Massachusetts General Hospital, Charlestown, MA 02129, keastas@nmr.mgh.harvard.edu.

Author Contributions

H. Y. Im, R. B. Adams, and K. Kveraga developed the study concept, and all authors contributed to the study design. Testing and data collection were performed by H.Y. Im, C. A. Cushing, T. G. Steiner, and D. N. Albohn. H. Y. Im analyzed the data and all the authors wrote the manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of the article.

Data collection was conducted at the Pennsylvania State University. Informed written consent was obtained in all studies according to the procedures of the Institutional Review Board at the Pennsylvania State University. The participants received a course credit for their participation.

Introduction

We routinely encounter groups of people at work, school, or social gatherings. In real-life situations, we often need to make quick decisions about which group of people to approach or avoid, and facial expressions of the group members play an important role in our judgments of their intent and predisposition. Because such decisions usually need to be made rapidly in many social situations, serial scrutiny of each individual's facial expression is slow and becomes increasingly inefficient as the size of the crowd grows. Instead, extracting summary statistics (e.g., the average) through a process known as ensemble coding¹⁻⁴ is a more efficient way to process an array of similar objects.

A large body of evidence has shown that the visual system can rapidly extract the average of multiple stimulus features such as orientation^{5,6}, size^{7,8}, and motion direction⁹ of groups of objects in an array. Ensemble coding provides precise global representation^{1,7,8,10}, with little or no conscious perception^{6,7,11-13} or sampling of individual members in a set^{14,15}. Recent work has further shown that ensemble coding occurs for even more complex objects, such as averaging emotion from sets of faces^{2,3,16-19}, facial identity^{16,20-23}, as well as a crowd's movements^{24,25} and gaze direction^{26,27}.

Face perception has great social importance, because emotional expressions forecast behavioral intentions of expressors²⁸⁻³⁰ and govern observers' fundamental social motivations accordingly. From a perceiver's perspective, for example, an angry face elicits an *avoidance* reaction while a happy face elicits *approach* reaction^{28,30-32}. To date, however, empirical work undertaken on ensemble perception of faces has largely concentrated on the efficiency and the fidelity of crowd perception^{2,3,16-23}, but not on how this process is socially relevant. To our knowledge, no studies have examined how humans make speeded social decisions about which crowd of faces to approach or avoid, based on extracted ensemble features of facial crowds (e.g., crowd emotion). We often engage in such affective appraisals to enhance our social life (e.g., looking for a more approachable group of people to have a chat with at a cocktail party), and, occasionally, to avoid danger (e.g., rapidly inferring intent to commit violence from the facial expressions of a mob on the street to escape in time and seek help from another group that looks kinder).

Therefore, in the current study we aimed to characterize the behavioral and neural mechanisms of social decision making based on rapidly extracted crowd emotion from groups of faces. Behaviorally, our goals were to examine 1) how extracting crowd emotion from two groups of faces was modulated by task demands reflecting different social motivations (approach or avoidance) and whether this processing was significantly lateralized across the visual field, and 2) how characteristics of facial crowds such as sex-linked identity cues and group size interact with the social motivations present when perceiving crowd emotion. These factors have received surprisingly little attention in the literature of ensemble perception of facial crowds, although they have been found to play a major role in affective processing elsewhere^{28-30,33-39}.

Neurally, our goal was to examine the brain networks and pathways mediating ensemble perception of crowd emotion and compare it to the neural processes underlying extraction of

emotion from, and choosing between, two individual faces. We focused on the dorsal and ventral visual streams, predicting their preferential involvement in processing crowd and individual emotion, respectively. The magnocellular (M) and parvocellular (P) pathways project primarily, but not exclusively, to the dorsal and ventral streams, respectively^{40,41}. The dorsal M-dominant pathway is suggested to support vision for action, non-conscious vision and detection of global and low-frequency information, whereas the ventral P-dominant pathway is suggested to support vision for perception, conscious vision, and analysis of local, high spatial frequency information^{42–52}. Given such distinctive properties and functions of the dorsal and ventral pathways, we hypothesized that decisions on rapidly extracted global information of crowd emotion may rely on the dorsal pathway-dominant processing, whereas decisions that involve comparison between two individual emotional faces may rely on the ventral pathway-dominant processing. We tested this hypothesis by using both whole brain and ROI analyses in the superior frontal gyrus (SFG) and the intraparietal sulcus (IPS) which have been implicated as conveying information via the M-pathway^{53–57} and the fusiform gyrus (FG) which is known to support P-pathway information^{38,56,58–62}. Finally, we also included an additional ROI in the amygdala given its central role in emotional processing^{63–68}. Unlike the other ROIs (IPS, SFG, and FG) that we predicted to be selectively engaged in reading of crowd versus individual emotion, we expect that the amygdala would be involved in both processes.

To accomplish our goals, we conducted a set of behavioral and fMRI experiments in which participants viewed visual stimuli containing two groups of faces with varying emotional expressions (single faces were also examined for a direct comparison in the fMRI study), presented in the left and right visual hemifields. Participants had to choose one of the two crowds (or individual faces) as rapidly as possible, to indicate which one they would avoid or approach. Unlike the estimation task where the absolute value is judged, the answers and the ease of the decision in such comparison task would vary depending on the task goal. For example, the decision to choose to *approach* a happy crowd vs. an emotionally neutral crowd should be quite clear and explicit. However, the same comparison (happy vs. neutral) becomes more ambiguous and implicit when observers have to decide which crowd they would rather *avoid*. This paradigm allows us to examine the role of observers' social motivation in comparing crowd emotion and its interaction with crowd size, sex-linked identity cues, and visual field of presentation.

Results

Experiment 1: behavioral study

Participants viewed two crowds of faces (Figure 1B), one in the left visual field (LVF) and one in the right visual field (RVF) for 1 second. They were instructed to fixate on the center fixation cross and to make a key press as quickly and accurately as possible to indicate which group of faces they would rather avoid (Experiment 1A) or approach (Experiment 1B). Rather than freely choosing, participants were explicitly informed that the correct answer was to choose either the crowd that looked angrier in the avoidance task and the crowd that looked happier on average in the approach task. This allowed us to create the task settings that are applicable in naturalistic social context, by instructing participants to make

relative comparisons between two crowds or faces in order to achieve the explicit social goal (e.g., avoiding a more threatening crowd or approaching a friendlier crowd).

Individual faces contained in each crowd were chosen from a set of 51 faces (Figure 1A) morphed from two highly intense, prototypical facial expressions (angry and happy) of the same person. The set contained six different identities (3 male and 3 female faces), taken from the Ekman face set⁶⁹. One visual field always contained a crowd with varied expressions that were nonetheless emotionally neutral on the average (i.e., the particular mix of happier and angrier expressions was at the midpoint between happy and angry). The other visual field contained a crowd that had a mix of expressions that was either happier or angrier on the average than the neutral crowd. Individual faces had all different emotional intensities, and half of individual faces were more intense in the neutral crowd than any expression in the emotional crowd. This is critical because it ensures that participants could not simply rely on finding the most intense (happy or angry) expression and base their decision on that face. Instead, they had to choose the crowd to approach or avoid based on the *average* emotion from a crowd to perform the task correctly.

While participants reported the task as being rather difficult, they could reliably perform the task at levels well above chance. The overall accuracies for both Experiments 1A and 1B were significantly higher than chance (avoidance task: 64.88% vs. 50%; approach task: 63.72% vs. 50%, all p 's < 0.001 from separate one-sample t tests), demonstrating that participants were able to extract the average crowd emotion from the two groups of faces and choose appropriately which group they would rather avoid or approach. Although accuracies for the avoidance task vs. approach task were not significantly different (64.88% vs. 63.72%: $t(40) = 1.330$, $p = 0.191$), the mean response time (RT)¹ was significantly slower for the avoidance task than for the approach task (1.17 vs. 0.98 seconds: $t(40) = 2.156$, $p < 0.04$, Cohen's $d = 0.666$). As shown in Figure 1D, neither accuracy nor RT was affected by the size of the facial crowd (8 vs. 12 faces total) in the avoidance task (accuracy: $t(40) = 0.113$, $p > .250$, Cohen's $d = 0.035$; RT: $t(40) = 0.010$, $p > .250$, Cohen's $d = 0.003$) or approach task (accuracy: $t(40) = 0.818$, $p > .250$, Cohen's $d = 0.153$; RT: $t(40) = -0.037$, $p > .250$, Cohen's $d = -0.011$), suggesting that extraction of crowd emotion does not require serial processing of each individual crowd member, but is processed in parallel. Because there was no effect of crowd size, we collapsed the data from the different crowd size conditions for further analyses.

Facilitation of task-congruent cues: avoiding angry and approaching happy crowds—In our morphing methods (Figure 1A), the emotional distance between the morphed faces could be quantified based on the arbitrary values of the emotional unit (EU) number, with zero being emotionally ambiguous (e.g., 50% happy and 50% angry), +25 being extremely happy (100% happy), and -25 being extremely angry (100% angry). Because the neutral crowd (EU of zero on average) was always presented on one side, the positive value of the emotional distance between the two crowds indicates that the other side

¹We also conducted RT analyses using each participant's median RT. Just as mean RT, median RT was significantly slower for the avoidance task than for the approach task (1.16 second vs. 0.97 second on average: $t(40) = 1.995$, $p < 0.05$, Cohen's $d = 0.632$). We also confirmed that median RTs yielded the same results for all the other findings reported in this manuscript (see Supplementary Results).

to be compared contained a happier crowd than the neutral crowd (e.g., +9 vs. 0: very happy vs. neutral and +5 vs. 0: somewhat happy vs. neutral) and the negative value of the emotional distance indicates that the other side contained an angrier crowd than the neutral crowd (e.g., -9 vs. 0: very angry vs. neutral and -5 vs. 0: somewhat angry vs. neutral). Such separation proved to be effective in systematically manipulating the difficulty of the task (Figure 1E): A repeated-measures analysis of variance showed a significant main effect of the emotional distance (four levels: -9, -5, +5, and +9) on performance accuracy in both avoidance task ($F(3,60) = 4.69, p < 0.01, \eta_p^2 = 0.29$) and approach task ($F(3,60) = 4.644, p < 0.01, \eta_p^2 = 0.219$). Post hoc Tukey's HSD pairwise comparison tests revealed higher accuracy for +9 than +5 of emotional distance, as well as higher accuracy for -9 than -5 of emotional distance (all p 's < 0.05), suggesting that accuracy increased when emotional distance between the two crowds being compared increased in both the avoidance and approach tasks.

Furthermore, contrast analyses also revealed that participants were more accurate for the crowd emotion that was congruent with the task goal – whether to approach or to avoid. That is, participants were more accurate when comparing angry versus neutral crowds (both levels of emotional distance: -9 and -5) than comparing happy versus neutral crowds (both levels of emotional distance: +9 and +5) during the *avoidance* task, statistically confirmed by the contrast of [-9 and -5 vs. +9 and +5]: $t(80) = 2.37, p < 0.03$. Conversely, during the approach task, participants were more accurate when comparing happy versus neutral crowds (both levels of emotional distance: +9 and +5) than comparing angry versus neutral crowds (both levels of emotional distance: -9 and -5): $t(80) = 2.12, p < 0.04$. The RTs showed similar trends toward faster RTs for comparisons involving task-congruent crowd emotion: angry vs. neutral (emotional distance of -9 and -5) for the avoidance task and happy vs. neutral (emotional distance of +9 and +5) for the approach task (Supplementary Result 1). Together, these results suggest that observers were most accurate and efficient when they had to choose angrier crowds over neutral for the avoidance task and happier crowds over neutral for the approach task. Thus, it appears that motivational information systematically modulates observers' evaluation of crowd emotion.

Right hemisphere dominance for goal-relevant crowd emotion—When participants judged which crowd they would avoid (Experiment 1A), choosing an angry over a neutral crowd is an easier, task-congruent decision (red frame shown in Figure 2A). On the other hand, choosing a neutral over a happy crowd introduces ambiguity into the decision, because the neutral crowd does not contain an explicit social cue (although it is less friendly than a happy crowd, it is not angry on average; gray frame shown in Figure 2A). In both types of decisions, we observed the right hemisphere dominance in which participants' accuracy was facilitated when the facial crowd to be chosen was presented in the LVF. Their accuracy was higher for an angry crowd presented in the LVF than the RVF when comparing an angry vs. a neutral crowd, and a neutral crowd presented in the LVF than the RVF when comparing a neutral vs. a happy crowd during the avoidance task (Figure 2A). A two-way repeated measures ANOVA with two factors of the visual field of presentation (LVF vs. RVF) and the emotional valence of the crowd to be chosen (angry vs. neutral) revealed the significant main effect of the visual field ($F(1,20) = 6.133, p < 0.03, \eta_p^2 = 0.235$), with the

accuracy for LVF presentation being greater than RVF presentation. The main effect of the emotional valence of the crowd to be chosen ($F(1,20) = 0.033, p = 0.858, \eta_p^2 = 0.002$) or the interaction ($F(1,20) = 0.818, p = 0.376, \eta_p^2 = 0.039$) was not significant. Post hoc Tukey's HSD pairwise comparison tests confirmed both the higher accuracy for an angry crowd in the LVF than an angry crowd in the RVF ($p < 0.05$) and the higher accuracy for a neutral crowd in the LVF than a neutral crowd in the RVF ($p < 0.04$). This result indicates hemispheric specialization for crowd emotion processing in which LVF/RH presentations are superior for crowd emotion to be chosen (an angry over a neutral crowd and a neutral over a happy crowd) for the avoidance task.

For the approach task (Experiment 1B), on the other hand, participants had to choose which of the two crowds they would rather approach. It is important to note that the emotional valence of the congruent social cue for the approach task is opposite to that for the avoidance task. For the approach task, choosing a happy over a neutral crowd is a task-congruent social decision (Figure 2B) whereas choosing a neutral over an angry crowd is a more ambiguous decision (gray frame shown in Figure 2B). Despite the emotional valence of a task-congruent social cue being flipped (e.g., angry for avoidance and happy for approach task), we again found a consistent pattern of hemispheric asymmetry: The participants' accuracy for crowd emotion to be chosen (a happy crowd when comparing a happy vs. a neutral crowd and a neutral crowd when comparing a neutral vs. an angry crowd) was facilitated when it was presented in the LVF than in the RVF (Figure 2B). A two-way repeated measures ANOVA showed the significant main effect of the visual field ($F(1,20) = 5.447, p < 0.05, \eta_p^2 = 0.232$), although the main effect of the emotional valence of the crowd to be chosen (happy vs. neutral: $F(1,20) = 0.358, p = 0.556, \eta_p^2 = 0.019$) or the interaction ($F(1,20) = 0.654, p = 0.428, \eta_p^2 = 0.035$) was not significant. Post hoc Tukey's HSD pairwise comparison tests also confirmed the higher accuracy both for a happy crowd in the LVF than the RVF ($p < 0.04$) and a neutral crowd in the LVF than the RVF ($p < 0.02$). Finally, the RT results also indicated faster processing of crowds containing task-congruent cues than of crowds containing task-incongruent cues, both for the avoidance and approach tasks (more details and statistics are reported in the Supplementary Result 1).

Because of the relative comparisons required in our study, in the avoidance task, a neutral crowd could be either the one to be avoided when compared with a happy crowd or the one *not* to be avoided when compared with an angry crowd. Similarly, in the approach task a neutral crowd could be either the one to be approached when compared with an angry crowd or the one *not* to be approached when compared with a happy crowd. One could imagine that on a given trial a neutral crowd was the avoidance/approach stimulus but on the next trial it is the non-avoided/non-approached stimulus depending on the other crowd to be compared with. If participants had used different cognitive strategies for the comparisons of an angry vs. a neutral crowd and of a happy vs. a neutral crowd within the same task, switching the response mapped to a neutral crowd from one trial to the next could significantly slow down participants' RT or impair their accuracy because of the switch cost^{70–74}. In order to examine whether such prior trial interference occurred, we compared participants' RT and accuracy when the response mapped onto a neutral crowd was switched (Switched trials) versus when the same type of comparison was repeated (Repeated trials). We found no difference in RT or accuracy between Switch trials vs. Repeated trials in the

avoidance task or the approach task (all p 's > 0.570, more details shown in Supplementary Result 9), suggesting that the prior trial interference effects and switching costs is minimal in our task.

Together, our results are consistent with previous studies that have shown that the right hemisphere is dominant for global processing whereas the left hemisphere is dominant for local processing^{75–79}. Furthermore, our results suggest that the RH dominance for global processing of crowd emotion is also modulated by the task goal at hand. Regardless of the actual emotional valence of the facial crowd, LVF/RH processing facilitated processing of a goal-relevant facial crowd: a facial crowd that looked “relatively angrier” was facilitated during the avoidance task whereas processing of a facial crowd that looked “relatively happier” was facilitated during the approach task.

Sex-specific identity cues that modulate crowd emotion perception—Because previous findings of individual face perception have documented that female- and male-specific facial features are perceptually confounded with happy and angry expressions, respectively^{33,80}, we examined whether processing of crowd emotion was also modulated by sex-specific facial identity cues. We compared the accuracy for male and female facial crowd stimuli (illustrated in Figure 3A) in the avoidance and approach tasks. Figures 3B and 3C show the accuracy on the avoidance task and approach task, plotted separately by the valence of emotional face images (happy and angry) to be compared to a neutral crowd and the sex of the face images (male vs. female). The two-way repeated-measures ANOVA showed that the main effects of the stimulus sex ($F(1,20) = 2.984$, $p = 0.100$) or of the emotional valence of the face images ($F(1,20) = 0.112$, $p = 0.741$) were not significant in the avoidance task. In the approach task, the main effect of the stimulus sex ($F(1,20) = 4.966$, $p < 0.04$, $\eta_p^2 = 0.199$), but not the main effect of the emotional valence ($F(1,20) = 0.588$, $p = 0.981$), was significant, with the accuracy for female crowds being greater. Importantly, we found the significant interaction between the sex and the emotion of the facial crowd both in the avoidance task ($F(1,20) = 4.908$, $p < 0.04$, $\eta_p^2 = 0.197$) and in the approach task ($F(1,20) = 4.678$, $p < 0.05$, $\eta_p^2 = 0.190$). Post hoc Tukey's HSD pairwise comparison test showed the higher accuracy for angry male than angry female crowds ($p < 0.05$) in the avoidance task and the higher accuracy for happy female than happy male in the approach task ($p < 0.04$). These results suggest that integration of crowd emotion from emotional faces is also influenced by sex-specific identity cues.

Comparing the differing task demands for avoidance and for approach, we also observed a modulation by task demands. Contrast analyses comparing the accuracy for angry males to the other three conditions (happy males, angry females, and happy females) revealed that participants were most accurate for comparing an angry male crowd vs. a neutral male crowd than the other three conditions during the avoidance task ($t(80) = 2.119$, $p < 0.05$), suggesting that facial anger and masculine features both conveyed threat cues and interacted to facilitate decisions to avoid a crowd. Conversely, participants were most accurate in comparing a happy female crowd vs. a neutral female crowd than the other three conditions during the approach task ($t(80) = 2.473$, $p < 0.02$). Although the sex of the faces in our crowd stimuli modulated the perception of crowd emotion, we found that the sex of the

participants did not influence perception of crowd emotion either in the avoidance or the approach task (Supplementary Result 2).

Reading crowd emotion from faces of different identities: Control experiments 1A and B

—In real-world situations, as compared to a laboratory, we never view one person's various emotional expressions simultaneously. Rather, we encounter groups of individuals who differ not only in their emotional expression, but also identity, age, and gender cues. Because a majority of the previous studies of average crowd emotion (including the main experiments in the current study, but see ¹⁹) constructed each stimulus using faces of the same person that varied in emotionality, observers' ability to extract crowd emotion from facial expressions of the same individual vs. from different individuals has not yet been directly compared. Thus, we conducted two Control experiments 1A and 1B in which two new groups of participants (N= 40 total) were presented with crowd stimuli containing a mix of different identities (Supplementary Result 3). As in our main experiments, the task was to choose a relatively angrier crowd of the two to avoid (Avoidance task, Control experiment 1A) and to choose a relatively happier crowd to approach (Approach task, Control experiment 1B). As we reported in detail in Supplementary Result 3, we replicated the results of Main experiments 1A and 1B even with mixed presentation of different facial identities and with new cohorts of participants. We obtained not only similar overall accuracy and overall RT results, also the replication of the task-goal dependent facilitation of the accuracy and the right hemisphere dominance for crowd emotion to be chosen. These results suggest that facial identity cues do not significantly interfere with extraction of crowd emotion. These results also confirm the robustness of our main findings of task-goal dependent modulation and hemispheric asymmetry for task-congruent and task-incongruent decisions on crowd emotion.

Parallel, global processing of facial crowds: Control experiment 2 (Eye tracking study)

—Both in Experiments 1A (avoidance) and 1B (approach), we have found that presenting a larger number of faces in facial crowds (from 8 to 12 total) did not impair participants' accuracy or slow down their RT (Figure 1D). Likewise, in Control experiments 1A and 1B (Supplementary Result 3), we observed no set size effects on the accuracy or RT, suggesting that participants did not process each individual crowd member in a serial manner (see also Supplementary Result 8 for the further analyses and discussion). However, it is still possible that a total duration of 2.5 seconds of stimulus (1 second) and blank (1.5 second) presentation was sufficient for participants to saccade rapidly to only two or three faces in each visual field to make a judgment relying only on these sampled subsets. If this were the case, the flat slope between 8 and 12 faces both in the accuracy and the RT could also be observed as a result of serial processing of subsets, rather than parallel processing. To address this issue, we conducted Control experiment 2 where a new group of 18 participants performed the same avoidance task as in Experiment 1A, with their eye movement monitored and restricted (fixed to the center of the screen) throughout the experiment. We observed not only that both the accuracy and RT were comparable to our main results from Experiment 1A (and Control experiments 1A), but also that the effects of the set size, the emotional distance, and the goal-dependent hemispheric lateralization were replicated in this control eye-tracking experiment. All the details and statistics of Control

experiment 2 are reported in Supplementary Result 4. These results provide the evidence that extracting crowd emotion of facial groups does not necessarily require participants make eye saccades towards a subset of individual faces. Therefore, we conclude that crowd emotion can be extracted as a whole, in a parallel manner rather than relying on serial processing of individual faces or subjects of faces in the crowds.

Experiment 2: fMRI study

In the fMRI study, we scanned 30 participants, using only the avoidance task because of time and budgetary constraints. Participants were presented with stimuli containing either two facial crowds (Figure 1B) or two single faces presented in a crowd of scrambled masks (Figure 1C). Participants were asked to choose rapidly which of the two facial crowds (crowd emotion condition) or which of the two single faces (individual emotion condition) they would rather avoid, using an event-related design with crowd emotion and individual faces conditions randomly intermixed (See Methods for more details). We compared the patterns of brain activation when participants chose to avoid one of two crowds or one of two individual faces. If the processing of crowd emotion relies on the same mechanism that mediates single face perception, we would observe activations of the same brain network during the processing of crowd emotion, but perhaps to a larger degree and larger extent than during single face comparisons, given the greater complexity of the stimulus and difficulty of the crowd emotion task. Alternatively, if the processing of crowd emotion and of individual face emotion relies on qualitatively distinct processes, specifically mediated by dorsal and ventral visual pathways as we hypothesized, we would expect to observe differential brain activations in distinct sets of brain areas in dorsal and ventral visual pathways.

Behavioral results

The participants' overall accuracy for crowd emotion condition in the fMRI study was 63.16%, not significantly different from that we observed from Experiment 1A (64.88%; $t(48) = -1.468, p = 0.149$). We further confirmed that the behavioral results of the crowd emotion condition in the fMRI study replicated other behavioral results in Experiment 1A and the control experiments: Participants' accuracy was higher when emotional distance between the two crowds being compared increased, when comparing angry (task-congruent cue) vs. neutral crowds than comparing happy vs. neutral crowds, and when comparing angry male vs. neutral male crowds than comparing angry female vs. neutral female crowds. The figures and statistical tests supporting these results are presented in Supplementary Result 5.

Critically, we again replicated the right hemisphere advantage for crowd emotion to be chosen, as in Experiment 1A (Figure 2A) and the control experiments (Supplementary Results 3 and 4). The participants' accuracy was higher when the crowd emotion to be chosen was presented in the LVF than the RVF, both in comparing an angry vs. a neutral crowd and comparing a neutral vs. a happy crowd (Figure 2C, color-filled bar graphs). This was confirmed by a significant main effect of the visual field of presentation (LVF vs. RVF: $F(1,29) = 4.560, p < 0.05$) from a two-way repeated measures ANOVA with the accuracy for the LVF being greater than the RVF. The main effect of the emotional valence of the crowd

to be chosen was not significant (angry vs. neutral: $F(1,29) = 1.141, p = 0.294$), although the interaction was significant ($F(1,29) = 9.361, p < 0.01$). Post hoc Tukey's HSD pairwise comparison tests also confirmed the higher accuracy for an angry crowd in the LVF than the RVF ($p < 0.03$) and for a neutral crowd in the LVF than the RVF ($p < 0.05$).

For the individual emotion condition, the overall accuracy was 65.92%, which was slightly but not significantly, higher than that for crowd emotion condition ($t(58) = -1.491, p = 0.106$). We also observed no difference in the RT for the crowd emotion condition vs. the individual emotion condition ($t(58) = 0.318, p = 0.751$). Even though only two faces were presented, and thus there was no need to extract the average crowd emotion, the level of accuracy and the response time for comparing two individual faces was similar to that for comparing two facial crowds. These results confirm that the difference in our fMRI findings comparing crowd emotion vs. individual emotion conditions is not due to a difference in task difficulty, but reflects qualitative differences in neural processing patterns and substrates.

Finally, we observed different patterns of hemispheric lateralization for the individual emotion condition (Figure 2C, bar graphs with outlines). Unlike the crowd emotion condition showing the right hemisphere advantage regardless of the emotional valence of the crowd emotion to be chosen, the individual emotion condition showed that participants' accuracy was higher when an angry face was presented in the LVF/RH and when a neutral face was presented in the RVF/LH. In addition, choosing an angry face over a neutral face was more accurate than choosing a neutral face over a happy face overall. A two-way repeated measures ANOVA showed that the main effect of the visual field of presentation was not significant for the individual emotion condition ($F(1,29) = 1.702, p = 0.202$), although the main effect of the emotional valence of the face to be chosen (angry vs. neutral: $F(1,29) = 18.511, p < 0.01$) and the interaction ($F(1,28) = 8.193, p < 0.01$) were significant. Post hoc Tukey's HSD pairwise comparison tests also confirmed the higher accuracy for an angry face in the LVF than in the RVF ($p < 0.05$) and the higher accuracy for an angry face than a neutral face in the LVF ($p < 0.05$), but not in the RVF ($p > 0.76$). Our findings in the individual emotion condition are consistent with previous findings suggesting that affective face processing in general is right-lateralized, with more marked laterality effects for negatively-valenced stimuli^{35,80–82}. Together, our behavioral data from the fMRI study replicate our main findings from Experiment 1A (and other control experiments) and provide further evidence that perception of crowd emotion and individual emotion engage different patterns of hemispheric specialization.

fMRI results

Distinct neural substrates for crowd emotion vs. individual emotion processing

—Our main goal in the fMRI experiment was to characterize the neural substrates involved in participants' avoidance decision between two facial crowds vs. those mediating decisions between two individual faces. Figure 4A shows the brain regions activated when participants were comparing two crowds (labeled in red) vs. comparing two individual faces (labeled in blue) and vice versa (The complete list of activations is reported in the Table 1). Using the contrast Crowd emotion - Individual emotion conditions, we observed that comparing two facial crowds in the avoidance task showed greater cortical

Author Manuscript

activations in the parietal and frontal regions along the dorsal stream (e.g., IPS, SPL, SFG, MFG, IFG, and OFC) in both hemispheres. On the other hand, the Individual emotion condition evoked greater activation in the regions along the ventral stream (FG, PHC, RSC, TP), ACC/PCC, and vmPFC. In order to verify the robustness of our results, we also used several different thresholding parameters ($p < 0.001$, uncorrected; $k = 5$ in Supplementary Table 1 and $p < 0.0005$, uncorrected; $k = 5$ in Supplementary Table 2) and observed similar patterns with greater activations in the regions along the dorsal pathway for crowd emotion condition and in the regions along the ventral pathway for individual emotion condition.

Author Manuscript

We further examined responses in our regions of interest (IPS, SFG, FG, and amygdala). We chose these ROIs based on the prior work that showed involvement of SFG in dorsal stream processing and its intrastream functional connectivity with IPS^{53–57}, the work showing the major involvement of FG in ventral stream information processing^{38,56,58–62}, and the wealth of data supporting the link between the amygdala and emotional functions in general^{63–68}. These regions were functionally restricted based on an unbiased contrast of all the visual conditions minus baseline (average activation of voxels) using random effects models (height: $p < 0.01$, uncorrected; extent: 5 voxels), within the anatomical label for each ROI (obtained by the anatomical parcellation of the normalized brain⁸³). As shown in Figure 4B, we observed that both the left and right IPS and SFG showed greater activation for the crowd emotion condition than the individual emotion condition, whereas the bilateral FG showed greater activation for the individual emotion condition than crowd emotion condition (See Supplementary Figure 10 for the breakdown of the emotional valence for each of the condition as well). This observation was confirmed by paired t-tests (two-tailed), conducted separately for each ROI (crowd emotion > individual emotion: $p < 0.01$ for the bilateral IPS and SFG; individual emotion > crowd emotion: $p < 0.01$ for the bilateral FG). These results suggest that the regions in the dorsal visual pathway (e.g., IPS and SFG) and in the ventral visual pathway (e.g., FG) show differential responsivity to the crowd emotion and the individual emotion comparisons, respectively. Unlike these three ROI's that showed selective responses to crowd emotion (IPS and SFG) and individual emotion (FG) conditions, we observed that the bilateral amygdala did not distinguish crowd emotion and individual emotion conditions. The amygdala showed similar level of responsivity to both conditions, confirmed by the non-significant t-test between crowd vs. individual emotion conditions (two-tailed): $p > 0.693$). The null result on the amygdala activation suggests that the differential responses of IPS, SFG, and FG we observed here are pathway-specific.

Author Manuscript

Furthermore, we found that the activity in these ROIs differentially predicted the participants' behavioral accuracy for the crowd emotion and individual emotion conditions. As shown in Figure 4C, the responses of the bilateral IPS and the right SFG positively correlated with participants' accuracy for the crowd emotion condition (left IPS: $r = 0.424$, $p < 0.01$; right IPS: $r = 0.462$, $p < 0.01$; left SFG: $r = 0.337$, $p = 0.069$; right SFG: $r = 0.454$, $p < 0.02$), whereas the activity of the bilateral FG positively correlated with the accuracy for the individual emotion condition (left FG: $r = 0.568$, $p < 0.01$; right FG: $r = 0.498$, $p < 0.01$). However, we did not find any significant correlation between participants' amygdala activity and behavioral accuracy for crowd emotion condition (left amygdala: $r = 0.266$, $p = 0.155$; right amygdala: $r = 0.192$, $p = 0.310$) or individual emotion condition (left amygdala: $r = 0.133$, $p = 0.482$; right amygdala: $r = 0.254$, $p = 0.176$). Together, our fMRI results provide

the evidence for the differential contributions of the dorsal and ventral visual pathways to crowd emotion and individual emotion processing, respectively.

Finally, we verified that our ROIs are the regions selectively engaged in M-pathway (IPS and SFG) and P-pathway (FG) processing, by conducting a pilot study using functional localizer scans for M-pathway and P-pathway regions (Supplementary Result 7). In this pilot study, we used non-face stimuli, sinusoidal counter-phase flickering gratings biased to engaged the M-pathway (a low-luminance contrast, varying gray-on-gray grating with a low spatial frequency and 15 Hz flicker) and biasing P-pathway (an isoluminant, high color-contrast red-green grating with high spatial frequency with slow, 5 Hz flicker), following the method detailed in Denison et al⁸⁴. These types of stimuli have been known to selectively bias M-pathway and P-pathway processing in previous studies using object, letter, scene, and face stimuli^{44,45,50,84–87}. Although different types of stimuli and different cohorts of participants were employed, the localizer scans showed activations in foci adjacent to and overlapping with our ROIs including the bilateral IPS and SFG activated by M-pathway stimuli and the bilateral FG activated by P-pathway stimuli.

Brain activations for comparing angry vs. neutral crowds and neutral vs. happy crowds—Our secondary interest was to examine the different patterns of brain activation when participants had to choose an angry over a neutral crowd (obvious and task-congruent comparison) vs. when they had to choose a neutral crowd over a happy crowd (less-obvious comparison). As shown in Figure 5A, we found greater activations evoked by the avoidance comparison between an angry vs. a neutral crowd in the brain areas along the dorsal visual pathway (SFG, MFG, IFG, SMA, and OFC) and the vIPFC, PCC, amygdala, and PHC in the right hemisphere, and in the left SPL (Figure 5A). Overall, many of these regions overlapped with the areas showing preference for crowd emotion than individual emotion processing (Figure 4A and Table 1), suggesting that task-congruent decision is dominant in the crowd emotion processing of these dorsal pathway regions. For the avoidance comparison between a neutral vs. a happy crowd, we observed greater activations in the bilateral MTG and SFG, STS, PCC, PHC, ACC, mPFC in the left hemisphere. Some of these areas including the MTG, PCC, and ACC are also known to be engaged in emotional conflict resolution, coupled with DLPFC which is responsible for integration of emotional and affective processes^{88,89,132}. Thus, this result suggests that choosing a neutral vs. a happy crowd as a relative option to be avoided may engage an additional control mechanism. The control mechanism in these areas appears to be responsible for resolution of possible emotional conflict between the task goal (e.g., avoidance) and the motivational value elicited by a happy facial crowd (e.g., approach). The complete list of the brain activations in Figure 5A is shown in Supplementary Table 3. Finally, comparing an angry vs. a neutral face evoked greater activation in the brain areas including the bilateral AG, MTG, PHC and PCC/PC, the right TP, PFC, and OFC, whereas comparing a neutral vs. a happy face evoked greater activation in the areas including the left IFG, insula, TP, and the SMA (shown in Figure 5B). The complete list of the brain activations in Figure 5B is shown in Supplementary Table 4.

Discussion

The goal of this study was to characterize the functional and neural mechanisms that support crowd emotion processing. We had four main findings: 1) Goal-dependent hemispheric asymmetry for crowd emotion processing in which presenting a facial crowd to be chosen (e.g., either an angry or a neutral crowd for the avoidance task and either a happy or a neutral crowd for the approach task) in the LVF/RH facilitated participants' accuracy, 2) Higher accuracy overall in identifying facial crowds containing task-congruent crowd emotion (e.g., angry crowd to avoid, happy crowd to approach); 3) Higher accuracy when sex-linked identity cues were congruent with the task goal (e.g., angry male crowds to avoid and happy female crowds to approach); and 4) Preferentially activated dorsal visual stream in crowd emotion processing, with IPS and SFG predicting behavioral crowd emotion efficiency; and 5) Preferentially activated ventral visual stream in individual face emotion processing, with fusiform cortex activity predicting the accuracy of decisions about individual face emotion.

The goal-dependent modulation of crowd emotion processing suggests that the mechanism underlying the reading of crowd emotion is highly flexible and adaptive, allowing perceivers to focus most keenly on desired outcomes in dynamic social contexts (e.g., to avoid unfriendly crowds or to approach friendly ones). Neither the stimulus display nor the response characteristics changed between avoidance and approach tasks: the only difference was the decision (approach or avoid) that was mapped to the response. The same visual stimuli containing facial crowds appear to be biased differently depending on whether the task goal was to avoid or to approach. Stimulus gender also interacted with the processing of crowd emotion, in a manner relevant to the current goal. Such visual integration of compound social cues (e.g., gender, emotion, race, eye gaze, body language, etc.) has been well incorporated into the theories of mechanisms underlying single face perception⁹⁰. However, the roles of these compound social cues in ensemble coding of facial crowds have not been examined. The current study provides evidence that intrinsic (social motivation shaped by the task instruction in this study) and extrinsic (e.g., emotional expressions and sex of the crowds) factors also mutually facilitate the reading of crowd emotion in a manner that is functionally related to the task at hand. Future study will be needed to further explore how other relevant information about environment and context modulate perception of crowd emotion and individual emotion differently and how perceived crowd and individual emotion yield different emotional intensity and confidence of observers in order to initiate an appropriate social response toward it.

Our findings also provide evidence that the processing that supports social decisions on crowd emotions are highly lateralized, in a manner that is relevant to the current task goal. Lateralized behavioral responses provide an opportunity to study the hemispheric asymmetries that enable cognitive functions. This hemispheric asymmetry enables flexible and adaptive processing optimized for the current task goal in dynamic environments⁹¹, supporting the selection of appropriate, and inhibition of inappropriate, responses⁹². This is particularly useful when a large number of complex stimuli (such as a crowd of emotional faces) and competing cognitive goals tax the processing capacity of the visual system, as was the case in our task. The pattern of hemispheric lateralization in affective processing has

been traditionally thought to rely on emotional valence^{35,93}, with aversive or negative stimuli lateralized to the right hemisphere (RH) and positive, approach-evoking stimuli lateralized to the left hemisphere (LH). However, we instead found that the lateralization effects in crowd emotion processing are actually goal-dependent, rather than being driven purely by stimulus valence. Our findings demonstrate that the same emotional stimuli can be biased differently in RH and LH, depending on the task goal and observers' intent. Specifically, the current task goal (approach or avoidance) biased processing such that LVF/RH was superior for recognizing a "relatively angrier" facial crowd during the avoidance task and for processing a "relatively happier" facial crowd during the approach task. Since the task-relevant emotions were anger and happiness for the avoidance and the approach tasks, respectively, our results suggest that global processing of ensemble face emotion most strongly engages the RH in general, consistent with a popular concept of the dominant role of RH in global processing⁷⁵⁻⁷⁹. Unlike the crowd emotion condition, however, we observed a RH advantage for an angry face and a LH advantage for a neutral face. This also conforms to the traditional framework of single face processing, with RH preference for aversive or negative face stimuli and LH preference for positive, approach-evoking stimuli^{31,35,82,94}. Together, our data suggest that reading crowd emotion and single face emotion show different patterns of the hemispheric lateralization, with crowd emotion processing showing the flexible modulation of the RH dominance depending on the task goal, rather than being based purely on stimulus valence.

It is worth noting that the inference should be made very carefully from the divided visual field paradigm because it is a relatively indirect approach to localizing hemispheres with cognitive functions⁹⁵. In particular, the interpretation of the results becomes very difficult when participants shift their gaze. To ensure that participants did not move their eyes while they performed the current task, we explicitly instructed participants to initiate each trial only after they fixated the central fixation cross. Moreover, we verified from the control eye-tracking experiment (Control experiment 2) that we could replicate our results from the main Experiments (1A and 2) when participants' eye movements were monitored and restricted (Supplementary Result 4). Finally, we also replicated this hemispheric lateralization for crowd emotion processing in all the main and control experiments (see Supplementary Result 6). Therefore, we conclude that our results are robust to confounding factors such as variability in facial stimuli, differences in experimental settings (behavioral, eye-tracking, and fMRI), and participants' eye movements and variability across different cohorts of participants.

For a given pair of facial crowds to be compared in our study, participants had to choose which one of the two is to be avoided (avoidance task) or approachable (approach task) *relatively more*. Although both an angry crowd compared to a neutral crowd and a neutral crowd compared to a happy crowd showed a RH advantage during the avoidance task, these two different types of comparisons may yield a different amount of emotional ambiguity and task difficulty. In our fMRI study, we found greater activations evoked by goal-congruent comparison (an angry vs. a neutral crowd) in the brain areas along the dorsal visual pathway (SFG, MFG, IFG, SMA, and OFC) as well as the vIPFC, PCC, amygdala, and PHC in the right hemisphere, and in the left SPL. This result suggests that task-congruent decision may occur predominantly in the crowd emotion processing of these dorsal pathway regions, with

the RH more pronounced. When comparing a neutral vs. a happy crowd, however, we observed greater activations in the bilateral MTG and the left SFG, STS, PCC, PHC, ACC, mPFC. The MTG, PCC, and ACC are known to be engaged in emotional conflict resolution, coupled with DLPFC, which is responsible for integration of emotional and affective processes^{19,88,89}. When a goal-incongruent emotional stimulus interferes with a goal-congruent emotional stimulus, emotional conflict occurs⁹⁶. Thus, this finding suggests that choosing a neutral over a happy crowd during the avoidance task may involve an emotional conflict control mechanism^{97,98} in which possible interference by goal-incongruent facial crowd (e.g., a happy crowd) is suppressed and ambiguity of a neutral crowd is resolved as a relatively better option for avoidance decision.

Different patterns of hemispheric lateralization for the crowd emotion vs. individual emotion processing suggest that they may rely on qualitatively distinct systems. Much behavioral evidence has been accumulated supporting this notion^{6,7,11–15,21,99–103}, but see 104, although only a few recent fMRI studies have compared the neural representations of ensemble coding and individual processing^{105–107}. Cant and Xu^{105,106} showed that PPA and LO were preferentially engaged in texture perception and object processing, respectively; and Huis in't Veld and de Gelder¹⁰⁷ showed the greater anticipatory and action preparation activity in the areas including IPL, SPL, SFG, and premotor cortex for interactive body movement of a group of panicked people, compared to an unrelated movement of individuals. However, unlike prior work that used stimuli of simple texture patches and objects^{105,106} or that removed the information about facial expression of people from their blurred video clips¹⁰⁷, the current study examined the distinct neutral substrates underlying the processing of facial crowds with varying emotional expressions compared to the processing of individual emotion expressions, providing evidence for distinct mechanisms supporting them.

The benefit of having distinct systems for ensemble coding and individual object processing is that these two processes can serve complementary functions. Global information extracted via ensemble coding influences processing of individual objects in many different ways. Because ensemble coding compresses properties of multiple objects into a compact description with a higher level of abstraction¹⁰⁸, it allows observers to surmount the severe limitations on individual object processing^{109–113}, imposed by attention or working memory^{114–117}. Furthermore, the global information of ensembles allows for an initial, rough analysis of visual inputs, which then biases and facilitates the processing of individual objects^{108,109,113,118}. For example, extracted ensemble representation influences the individual object processing, by guiding detection of outliers in a set (e.g., pop-out visual search^{108,118}), facilitating selection of an individual object at the center location of a set¹¹³, and biasing memory for individual objects towards the global mean¹¹⁹. Our results from the whole brain and the ROI analyses both support our hypothesis that the dorsal visual stream contributes to global processing for crowd emotion extraction whereas the ventral visual stream contributes more to object-based (local) processing of emotion in individual faces.

Processing of crowd emotion appears to be achieved in a global, parallel fashion, rather than serially for the following reasons. First of all, we found that the RTs for the crowd emotion condition were equivalent to those for the individual emotion, despite many more faces that needed to be processed in the crowd condition compared to the individual face condition

(see Supplementary Result 8). Second, simply sampling any one or two individual faces from each crowd in our stimuli would lead to the chance level of accuracy, because a half of the individual members in the neutral crowd were always more intense than a half of the angry crowd in our stimuli. Thus, sampling faces with an extreme emotional expression can not be an explanation for our findings. Third, the participants were equally accurate and fast when they viewed the facial crowds containing 8 faces or 12 faces. Finally, fixating eyes to the center did not impair participants' accuracy or RT for extracting crowd emotion, indicating that making saccades towards individual faces in the crowd to foveate them is not essential to effectively evaluate crowd emotion in our task. Therefore, we suggest that extracting crowd emotion relies on a parallel, global process, rather than on a sequential sampling of individual members¹⁰⁴. In various feature dimensions, the notion of global averaging has been previously tested, by using empirical approaches showing that multiple stimuli were integrated²², ideal observer analysis^{3,15}, equivalent noise²⁶, or general linear modeling¹⁷. Consistent with this prior work, the current findings suggest that people do average different facial expressions to make social decisions about facial crowds, and such ensemble coding of crowds of faces is achieved via a distinct mechanism from that supporting individual object processing.

To conclude, here we have reported evidence for distinct mechanisms dedicated to processing of crowd emotion and individual face emotion, which are biased towards different visual streams (dorsal vs. ventral), and show different patterns of hemispheric lateralization. The differential engagement of the dorsal stream regions and the complementary functions of the left and right hemispheres both suggest that processing of crowd emotion is specialized for action execution that is highly flexible and goal-driven, allowing us to trigger a rapid and appropriate reaction to our social environment. Furthermore, we have shown that observers' goals – to avoid or approach – can exert powerful influences on the perception accuracy of crowd emotion, highlighting the importance of understanding the interplay of ensemble coding of crowd emotion and social vision.

General Methods

Participants

In Experiment 1, a total of 42 undergraduate students participated: 21 subjects (12 female) participated in the avoidance task (Exp.1A) and a different cohort of 21 participants (11 female) participated in the approach task (Exp.1B). A power analysis (N^{*120}) based on a pilot run of this experiment with four subjects indicated that 21 subjects were enough to achieve at least 80% power. No subjects were excluded from the behavioral data analysis. In Experiment 2, a new group of 32 (18 female) undergraduate students participated. Two participants were excluded from further analyses because they made too many late responses (e.g., RTs longer than 2.5s). Thus, the behavioral and fMRI analyses for Experiment 2 were done with a sample of 30 participants. All the participants had normal color vision and normal or corrected-to-normal visual acuity. Their informed written consent was obtained according to the procedures of the Institutional Review Board at the Pennsylvania State University. The participants received monetary compensation or a course credit.

Apparatus and stimuli

Stimuli were generated with MATLAB and Psychophysics Toolbox^{121,122}. In each crowd stimulus (Figure 1A), either 4 or 6 morphed faces were randomly positioned in each visual field (right and left) on a grey background. Therefore, our facial crowd stimuli comprised either 8 or 12 faces. We used a face-morphing software (Norrkross MorphX) to create a set of 51 morphed faces from two highly intense, prototypical facial expressions of the same person for a set of six different identities (3 male and 3 female faces), taken from the Ekman face set⁶⁹. The morphed face images were controlled for luminance, and the emotional expression of the faces ranged from happy to angry (Figure 1A), with 0 in Emotional Unit (EU) being neutral (morph of 50% happy and 50% angry), +25 in EU being the happiest (100% happy), and -25 in EU being the angriest (100% angry). Because the morphed face images were linearly interpolated (in 2% increments) between two extreme faces, they were separated from one another by EU of intensity such that Face 1 was one EU happier than Face 2, and so on. Therefore, the larger the separation between any two morphed faces in EU, the easier it was to discriminate them. Such morphing approach was adapted from the previous studies on ensemble coding of faces¹⁶.

Since the previous literature on averaging of other visual features showed that the range of variation is an important determinant of averaging performance (e.g., size or hue^{123,124}), we kept the range of faces the same (i.e., 18 in emotional units) across the two set sizes. This emotional unit was determined from initial pilot work in order to ensure that each individual face is distinguishable from one another within the range and that the task is not so easy as to produce accuracy ceiling effects. One of the two crowds in either left or right visual field always had the mean value of zero in emotional units, which is neutral on average, and the other had the emotional mean of +9 (very happy; morphing of angry 32% and happy 68%), +5 (somewhat happy; morphing of angry 40% and happy 60%), -9 (very angry; morphing of angry 68% and happy 32%), and -5 (somewhat angry; morphing of angry 60% and happy 40%). Thus, the sign of such offset between the emotional and neutral crowds in EU indicates the valence of the emotional crowd compared to the neutral: The positive values indicate more positive (happier) crowd emotion compared with the neutral and the negative values indicate more negative (angrier) mean emotion.

In order to avoid the possibility that participants simply sampled one or two single faces from each set and compared them to do the crowd emotion task, we ensured that 50% of the individual faces in the neutral set were more expressive than 50% of the individual faces in the emotional sets to be compared. For example, half of the members of the neutral set were angrier than a half of the members of the angry crowd. This manipulation allowed us to assess whether participants used such “sampling strategy¹⁰⁴” rather than extracting an average, because sampling one or two members in a set would yield 50% of accuracy in this setting.

Stimuli for the individual emotion condition (Figure 1C; only included in the fMRI study) comprised one emotional face (either angry or happy) and one neutral face from the same set of morphed face images randomly positioned in the same invisible frame surrounding the crowd stimuli in each visual field. The offsets between the emotional and neutral faces remained the same as those in facial crowd stimuli. To ensure that the difference is not due

to the confound of simply having more “stuff” in crowd emotion condition, compared to the individual emotion condition, we included scrambled faces in the individual emotion condition so that the same number of the face-like blobs were presented as in the crowd emotion condition. This ensured that any differences are not due to low-level visual differences in the stimulus displays, but rather to how many resolvable emotional faces participants had to discriminate on each trial (2 vs. 8 or 12).

On one half of the trials, the emotional stimulus (i.e., happy or angry: ± 5 or ± 9 EU away from the mean) was presented in the left visual field and the neutral stimulus was presented in the right visual field, and it was switched for the other half of the trials. Each face image subtended $2^\circ \times 2^\circ$ of visual angle, and face images were randomly positioned within an invisible frame subtending $13.29^\circ \times 18.29^\circ$, each in the left and right visual fields. The distance between the proximal edges of the invisible frames in left and right visual fields was 3.70° .

Procedure

Participants in Experiment 1 sat in a chair at individual cubicles about 61 cm away from a computer with a 48 cm diagonal screen (refresh rate = 60 Hz). Participants in Experiment 2 were presented with the stimuli rear-projected onto a mirror attached to a 64-channel head coil in the fMRI scanner. Figure 1B illustrates a sample trial of the experiment. Participants were presented with visual stimuli for 1 second, followed by a blank screen for 1.5 seconds. The participants were instructed to make a key press as soon as possible to indicate which of the two crowds of faces or two single faces on the left or right they would rather avoid (Experiment 1A and Experiment 2) or approach (Experiment 1B). Participants pressed ‘f’ key for choosing the LVF and ‘j’ key for choosing the RVF in the behavioral experiments (1A and 1B). In the fMRI scanner (Experiment 2), they pressed ‘1’ key for the LVF and ‘4’ key for the RVF using the response box. They used both left and right index fingers. Key-response assignment was not counterbalanced in order to maintain the automatic and consistent stimulus-response compatibility¹²⁵ (left key for the LVF and right key for the RVF). They were explicitly informed that the correct answer was to choose either the crowd or the face showing a more negative (e.g., angrier) emotion for the avoidance task and a more positive (e.g., happier) emotion for the approach task. Responses that were made after 2.5 seconds were considered late and excluded from data analyses. Feedback for correct, incorrect, or late responses was provided after each response. Before the actual experiment session, participants were provided with 20 practice trials that were conceptually identical to the actual trials.

In Experiment 1, half of the participants performed the avoidance task and the other half performed the approach task. Experiment 1 had a 4 (emotional distance between facial crowds, -9 , -5 , 5 , or 9) \times 2 (visual field of presentation, LVF and RVF) \times 2 (set size: 4 or 6 faces in each visual field) design, and the sequence of total 320 trials (20 repetitions per condition) was randomized. In Experiment 2 (fMRI), all the participants performed the avoidance task. Because we needed more trials for statistical power for fMRI data analyses and we observed no effect by the number of crowd members on crowd emotion perception (Figure S1), we only used crowd stimuli containing 4 faces in Experiment 2. Thus,

Experiment 2 had a 2 (stimulus type: crowd and individual) x 4 (emotional distance) x 2 (visual field of presentation) design and 112 additional null trials (background trial without visual stimulation). The sequence of total 624 trials including 512 experimental trials (32 repetitions per condition) and 112 null trials was optimized for hemodynamic response estimation efficiency using the *optseq2* software (<https://surfer.nmr.mgh.harvard.edu/optseq/>).

fMRI data acquisition and analysis

fMRI images of brain activity were acquired using a 3 T scanner (Siemens Magnetom Prisma) located at The Pennsylvania State University Social, Life, and Engineering Sciences Imaging Center. High resolution anatomical MRI data were acquired using T1-weighted images for the reconstruction of each subject's cortical surface (TR = 2300 ms, TE = 2.28 ms, flip angle = 8°, FoV = 256 x 256 mm², slice thickness = 1 mm, sagittal orientation). The functional scans were acquired using gradient-echo EPI with a TR of 2000 ms, TE of 28ms, flip angle of 52° and 64 interleaved slices (3 x 3 x 2 mm). Scanning parameters were optimized by manual shimming of the gradients to fit the brain anatomy of each subject, and tilting the slice prescription anteriorly 20–30° up from the AC-PC line as described in the previous studies^{44,126,127}, to improve signal and minimize susceptibility artifacts in the brain regions including OFC and amygdala¹²⁸. We acquired 780 functional volumes per subject in four functional runs, each lasting 6.5 min.

The acquired fMRI images were pre-processed using SPM8 (Wellcome Department of Cognitive Neurology). The functional images were corrected for differences in slice timing, realigned, corrected for movement-related artifacts, coregistered with each participant's anatomical data, normalized to the Montreal Neurological Institute template, and spatially smoothed using an isotropic 8-mm full width half-maximum Gaussian kernel. Outliers due to movement or signal from preprocessed files, using thresholds of 3 SD from the mean, 0.75 mm for translation and 0.02 radians rotation, were removed from the data sets, using the ArtRepair software¹²⁹. Subject-specific contrasts were estimated using a fixed-effects model. These contrast images were used to obtain subject-specific estimates for each effect. For group analysis, these estimates were then entered into a second-level analysis treating participants as a random effect, using one-sample t-tests at each voxel. Six contrasts of our interest were used: 1) Crowd emotion minus Individual emotion, 2) Individual emotion minus Crowd emotion, 3) Angry crowd minus Happy crowd, 4) Happy crowd minus Angry crowd, 5) Angry individual minus Happy individual, and 6) Happy individual minus Angry individual. These contrasts were thresholded at $p < 0.05$ (FWE whole-brain corrected) and a minimum cluster size of 5 voxels. For visualization and anatomical labeling purposes, all group contrast images were overlaid onto the inflated group average brain, by using 2D surface alignment techniques implemented in FreeSurfer¹³⁰.

For the region of interest (ROI) analyses, we extracted the BOLD activity from the bilateral IPS, SFG, FG, and amygdala. We defined a separate contrast between all the visual stimulation trials (all trials containing stimuli) vs. background (Null trials). From this contrast, we localized each of the ROIs based on the peak activation within the anatomical label obtained by the anatomical parcellation of the normalized brain⁸³. The [x y z]

coordinates for these ROIs were $[-30/30 -67/-70 52/42]$ for IPS (L/R), $[-27/36 5/6 66/58]$ for SFG (L/R), $[-30/33 -58/-58 -10/-10]$ for FG (L/R), and $[-18/18 -1/-4 -18/-12]$ for amygdala (L/R). The coordinates for the right IPS and the right SFG were adjacent to the regions that have been reported in the previous study as showing the robust intra-stream connectivity (dorsal visual stream⁵⁶). Moreover, the coordinate for the right FG has been also localized as the right FFA (Fusiform Face Area) in the activation maps by Spiridon, Fischl, & Kanwisher¹³¹. The beta weights were extracted for crowd emotion and individual emotion conditions using the rfxplot toolbox (<http://rfxplot.sourceforge.net>) for SPM. We defined a 6mm sphere around the [x y z] coordinate for each of our ROIs. Using the rfxplot toolbox in SPM8, we extracted all the voxels from each individual participant's functional data within that sphere. In order to achieve our main goal to map out the brain regions that are preferentially engaged in crowd emotion processing vs. individual emotion processing, we collapsed the four different levels of emotional distance (-9, -5, +5, and +9) and the two levels of visual field of presentation (LVF and RVF). On the extracted beta estimates for the two main conditions (Crowd emotion and Individual emotion) from each of the ROIs, we conducted the paired t-tests to compare the % signal change between Crowd emotion vs. Individual emotion conditions and the correlation analyses to examine the relationship with the behavioral accuracy measurements. In Supplementary Figure 10, we also report the beta estimates from the breakdown of the emotional valence for Crowd emotion and Individual emotion conditions (four conditions total: Angry Crowd, Happy Crowd, Angry Individual, and Happy Individual).

Data availability

The datasets generated during and/or analyzed during the current study are available from the first author (him3@mgh.harvard.edu) or the corresponding author (kestras@nmr.mgh.harvard.edu) on reasonable request.

Code availability

All the MATLAB codes for the behavioral and fMRI analyses presented in this article are available from the first author (him3@mgh.harvard.edu) or the corresponding author (kestras@nmr.mgh.harvard.edu) upon the request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health R01MH101194 to K.K. and to R.B.A., Jr.

References

1. Alvarez GA. Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*. 2011; 15(3):122–131. [PubMed: 21292539]
2. Fischer J, Whitney D. Object-level visual information gets through the bottleneck of crowding. *Journal of Neurophysiology*. 2011; 106:1389–1398. [PubMed: 21676930]

3. Haberman J, Harp T, Whitney D. Averaging facial expression over time. *Journal of Vision*. 2009; 9:1–13.
4. Haberman J, Whitney D. Efficient summary statistical representation when change localization fails. *Psychonomic Bulletin and Review*. 2011; 18:855–859. [PubMed: 21748419]
5. Dakin SC, Watt RJ. The computation of orientation statistics from visual texture. *Vision Research*. 1997; 37:3181–3192. [PubMed: 9463699]
6. Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M. Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*. 2001; 4:739–744. [PubMed: 11426231]
7. Ariely D. Seeing sets: Representation by statistical properties. *Psychological Science*. 2001; 12:157–162. [PubMed: 11340926]
8. Chong SC, Treisman A. Representation of statistical properties. *Vision Research*. 2003; 43:393–404. [PubMed: 12535996]
9. Watamaniuk SNJ, Sekuler R. Temporal and spatial integration in dynamic random-dot stimuli. *Vision Research*. 1992; 32:2341–2347. [PubMed: 1288010]
10. Halberda J, Sires SF, Feigenson L. Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science*. 2006; 17:572–576. [PubMed: 16866741]
11. Alvarez GA, Oliva A. The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*. 2008; 19:392–398. [PubMed: 18399893]
12. Choo H, Franconeri SL. Objects with reduced visibility still contribute to size averaging. *Attention, Perception & Psychophysics*. 2010; 72:86–99.
13. Corbett JE, Oriet C. The whole is indeed more than the sum of its parts: Perceptual averaging in the absence of individual item representation. *Acta Psychologica*. 2011; 138:289–301. [PubMed: 21903186]
14. Haberman J, Whitney D. The visual system discounts emotional deviants when extracting average expression. *Attention, Perception, & Psychophysics*. 2010; 72:1825–1838.
15. Im HY, Halberda J. The effects of sampling and internal noise on the representation of ensemble average size. *Attention, Perception, & Psychophysics*. 2013; 75:278–286.
16. Haberman J, Whitney D. Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*. 2007; 17:R751–R753. [PubMed: 17803921]
17. Hubert-Wallander B, Boynton GM. Not all summary statistics are made equal: Evidence from extracting summaries across time. *Journal of Vision*. 2015; 15:1–12.
18. Ji L, Chen W, Fu X. Different roles of foveal and extrafoveal vision in ensemble representation for facial expressions. *Engineering Psychology and Cognitive Ergonomics Lecture Notes in Computer Science*. 2014; 8532:164–173.
19. Yang JW, Yoon KL, Chong SC, Oh KJ. Accurate but pathological: Social anxiety and ensemble coding of emotion. *Cognitive Therapy and Research*. 2013; 37:572–578.
20. De Fockert JW, Wolfenstein C. Rapid extraction of mean identity from sets of faces. *Quarterly Journal of Experimental Psychology*. 2009; 62:1716–1722.
21. Leib AY, Puri AM, Fischer J, Bentin S, Whitney D, Robertson L. Crowd perception in prosopagnosia. *Neuropsychologia*. 2012; 50:1698–1707. [PubMed: 22521873]
22. Leib AY, Fischer J, Liu Y, Qiu S, Robertson L, Whitney D. Ensemble crowd perception: A viewpoint-invariant mechanism to represent average crowd identity. *Journal of Vision*. 2014; 14:1–13.
23. Neumann MF, Schweinberger SR, Burton AM. Viewers extract mean and individual identity from sets of famous faces. *Cognition*. 2013; 128:56–63. [PubMed: 23587844]
24. Brunyé TT, Howe JL, Mahoney CR. Seeing the crowd for the bomber: Spontaneous threat perception from static and randomly moving crowd simulations. *Journal of Experimental Psychology: Applied*. 2014; 20:303–322. [PubMed: 25347410]
25. Sweeny TD, Haroz S, Whitney D. Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance*. 2013; 39:329–337. [PubMed: 22708744]
26. Florey J, Clifford CW, Dakin S, Mareschal I. Spatial limitations in averaging social cues. *Scientific Reports*. 2016; 6:32210. [PubMed: 27573589]

27. Sweeny TD, Whitney D. Perceiving crowd attention: Ensemble perception of a crowd's gaze. *Psychological Science*. 2014; 25:1903–1913. [PubMed: 25125428]
28. Adams RB Jr, Ambady N, Macrae CN, Kleck RE. Emotional expressions forecast approach-avoidance behavior. *Motivation & Emotion*. 2006; 30:179–188.
29. Horstmann G. What do facial expressions convey: Feeling states, behavioral intentions, or action requests? *Emotion*. 2003; 3:150–166. [PubMed: 12899416]
30. Marsh AA, Ambady N, Kleck RE. The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion*. 2005; 5:119–124. [PubMed: 15755225]
31. Davidson RJ. Anterior cerebral asymmetry and the nature of emotion. *Brain & Cognition*. 1992; 20:125–151. [PubMed: 1389117]
32. Elliot AJ. Approach and avoidance motivation and achievement goals. *Educational Psychologist*. 1999; 34:169–189.
33. Adams RB Jr, Hess U, Kleck RE. The Intersection of Gender-Related Facial Appearance and Facial Displays of Emotion. *Emotion Review*. 2015; 7:5–13.
34. Bargh JA, Chen M, Burrows L. Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality & Social Psychology*. 1996; 71:230–244. [PubMed: 8765481]
35. Davidson RJ, Irwin W. The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*. 1999; 3:11–21. [PubMed: 10234222]
36. De Renzi E. Prosopagnosia in two patients with CT scan evidence of damage confined to the right hemisphere. *Neuropsychologia*. 1986; 24:385–389. [PubMed: 3736820]
37. Fabes RA, Martin CL. Gender and age stereotypes of emotionality. *Personality and Social Psychology Bulletin*. 1991; 17:532–540.
38. Kanwisher N, McDermott J, Chun MM. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*. 1997; 17:4302–4311. [PubMed: 9151747]
39. Wada Y, Yamamoto T. Selective impairment of facial recognition due to a haematoma restricted to the right fusiform and lateral occipital region. *Journal of Neurology, Neurosurgery, and Psychiatry*. 2001; 71:254–257.
40. Merigan WH, Maunsell JH. How parallel are the primate visual pathways? *Annual Review of Neuroscience*. 1993; 16:369–402.
41. Sawatari A, Callaway EM. Convergence of magno- and parvocellular pathways in layer 4B of macaque primary visual cortex. *Nature*. 1996; 380:442–446. [PubMed: 8602243]
42. Freud E, Plaut DC, Behrmann M. 'What' is happening in the dorsal visual pathway. *Trends in Cognitive Sciences*. 2016; 20:773–784. [PubMed: 27615805]
43. Goodale MA, Milner AD. Separate visual pathways for perception and action. *Trends in Neurosciences*. 1992; 15:20–25. [PubMed: 1374953]
44. Kveraga K, Boshyan J, Bar M. The magnocellular trigger of top-down facilitation in object recognition. *Journal of Neuroscience*. 2007; 27:13232–13240. [PubMed: 18045917]
45. Kveraga K, Ghuman AS, Bar M. Top-down predictions in the cognitive brain. *Brain and Cognition*. 2007; 65:145–168. [PubMed: 17923222]
46. Livingstone MS, Hubel DE. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*. 1988; 240:740–749. [PubMed: 3283936]
47. Milner AD, Goodale MA. *The Visual Brain in Action*. Oxford University Press; Oxford: 1995.
48. Milner AD, Goodale MA. Two visual systems re-viewed. *Neuropsychologia*. 2008; 46:774–785. [PubMed: 18037456]
49. Schiller PH, Logothetis NK. The color-opponent and broad-band channels of the primate visual system. *Trends in Neuroscience*. 1990; 13:392–398.
50. Thomas C, Kveraga K, Huberle E, Karnath HO, Bar M. Enabling global processing in simultanagnosia by psychophysical biasing of visual pathways. *Brain*. 2012; 135:1578–1585. [PubMed: 22418740]

51. Vuilleumier P, Armony JL, Driver J, Dolan RJ. Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*. 2003; 6:624–631. [PubMed: 12740580]
52. Winston JS, Vuilleumier P, Dolan RJ. Effects of low-spatial frequency components of fearful faces on fusiform cortex activity. *Current Biology*. 2003; 13:1824–1829. [PubMed: 14561410]
53. Courtney SM, Ungerleider LG, Keil K, Haxby JV. Object and spatial visual working memory activate separate neural systems in human cortex. *Cerebral Cortex*. 1996; 6:39–49. [PubMed: 8670637]
54. Courtney SM, Petit L, Maisog JM, Ungerleider LG, Haxby JV. An area specialized for spatial working memory in human frontal cortex. *Science*. 1998; 279:1347–1351. [PubMed: 9478894]
55. Sala JB, Courtney SM. Binding of what and where during working memory maintenance. *Cortex*. 2007; 43:5–21. [PubMed: 17334204]
56. Takahashi E, Ohki K, Kim DS. Dissociation and convergence of the dorsal and ventral visual streams in the human prefrontal cortex. *Neuroimage*. 2013; 65:488–498. [PubMed: 23063444]
57. Wilson FA, Scalaidhe SP, Goldman-Rakic PS. Dissociation of object and spatial processing domains in primate prefrontal cortex. *Science*. 1993; 260:1955–1958. [PubMed: 8316836]
58. Denys K, Vanduffel W, Fize D, Nelissen K, Peuskens H, Van Essen D, Orban GA. The processing of visual shape in the cerebral cortex of human and nonhuman primates: a functional magnetic resonance imaging study. *Journal of Neuroscience*. 2004; 24:2551–2565. [PubMed: 15014131]
59. Grill-Spector K, Malach R. The human visual cortex. *Annual Review of Neuroscience*. 2004; 27:649–677.
60. Haxby JV, Grady CL, Horwitz B, et al. Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Science, USA*. 1991; 88:1621–1625.
61. Purves, D., Augustine, GJ., Fitzpatrick, D., Hall, WC., LaMantia, A-S., McNamara, JO., Williams, SM. *Neuroscience*. 3. Massachusetts U.S.A: Sinauer Associates Inc; 2004.
62. Taylor JC, Wiggett AJ, Downing PE. Functional MRI analysis of body and body part representations in extrastriate and fusiform body parts. *Journal of Neurophysiology*. 2007; 98:1626–1633. [PubMed: 17596425]
63. LeDoux, JE. *The Emotional Brain*. Simon and Schuster; New York: 1996.
64. Morris JS, Frith CD, Perrett DI, Rowland D, Young AW, Calder AJ, Dolan RJ. A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*. 1996; 383:812–815. [PubMed: 8893004]
65. Damasio AR, Grabowski TJ, Bechara A, Damasio H, Ponto LL, Parvizi J, Hichwa RD. Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*. 2000; 3:1049–1056. [PubMed: 11017179]
66. Vuilleumier P, Armony JL, Driver J, Dolan RJ. Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron*. 2001; 30:829–841. [PubMed: 11430815]
67. Zald DH. The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Brain Research Reviews*. 2003; 41:88–123. [PubMed: 12505650]
68. Whalen PJ, Kagan J, Cook RG, Davis FC, Kim H, Polis S, McLaren DG, Somerville LH, McLean AA, Maxwell JS, Johnstone T. Human amygdala responsivity to masked fearful eye whites. *Science*. 2004; 306:2061. [PubMed: 15604401]
69. Ekman, P., Friesen, WV. *Pictures of facial affect*. Consulting Psychologists Press; Palo Alto, CA: 1976.
70. Botvinick MM, Cohen JD, Carter CS. Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*. 2004; 8:539–545. [PubMed: 15556023]
71. Jonides J, Nee DE. Brain mechanisms of proactive interference in working memory. *Neuroscience*. 2006; 139:181–193. [PubMed: 16337090]
72. Monsell S. Task switching. *Trends in Cognitive Sciences*. 2003; 7:134–140. [PubMed: 12639695]
73. Rogers RD, Monsell S. The costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*. 1995; 124:207–231.

74. Theios, J. The components in response latency in simple human information processing tasks. In: Rabbitt, PMA., Dornic, S., editors. Attention and performance V. New York: Academic Press; 1975. p. 418-440.
75. Christie J, Ginsberg JP, Steedman J, Fridriksson J, Bonilha L, Rorden C. Global versus local processing: seeing the left side of the forest and the right side of the trees. *Frontiers in Human Neuroscience*. 2012; 6:28. [PubMed: 22371700]
76. Delis DC, Robertson LC, Efron R. Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia*. 1986; 24:205–214. [PubMed: 3714025]
77. Robertson LC, Lamb MR, Knight RT. Effects of lesions of temporal-parietal junction on perceptual and attentional processing in humans. *Journal of Neuroscience*. 1988; 8:3757–3769. [PubMed: 3193178]
78. Robertson LC, Ivry R. Hemispheric asymmetry: attention to visual and auditory primitives. *Current Directions in Psychological Science*. 2000; 9:59–63.
79. Yovel G, Levy J, Yovel I. Hemispheric asymmetries for global and local visual perception: effects of stimulus and task factors. *Journal of Experimental Psychology: Human Perception and Performance*. 2001; 27:1369–1385. [PubMed: 11766931]
80. Becker DV, Kenrick DT, Neuberg SL, Blackwell KC, Smith DM. The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*. 2007; 92:179–190. [PubMed: 17279844]
81. Borod JC, Cicero BA, Obler LK, Welkowitz J, Erhan HM, Santschi C, Grunwald IS, Agosti RM, Whalen JR. Right hemisphere emotional perception: evidence across multiple channels. *Neuropsychology*. 1998; 12:446–458. [PubMed: 9673999]
82. Silberman EK, Weingartner H. Hemispheric lateralization of functions related to emotion. *Brain and Cognition*. 1986; 5:322–353. [PubMed: 3530287]
83. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automatic anatomical labelling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*. 2002; 15:273–289. [PubMed: 11771995]
84. Denison RN, Vu AT, Yacoub E, Feinberg DA, Silver MA. Functional mapping of the magnocellular and parvocellular subdivisions of human LGN. *Neuroimage*. 2014; 102:358–369. [PubMed: 25038435]
85. Adams RB Jr, Im HY, Cushing C, Ward N, Boshyan J, Steiner T, Albohn D, Kveraga K. Compound facial threat cue perception: Contributions of visual pathways, aging, and anxiety. *Journal of Vision*. 2016; 16:1375.
86. Im, HY., Adams, RB., Jr, Boshyan, J., Ward, N., Cushing, C., Kveraga, K. Anxiety modulates perception of facial fear in a pathway-specific, lateralized manner. under review doi: <https://doi.org/10.1101/141838>
87. Kveraga, K. Threat perception in visual scenes: dimensions, action and neural dynamics. In: Kveraga, K., Bar, M., editors. Scene Vision: making sense of what we see. Cambridge: MIT Press; 2014. p. 291-307.
88. Erk S, Kleczar A, Walter H. Valence-specific regulation effects in a working memory task with emotional context. *Neuroimage*. 2007; 37:623–632. [PubMed: 17570686]
89. Heller W, Nitscke JB. Regional brain activity in emotion: A framework for understanding cognition in depression. *Cognition & Emotion*. 1997; 11:637–661.
90. Adams RB Jr, Kveraga K. Social vision: Functional forecasting and the integration of compound social cues. *Review of Philosophy and Psychology*. 2015; 6:591–610.
91. Rogers, LJ., Vallortigara, G., Andrew, RJ. Divided brains: The biology and behavior of brain asymmetries. Cambridge, UK: Cambridge University Press; 2013.
92. Scott WA. Cognitive complexity and cognitive flexibility. *American Sociological Association*. 1962; 25:405–414.
93. Craig AD. Forebrain emotional asymmetry: a neuroanatomical basis? *Trends in Cognitive Sciences*. 2005; 9:566–571. [PubMed: 16275155]
94. Davidson, RJ. Brain Asymmetry, Cerebral asymmetry, emotion, and affective style. Davidson, RJ., Hugdahl, K., editors. MIT Press; Cambridge, MA: 1995. p. 361-387.

95. Ivry, RB., Robertson, LC. The two sides of perception. Cambridge (MA): MIT Press; 1998.
96. Carroll NC, Young AW. Priming of emotion recognition. *Quarterly Journal of Experimental Psychology*. 2005; 58:1173–1197. [PubMed: 16194954]
97. Norman DA, Shallice T. Attention to action: willed and automatic control of behavior. *Cognitive neuroscience: A reader*. 2000:376–390.
98. Posner MI, Rueda MR, Kanske P. The handbook of psychophysiology. 2007; Ch. 18:410–414.
99. Cant JS, Sun SZ, Xu Y. Distinct cognitive mechanisms involved in the processing of single objects and object ensembles. *Journal of Vision*. 2015; 15:1–21.
100. Chong SC, Joo SJ, Emmanouil T, Treisman A. Statistical processing: Not so implausible after all. *Perception & Psychophysics*. 2008; 70:1327–1334. [PubMed: 18927015]
101. Chong SC, Evans KK. Distributed vs. Focused attention (count vs. estimate). *Wiley Interdisciplinary Reviews: Cognitive Science*. 2011; 2:634–638. [PubMed: 22121459]
102. Chong SC, Treisman A. Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*. 2005; 67:1–13. [PubMed: 15912869]
103. Haberman J, Brady TF, Alvarez GA. Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*. 2015; 144:432–446. [PubMed: 25844624]
104. Myczek K, Simons DJ. Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception & Psychophysics*. 2008; 70:772–788. [PubMed: 18613626]
105. Cant JS, Xu Y. The impact of density and ratio on object-ensemble representation in human anterior-medial ventral visual cortex. *Cerebral Cortex*. 2015; 25:4226–4239. [PubMed: 24964917]
106. Cant JS, Xu Y. The contribution of object shape and surface properties to object-ensemble representation in anterior-medial ventral visual cortex. *Journal of Cognitive Neuroscience*. 2017; 29:398–412. [PubMed: 27676614]
107. Huis in 't Veld EMJ, de Gelder B. From individual fear to mass panic. The neurological basis of crowd perception. *Human Brain Mapping*. 2015; 36:2338–1351. [PubMed: 25716010]
108. Utochkin IS. Ensemble summary statistics as a basis for rapid visual categorization. *Journal of Vision*. 2015; 15:1–14.
109. Brady TF, Alvarez GA. No evidence for a fixed object limit in working memory: Ensemble representations inflate estimates of working memory capacity for complex objects. *Journal of Experimental Psychology: Learning, Memory and Cognition*. 2015; 41:921–929.
110. Cohen MA, Dennett DC, Kanwisher N. What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences*. 2016; 19:324–335.
111. Feigenson, L. Objects, sets, and ensemble. In: Dehaene, S., Brannon, E., editors. *Space, Time, and Number in the Brain: searching for the Foundations of Mathematical Thought*. London: Elsevier; 2011.
112. Im HY, Chong SC. Mean size as a unit of visual working memory. *Perception*. 2014; 43:663–676. [PubMed: 25223110]
113. Im HY, Park WJ, Chong SC. Ensemble statistics as units of selection. *Journal of Cognitive Psychology*. 2015; 27:114–127.
114. Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*. 2001; 24:87–185. [PubMed: 11515286]
115. Luck S, Vogel EK. The capacity of visual working memory for features and conjunctions. *Nature*. 1997; 390:279–281. [PubMed: 9384378]
116. Pylyshyn ZW, Storm RW. Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*. 1988; 3:179–197. [PubMed: 3153671]
117. Treisman AM, Gelade G. A feature-integration theory of attention. *Cognitive Psychology*. 1980; 12:97–136. [PubMed: 7351125]
118. Haberman, J., Whitney, D. Ensemble Perception: Summarizing the scene and broadening the limits of visual processing. In: Wolfe, J., Robertson, L., editors. *From Perception to Consciousness: Searching with Anne Treisman*. Oxford University Press; 2012.

119. Brady TF, Alvarez GA. Hierarchical encoding in visual working memory: ensemble statistics bias memory for individual items. *Psychological Science*. 2011; 22:384–392. [PubMed: 21296808]
120. Cohen, J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: LEA; 1988.
121. Brainard DH. The Psychophysics Toolbox. *Spatial vision*. 1997; 10:433–436. [PubMed: 9176952]
122. Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial vision*. 1997; 10:437–442. [PubMed: 9176953]
123. Maule J, Franklin A. Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of Vision*. 2015; 15:1–18.
124. Utochkin IS, Tiurina NA. Parallel averaging of size is possible but range-limited: a reply to Marchant, Simons, and De Fockert. *Acta Psychologica*. 2014; 146:7–18. [PubMed: 24361740]
125. Fitts PM, Seeger CM. S-R compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*. 1953; 46:199–201. [PubMed: 13084867]
126. Deichmann R, Gottfried JA, Hutton C, Turner R. Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage*. 2003; 19:430–441. [PubMed: 12814592]
127. Wall MB, Walker R, Smith AT. Functional imaging of the human superior colliculus: an optimised approach. *Neuroimage*. 2009; 47:1620–1627. [PubMed: 19505584]
128. Kringelbach ML, Rolls ET. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*. 2004; 72:341–372. [PubMed: 15157726]
129. Mazaika, PK., Hoefft, F., Glover, GH., Reiss, AL. Methods and software for fMRI analysis for clinical subjects. Poster Session Presented at the Meeting of Human Brain Mapping; 2009; San Francisco, CA. 2009.
130. Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM. Automatically parcellating the human cerebral cortex. *Cerebral Cortex*. 2004; 14:11–22. [PubMed: 14654453]
131. Spiridon M, Fischl B, Kanwisher N. Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping*. 2006; 27:77–89. [PubMed: 15966002]
132. Yang Q, Wang X, Yin S, Zhao X, Tan J, Chen A. Improved emotional conflict control triggered by the processing priority of negative emotion. *Scientific Reports*. 2016; 6:24302. [PubMed: 27086908]

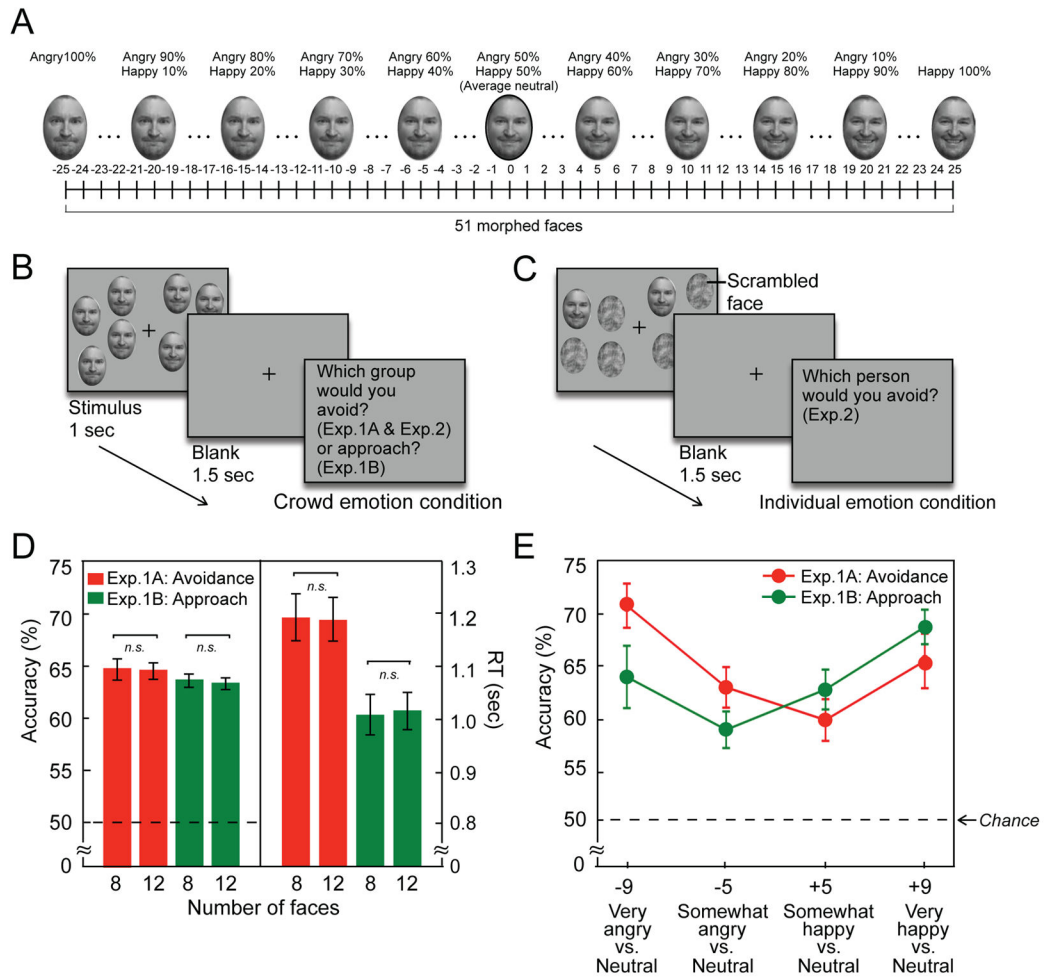


Figure 1. Sample face images, sample trials of crowd emotion and individual emotion conditions, and the results from Experiment 1

(A) Some examples of 51 morphed faces from two extremely angry and happy faces of the same person, with Face -25 in Emotional Unit (EU) being extremely angry, Face 0 being neutral, and Face +25 being extremely happy. (B) A sample trial of crowd emotion condition. (C) A sample trial of individual emotion condition (included in the fMRI study). (D) The effect of the number of faces on the accuracy and RT in Experiment 1A (avoidance task, red bars) and in Experiment 1B (approach task, green bars). The error bars indicate the standard error of the mean (SEM). (E) The effect of the similarity in average emotion between facial crowds on crowd emotion processing: Participants' accuracies on Experiment 1A (avoidance task, red line) and Experiment 1B (approach task, green line) are plotted as a function of the emotional distance in EU between two facial crowds to be compared.

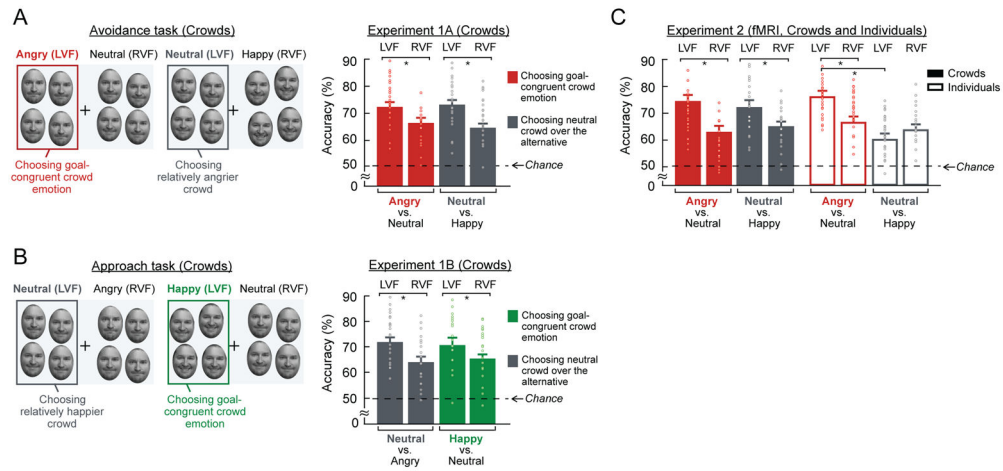


Figure 2. Right hemisphere dominance for goal-relevant crowd emotion in the crowd emotion processing

(A) Participants’ accuracy for the avoidance task (Experiment 1A), separately plotted for when the crowd to be chosen (an angry crowd when compared to a neutral and a neutral crowd when compared to a happy crowd) is presented in the LVF vs. RVF. Because the correct answer was to choose a “relatively angrier” crowd than the other, participants had to choose an angry over a neutral crowd and a neutral over a happy crowd. Participants’ accuracy was greater for the LVF than the RVF, regardless of the actual emotional valence of the crowd to be chosen (both an angry and a neutral crowd). (B) Participants’ accuracy for the approach task (Experiment 1B), separately plotted for when the crowd to be chosen (a happy crowd when compared to a neutral and a neutral crowd when compared to an angry crowd) is presented in the LVF vs. RVF. Note that the valence of the goal-relevant crowd emotion is switched from angry to happy in the approach task. Participants’ accuracy was greater for the LVF than the RVF, regardless of the actual emotional valence of the crowd to be chosen (both a happy and a neutral crowd). (C) Participants’ accuracy for Experiment 2 (fMRI study). Accuracies both for crowd emotion and individual emotion conditions are plotted for the LVF and RVF, separately. As in Experiment 1, participants’ accuracy for crowd emotion condition was greater for LVF than RVF both when they had to choose an angry over a neutral crowd and a neutral over a happy crowd to avoid. However, patterns were different for individual condition: Accuracies were greater for an angry face in the LVF and a neutral face in the RVF. The error bars indicate SEM. Points in (A)-(C) represent data from individual participants.

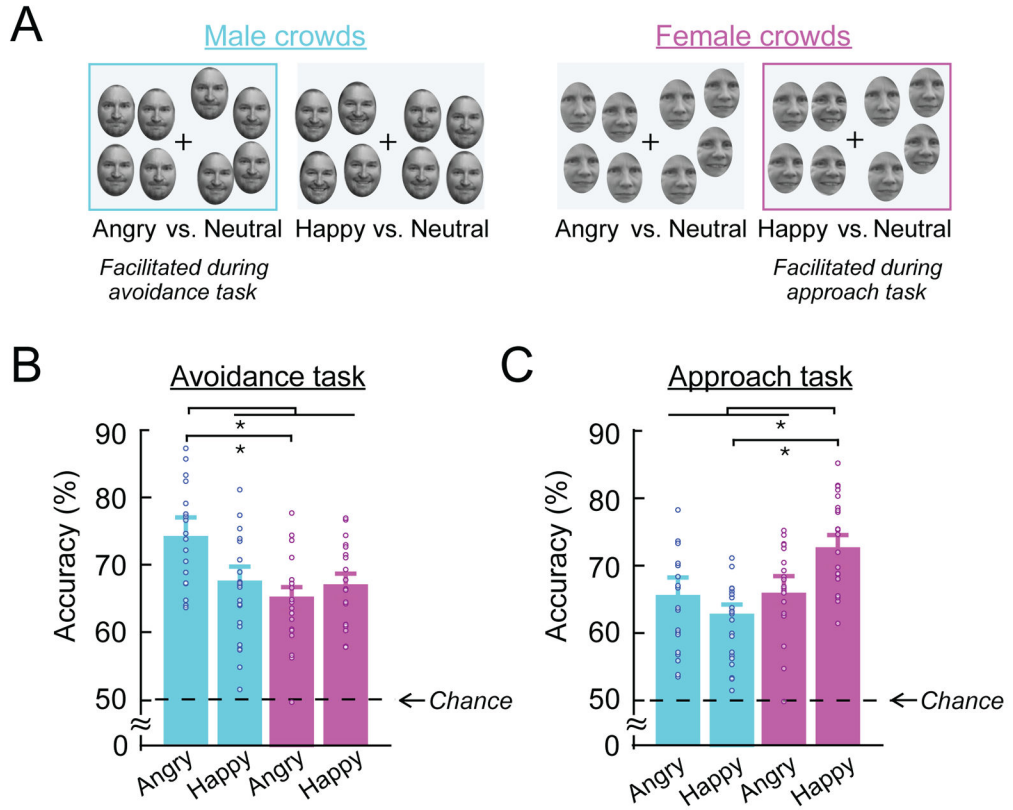


Figure 3. The effect of the sex-specific identity cue of facial crowds on crowd emotion perception (A) Sample crowd stimuli for male (in cyan) and female (in magenta) crowds. (B) Participants’ accuracy for the avoidance task (Experiment 1A) for sex of facial crowds (male crowds vs. female crowds) and for the emotional valence of an emotional crowd (Angry vs. Happy). Angry male crowds were identified most accurately in the avoidance task. (C) Participants’ accuracy for the approach task (Experiment 1B). In the approach task, happy female crowds were identified most accurately. The error bars indicate SEM. Points in (B)-(C) represent data from individual participants.

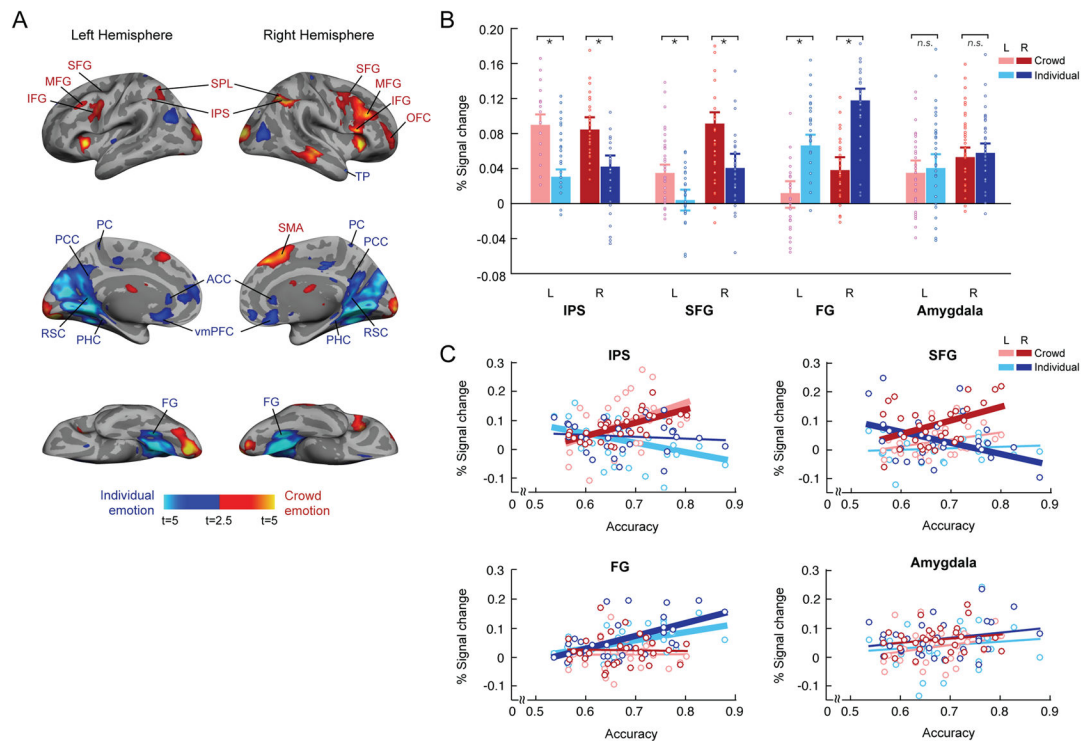


Figure 4. Distinct neural pathways preferentially involved in dorsal and ventral visual pathways for crowd emotion and individual emotion processing, respectively

(A) The brain areas that showed greater activation when participants were making avoidance decision by comparing two crowds are shown in red and the brain areas that showed greater activation for comparing two single faces are shown in blue. The activations were thresholded at $p < 0.05$, FWE-corrected. The complete list of activations is labeled and reported in the Table 1, and the results of the same contrasts at different thresholds are reported in the Supplementary Tables 1 and 2. (B) The percent signal change of our ROI's (bilateral IPS, SFG, FG, and amygdala) when participants were making avoidance decision by comparing two crowds (pink and red bars for the left and right hemispheres, respectively) and two single faces (light and dark blue bars for the left and right hemispheres, respectively). Points represent data from individual participants. (C) The correlation between the percent signal change and the participants' accuracy for crowd emotion condition (pink and red dots for the left and right hemispheres, respectively) and for individual emotion condition (in light and dark blue dots for the left and right hemispheres, respectively), with overlaid linear regression lines with same colors. Thick regression lines indicate statistically significant correlation (at $p < 0.05$) between the individual participants' accuracy and the percent signal change of each ROI.

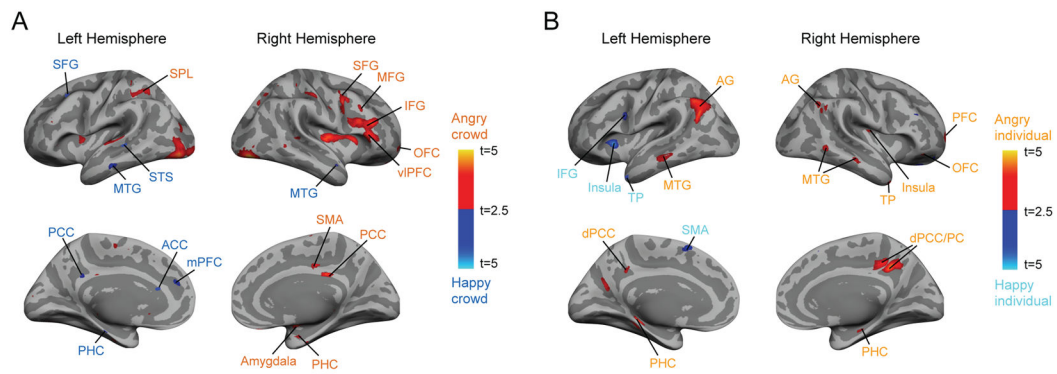


Figure 5. Brain activations by different types of avoidance comparisons: an angry vs. a neutral crowd (A) or face (B) vs. a neutral vs. a happy crowd (A) or face (B)

(A) The brain areas that showed greater activation when participants were making avoidance decision by comparing an angry crowd vs. a neutral crowd (shown in orange labels) and the brain areas that showed greater activation for comparing a neutral crowd vs. a happy crowd (shown in blue labels). (B) The brain areas that showed greater activation when participants were making avoidance decision by comparing an angry face vs. a neutral face (shown in yellow labels) and the brain areas that showed greater activation for comparing a neutral crowd vs. a happy face (shown in cyan labels).

Table 1

Significantly activated areas in mean response for [1] Crowd emotion minus Individual emotion and [2] Individual emotion minus Crowd emotion contrasts. The activations were thresholded at $p < 0.05$, FWE-corrected. – indicates that this cluster is part of a larger cluster immediately above.

Activation location	MNI Coordinates			t-value	Extent
	x	y	z		
Crowd emotion > Individual emotion					
L Visual cortex (BA18)	-15	-88	-14	7.281	1210
	-21	-97	14	4.998	-
L Visual cortex (BA19)	-33	-79	-16	3.888	-
R Visual cortex (BA18)	18	-100	12	5.858	666
R Superior frontal gyrus	33	8	64	5.686	3428
R Middle frontal gyrus	36	32	46	6.267	-
L Superior frontal gyrus	-24	9	64	3.112	46
L Middle frontal gyrus	-42	26	38	3.067	380
	-42	50	26	3.024	91
	-42	56	7	3.098	-
R Anterior Insula	30	26	2	5.864	3428
L Anterior Insula	-33	20	4	5.858	176
R Intraparietal sulcus	39	-58	48	5.607	967
R Superior parietal lobule	36	-64	52	5.051	-
L Intraparietal sulcus	-39	-49	38	4.677	368
R Superior temporal sulcus	48	-25	-10	4.709	404
L Superior temporal sulcus	-54	-43	-8	3.303	26
R Supramarginal gyrus	54	-46	32	2.556	967
L Supramarginal gyrus	-51	-52	56	2.774	368
	-51	-43	3	2.928	9
R Cerebellum	33	-52	-44	2.876	15
L Cerebellum	-33	-55	-42	3.901	178
R Supplementary Motor Area	9	32	48	5.000	511
L Superior parietal lobule	-30	-67	48	3.477	368
L Premotor cortex	-39	5	36	4.364	380

Activation location	MINI Coordinates			t-value	Extent
	x	y	z		
Crowd emotion > Individual emotion					
L Inferior frontal gyrus	-27	26	22	3.586	21
R Caudate	12	20	6	3.046	18
R brainstem	9	-16	-8	2.804	13
Individual emotion > Crowd emotion					
R Visual cortex (BA19)	12	-64	-8	8.054	5336
R Fusiform gyrus	27	-55	-14	5.998	-
R Parahippocampal cortex	21	-43	-10	4.943	-
R Retrosplenial cortex	18	-61	12	4.891	-
L Visual cortex (BA19)	-18	-55	-6	6.742	5336
L Fusiform gyrus	-27	-46	-12	5.699	-
L Parahippocampal cortex	-27	-43	-10	5.623	-
L Primary visual cortex (BA17)	-6	-73	6	6.421	-
L Retrosplenial cortex	-18	-55	4	5.281	-
L Posterior cingulate cortex	-9	-52	32	-2.858	-
R Posterior cingulate cortex	12	-25	42	4.384	91
R Temporal pole	51	17	-39	4.564	34
	27	14	-32	2.849	8
	42	8	-28	3.239	9
L Precuneus	-12	-43	60	3.258	71
	-6	-61	66	2.796	8
R Precuneus	6	-52	64	2.863	26
R Visual cortex (BA19)	48	-79	16	4.246	272
L Ventromedial prefrontal cortex	-18	14	-22	3.581	23
R Ventromedial prefrontal cortex	9	26	-12	4.031	406
	9	47	-8	3.452	374
L Middle cingulate cortex	-12	-28	40	3.222	65
R Middle cingulate cortex	6	-13	44	2.992	15
Anterior cingulate cortex	0	32	6	3.446	406

Activation location	MINI Coordinates			t-value	Extent
	x	y	z		
Crowd emotion > Individual emotion					
L Angular gyrus	-42	-67	24	3.594	228
	-18	14	-22	3.336	23
L Posterior Insula	-39	-16	0	3.346	28
L Insula	-42	2	6	2.994	11
R Posterior Insula	44	-4	-8	3.172	58
L Medial prefrontal cortex	-12	47	40	2.731	374
L Supramarginal gyrus	-57	-37	28	3.251	40
R Hippocampus	27	-13	-22	2.871	16
L Orbitofrontal cortex	-45	29	-8	2.804	6