



Differential neurodynamics and connectivity in the dorsal and ventral visual pathways during perception of emotional crowds and individuals: a MEG study

Hee Yeon Im^{1,2,3} · Cody A. Cushing^{2,4} · Noreen Ward² · Kestutis Kveraga^{2,3}

Accepted: 3 February 2021
© The Psychonomic Society, Inc. 2021

Abstract

Reading the prevailing emotion of groups of people (“crowd emotion”) is critical to understanding their overall intention and disposition. It alerts us to potential dangers, such as angry mobs or panicked crowds, giving us time to escape. A critical aspect of processing crowd emotion is that it must occur rapidly, because delays often are costly. Although knowing the timing of neural events is crucial for understanding how the brain guides behaviors using coherent signals from a glimpse of multiple faces, this information is currently lacking in the literature on face ensemble coding. Therefore, we used magnetoencephalography to examine the neurodynamics in the dorsal and ventral visual streams and the periamygdaloid cortex to compare perception of groups of faces versus individual faces. Forty-six participants compared two groups of four faces or two individual faces with varying emotional expressions and chose which group or individual they would avoid. We found that the dorsal stream was activated as early as 68 msec after the onset of stimuli containing groups of faces. In contrast, the ventral stream was activated later and predominantly for individual face stimuli. The latencies of the dorsal stream activation peaks correlated with participants’ response times for facial crowds. We also found enhanced connectivity earlier between the periamygdaloid cortex and the dorsal stream regions for crowd emotion perception. Our findings suggest that ensemble coding of facial crowds proceeds rapidly and in parallel by engaging the dorsal stream to mediate adaptive social behaviors, via a distinct route from single face perception.

Keywords Ensemble coding · Crowd emotion · Magnetoencephalography · Face perception · Affective processing

Introduction

We routinely encounter groups of people at work, school, social gatherings, and now on group communication platforms on the Internet. Reading the overall mood of the group from facial expressions is important for guiding our attitude and responses toward them. For example, reading the intention of violence from the facial expressions of a crowd on the

street allows us to sense potential danger and escape the situation. Reading the mood and receptiveness of an audience during presentations allows us to tailor our own behavior, perhaps by explaining something in more detail in a lecture or adopting an affective state suitable to the group mood. Such assessment of crowd emotion must be rapid and flexible to support dynamic interactions (Elias, Dyer, & Sweeny, 2017; Haberman & Whitney, 2007), because there is rarely enough time to serially analyze each person’s facial expression in many social situations. One way to achieve this level of speed and efficiency is to represent the groups of faces as a higher-level description in the form of ensemble statistics. Ensemble statistics can be extracted from different feature dimensions as an abstract and global description of groups of multiple items, such as average (Alvarez & Oliva, 2008; Ariely, 2001; Chong & Treisman, 2003; Dakin & Watt, 1997; Im, Park, & Chong, 2015; Im, Tiurina, & Utochkin, 2020; Maule & Franklin, 2015), approximate numerosity (Burr & Ross, 2008; Chong & Evans, 2011; Gallistel & Gelman, 2000; Halberda, Sires, & Feigenson, 2016; Im, Zhong, & Halberda, 2016; Utochkin &

✉ Kestutis Kveraga
kestas@nmr.mgh.harvard.edu

¹ Department of Psychology, University of British Columbia, Vancouver, BC, Canada

² Athinoula A. Martinos Center, Department of Radiology, Massachusetts General Hospital, Charlestown, MA, USA

³ Department of Radiology, Harvard Medical School, Boston, MA, USA

⁴ Department of Psychology, University of California, Los Angeles, CA, USA

Vostrikov, 2017), or variance (Im & Halberda, 2013; Morgan, Chubb, & Solomon, 2008; Norman, Heywood, & Kentridge, 2015; Solomon, 2010). It has been shown that human observers are capable of perceiving ensembles from complex, real-world objects that engage in high-level processing, for example, average emotion or identity from a crowd of faces (Herman & Whitney, 2009; Im et al., 2017; Ji & Pourtois, 2018; Leib et al., 2014), group behavior from biological motion of humans (Sweeny, Haroz, & Whitney, 2012), or a crowd's attentional focus from people's average eye gaze direction (Sweeny & Whitney, 2014). Ensemble coding appears to be an efficient and powerful algorithm that the human brain can rely on to facilitate people perception and social cognition in general.

Although the efficiency of ensemble processing has been well documented (Alvarez & Oliva, 2009; Chong & Treisman, 2003; Herman, Harp, & Whitney, 2009; Im et al., 2016), and recently incorporated into several cognitive models (Cohen, Dennett, & Kanwisher, 2016; McClelland & Bayne, 2016; Rensink, 2000; Wolfe, Vö, Evans, & Greene, 2011), the neural mechanisms supporting this feat have only begun to be explored. In the domain of face processing, a behavioral study in prosopagnosia patients reported an intact ability to extract the average identity of a crowd of faces despite the prosopagnosics' impairments in single face recognition (Leib, Puri, Fischer, Bentin, Whitney, & Robertson, 2012); this suggested a possibility of distinct neural mechanisms underlying perception of crowd and individual faces. To identify neural substrates mediating ensemble coding of emotional faces, we recently conducted an fMRI study in healthy participants and found that different subsets of brain areas were preferentially activated during the processing of crowds of emotional faces and individual faces (Im et al., 2017). In this study, we presented either two crowds of people's faces or two single faces to participants and asked them to make snap judgments about which of the two crowds or individuals they would rather avoid. The participants' behavioral responses were equally precise and fast when they were presented with only two single faces versus crowds of many faces (up to 12 total), suggesting that ensemble coding of facial crowds (i.e., crowd emotion) was quite efficient and possibly benefited from parallel processing. We also found that the dorsal visual stream regions, including the posterior parietal cortex, showed greater activation for crowd emotion processing, whereas the ventral visual stream regions, including the posterior fusiform cortex, showed greater activation for individual emotion processing, suggesting differential contributions of the dorsal and ventral visual streams to the processing of crowd emotion and individual emotion.

Characterizing the specific contributions of dorsal and ventral pathways to the perception of crowd emotion and emotion of a single face is challenging. Given the dense structural and functional connectivity between the parietal and temporal cortex, it is difficult to attribute processing to one or the other

pathway independent of the other, particularly in fMRI with its slow BOLD signal. Because dorsal and ventral visual regions are heavily interconnected (Borra et al., 2008; Budisavljevic, Dell'Acqua, & Castiell, 2018; Catani, Jones, & Ffytche, 2005; Zhong & Rockland, 2003), it is also important to establish whether crowd emotion is projected to the dorsal stream initially, or is the result of earlier processing of the ventrotemporal visual areas, such as the fusiform face area (FFA) whose output is then projected to the dorsal regions. This can be done with MEG by examining the latencies of activation in these pathways to each condition, as well as using spectral measures of regional and interregional communication, such as phase-locking analyses (Lachaux, Rodriguez, Martinerie, & Varela, 1999). Thus, building on our prior work in fMRI (Im et al., 2017), the current study examined the fine-scale temporal profiles of neurodynamics in the dorsal and ventral streams during crowd emotion perception and individual emotion perception. We tested our hypothesis that rapid extraction of crowd emotion is mediated by the early engagement of the dorsal visual stream in processing global information (Peyrin et al., 2010; Schyns & Oliva, 1994; Thomas, Kveraga, Huberle, Karnath, & Bar, 2012), motivated by the greater activations of the dorsal stream regions for crowd emotion perception compared with individual emotion in the previous fMRI study (Im et al., 2017). Conversely, we expected processing in the individual face comparison condition to predominantly activate the ventral stream regions, including the fusiform face regions, as has been observed in many studies on face perception (Haxby et al., 1991; Im et al., 2017; Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997).

Perceiving emotional faces and executing appropriate responses to them involves the integration of the perception, cognition, and action stages, relying on recurrent interactions of a number of widely distributed brain regions (for reviews, see Dolcos, Iordan, & Dolcos, 2011; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012). Such large-scale integration and interaction could be mediated by groups of neurons that oscillate within a specific frequency range or create neural oscillation synchronies over brief time windows during interregional communications (Hipp, Engel, & Siegel, 2011; Lachaux et al., 1999; Sauseng & Klimesch, 2008). Thus, measures of neuronal phase synchrony can provide sensitive measurements that reflect the integration and exchange of information between brain networks (Ehlers, Wills, Desikan, Phillips, & Havstad, 2014; Simões, Jensen, Parkkonen, & Hari, 2003). Moreover, phase-locking in a region can provide useful estimates of variance across trial types that are not directly available in typical evoked response analyses, with the benefit of being more robust against potential artifacts and individual differences in component latencies (Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). Therefore, the current study focused on frequency-specific MEG

responses by estimating the phase synchrony of neural oscillations within and between brain regions of interest (ROIs). In our recent MEG studies (Cushing et al., 2018; Cushing, Im, Adams Jr., Ward, & Kveraga, 2019), we have observed differences in the strengths of α -band (8–13 Hz) versus β -band (14–30 Hz) oscillations during perception of a fearful face stimuli, with greater phase-locking in the α -band when the face's emotional expression (e.g., fear) was combined with an ambiguous facial cue (e.g., direct eye gaze) but with greater phase-locking in the β -band when the fearful face was combined with a clear facial cue (e.g., averted eye gaze). The latter study (Cushing et al., 2019) further showed that these synchronization differences became more robust in the α -band for faces presented to the parvocellular (P) pathway (which projects predominantly to the ventral visual stream) and in the beta-band for faces presented to the magnocellular (M) pathway (whose projections dominate the dorsal visual stream), suggesting that different social affective cues conveyed by a single face yield different spectrottemporal profiles of the processing streams. It is not yet clear whether these spectrottemporal differences are the result of the cell properties in the P and M pathways or the cortical regions to which they predominantly project. However, motivated by these findings, we nonetheless focused on examining the α - and β -frequency bands to test for spectral phase-locking differences in processing emotional crowds versus individual faces.

As our ROIs, we selected posterior fusiform cortex (pFusi), posterior superior temporal sulcus (pSTS), posterior parietal cortex (PPC), and periamygdaloid cortex (PAC). Given the well-established, critical roles of the pFusi (including FFA) in the ventral stream plays in face perception in general (Fox, Moon, Iaria, & Barton, 2009; Kanwisher et al., 1997; Kanwisher & Yovel, 2006; Rossion et al., 2003), the temporal profile of pFusi activation was examined to characterize the ventral stream processing that we predict to be more engaged by individual emotion processing, based on our fMRI findings using a similar study paradigm (Im et al., 2017). Unlike ventral stream areas, such as pFusi, specific roles of subregions of the dorsal stream in visual processing are only beginning to be understood. To select ROIs in the dorsal stream, we likewise relied on our prior fMRI results (Im et al., 2017) and examined three different subregions of the dorsal stream. In our fMRI study, we found that pSTS and PPC showed greater activations to emotional crowd stimuli, while pFusi showed greater activations to individual emotional faces. Finally, we also included PAC in our ROIs for the current study, because it is part of the amygdaloid complex, which is heavily involved in processing affective stimuli, such as emotional faces (Morris et al., 1996; Pessoa, 2010; Zald, 2003). While we cannot be sure whether activity from PAC is limited to the cortical layers of the amygdaloid complex or also includes some activity from the subcortical amygdalar nuclei, our MEG face perception studies (Cushing et al., 2018, 2019) as well as many

studies by others (Dumas, Dubal, Attal, et al., 2013; Styliadis, Ioannides, Bamidis, & Papadelis, 2014) reliably show activity in this region evoked by emotional faces. We thus tested how activations of PAC, the affective processing hub of the limbic brain, were differentially related to those of the brain areas in the dorsal stream (PPC) and the ventral stream (pFusi) during the social affective processing that involves an individual face and a crowd of faces.

Relying on the temporal profiles of our ROIs and the inter-region connectivity, we aimed to assess differential contributions of brain regions in the dorsal and ventral streams to the perception of crowd emotion and individual emotion. We hypothesized that if the perception of crowd emotion required serial processing of individuated faces relying on the same mechanism as that for individual face processing, then the response time for crowd emotion would proportionally increase as the number of faces to be processed increases. Moreover, phase-locking would peak later for ensemble processing compared with individual processing, possibly both in the dorsal and ventral pathways, given a greater number of faces to be processed. Alternatively, if the perception of crowd emotion instead relied on efficient ensemble coding that is distinct from single face processing, then participants' processing speed would not increase with the number of faces to be processed (2 vs. 8). Moreover, phase-locking results would show 1) earlier peaks for crowd emotion perception, given the fact that there are many more faces to be processed for the same duration of stimulus exposure, compared to individual emotion perception, and 2) stronger phase-locking in the dorsal versus ventral pathways for crowd versus individual emotion processing, respectively.

Methods

Participants Forty-six participants (19–41 years old; 34 females and 12 males) from the Massachusetts General Hospital (MGH) and surrounding communities participated in this study. Thirty-eight participants were white, three were black, and five were Asian. No subjects were excluded from behavioral data analysis. All had normal or corrected-to-normal visual acuity and normal color vision. Most of the participants (41 participants among 46) were right-handed, whereas five participants were left-handed. Participants were first screened via a questionnaire to make sure they were eligible for MEG recording and subsequent MRI structural scans and had no history of mental illness or use of psychoactive medication. Informed consent was obtained from the participants following the Declaration of Helsinki. The experimental protocol was approved by the Institutional Review Board of MGH. The participants were compensated with \$50 for their participation in this study.

Apparatus and stimuli We kept our experimental designs, settings, and parameters of the current study as similar to those in our previous fMRI work (Im et al., 2017) as we could. For face sets, we used six different identities (3 male and 3 female faces), taken from the Ekman face set (Ekman & Friesen, 1976). We first created a set of 51 faces with different emotional intensities, by morphing two highly intense facial expressions (happy and angry) of the same person using a face-morphing software (Norrkross MorphX). The morphed face images were controlled for luminance, and emotional expressions of the morphed faces ranged from 100% happy to 100% angry, with different proportions of morphing between the two extreme faces by linearly interpolating in 2% increments (Fig. 1a). The morphed face images were linearly interpolated such that the larger the separation between any two morphed faces was, the easier it was to discriminate the two faces.

Stimulus images were generated with MATLAB and Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). In each crowd stimulus (shown as red outlined boxes in Fig. 1b), four different morphed faces were chosen from the continuum of the 51 faces. The faces were randomly positioned in the invisible frame (subtending $4.42^\circ \times 6.76^\circ$) in each visual field (right and left) on a grey background. Therefore, the facial crowd stimuli comprised eight individual faces total. The distance between the proximal edges of the invisible frames in the left and right visual fields was 6.76° .

We ensured that the range of emotional intensities of faces in each set is the same 18 steps, because the previous studies suggested that the range of variation in individual elements is an important determinant for the ease of averaging process at least for other visual features (e.g., size or hue; Maule & Franklin, 2015; Utochkin, & Tiurina, 2014). This range was determined based on our previous work (Im et al., 2017) in which we observed that at this range 1) each face was distinguishable from one another within a set and 2) the task was not so easy as to produce accuracy ceiling effects (e.g., overall accuracy with approximately 65%). Using the same range also allowed us to compare our current behavioral results to the previous results.

Stimuli for the individual emotion trials (shown as blue outlined boxes in Fig. 1b) comprised one emotional face (either angrier or happier than neutral) and one neutral face (i.e., 50% and 50% morphing of extremely happy and angry faces) from the same set of morphed face images. Individual faces were randomly positioned in the invisible frame with the same size as that of the crowd stimuli in each visual field. The offsets in emotional intensities 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 with the individual emotion condition, we included scrambled faces in the individual emotion condition so that

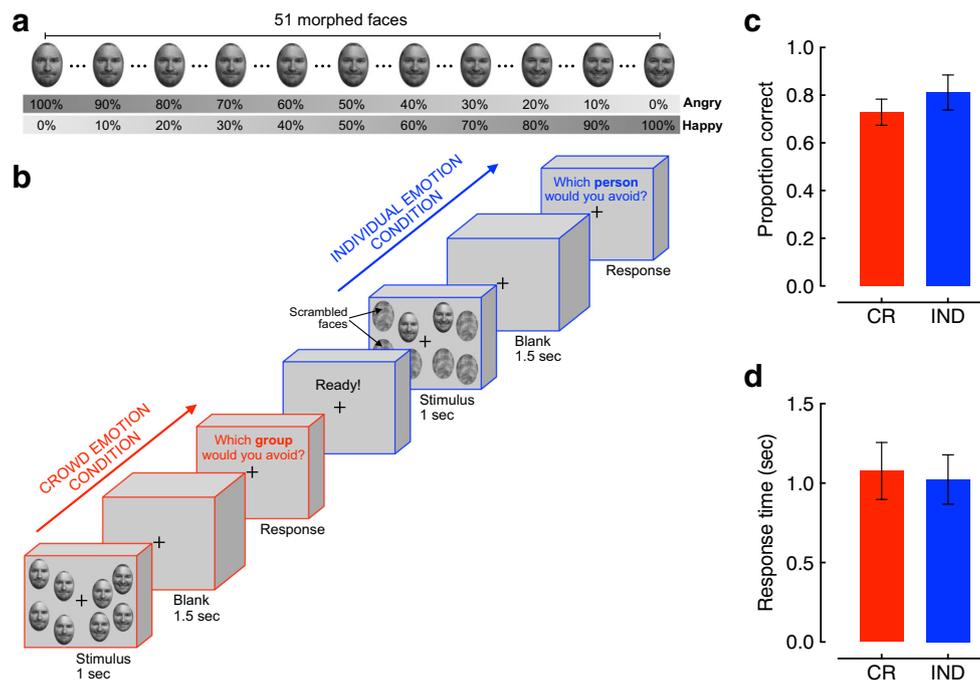


Fig. 1 Experimental designs and Behavioral results. **a** Examples of face stimuli morphed from two extreme angry and happy faces of the same person. **b** Sample sequence describing two experimental trials: one of crowd emotion condition (red cuboids), and the other of individual emotion condition (blue cuboids). Trials of crowd emotion and

individual emotion conditions were randomly intermixed. **c** Response accuracy for crowd emotion (red) and individual emotion (blue) conditions. **d** Response time for crowd emotion (red) and individual emotion (blue) conditions.

the same number of the face-like blobs were presented as in the crowd emotion condition.

For both trial types (crowd emotion and individual emotion trials), the emotional stimulus (i.e., happier or angrier than neutral) was presented in the left visual field, and the neutral stimulus (50% happy and 50% angry) was presented in the right visual field on one-half of the trials. It was switched for the other half of the trials.

Task design Stimuli were rear-projected onto a translucent screen placed 160 cm from the seated participant to create a 61.5-cm × 38.5-cm display. From this viewing distance, 1 pixel corresponded to 0.013° of visual angle. Fig. 1b illustrates a sample sequence of two successive trials. A visual stimulus, either containing individual faces or crowds of faces, was presented for 1 second, followed by a grey blank screen for 1.5 seconds. After the blank screen, the participants were instructed to make a key-press using a button box connected to the MEG system 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. They pressed the “1” key using their left index finger for choosing the left visual field (LVF) and the “4” key using their right index finger for choosing the right visual field (RVF). The key-response assignment was not counterbalanced in order to maintain the consistent and spatially coregistered stimulus-response mapping (S-R mapping: LVF for the left key and RVF for the right key; Fitts & Seeger, 1953). The written instructions were provided at the beginning of the experiment to explicitly inform the participants that the correct answer was to choose either the crowd or the face showing a more negative (e.g., angrier) emotion between the two. Participants’ responses that were made after 2.5 seconds were considered late and excluded from analyses of manual accuracy and response time (RT). Feedback for correct, incorrect, or late responses was provided visually after each response. Before the test session, participants completed 20 practice trials to get familiarized with the task and the trial sequence.

MEG data acquisition Magnetoencephalogram recordings were obtained with a 306-channel Neuromag Vectorview whole-head system (Elekta Neuromag) with 204 planar gradiometers and 102 magnetometers. The MEG scanner was enclosed in a magnetically shielded room with a shielding factor of 250,000 at 1 Hz (ImedcoAG). Before the recording session, four head position indicator (HPI) electrodes were attached to four locations to monitor head position in the dewar: two on each side of the forehead (right below the hairline) and one behind and above each ear of each participant. We also collected digitizer data by pointing the three cardinal landmarks (nasion plus right and left pre-auricular points) for each participant using a Polhemus FastTrack 3D system within a head-coordinate frame. HPI positions were recorded

within this frame, and approximately 200 points on the scalp and the face were recorded for co-registration with structural MRIs of the participants. We also recorded participants’ cardiac activity via ECG using two electrodes attached to the left and right chest and their eye movements and blinks using 4 EOG electrodes: 2 vertical electrodes (one placed just above the eyebrow, the other on the upper cheekbone just below the eye) and 2 horizontal electrodes placed on either side of the head between the eye and hairline. All data from MEG sensors and EOG and ECG electrodes were sampled at 600 Hz and were band-pass filtered online at 0.1–200 Hz during the MEG data acquisition.

MEG data processing

Pre-processing and averaging All recordings were pre-processed and averaged using a combination of the MNE analysis package (Gramfort, Luessi, Larson, et al., 2014), MNE-Python (Gramfort, Luessi, Larson, et al., 2013; <https://mne.tools/stable/index.html>), and the custom scripts both in Python and MATLAB. To remove noise from external sources, the signal-space projection was applied to the recordings (Fischl, Van Der Kouwe, Destrieux, et al., 2004; Tesche et al., 1995; Uusitalo & Ilmoniemi, 1997). Sensors that were visibly noisy were marked by experimenters during the recording session, then excluded from the analysis. Any other noisy or flat channels that were not picked up by experimenters during the recording were further inspected during the preprocessing stage where the channels that resulted in the rejection of 20% or more of epochs also were excluded from further analyses. On average, 4.96 channels were excluded from the participants eventually. Rejection parameters for further data quality inspection were set at 4,000 fT/cm for gradiometers, 4,000 fT for magnetometers, and 800 uV for EOG: Any epoch where any of these limits were exceeded also was excluded from further analyses. For time-course analysis, a low-pass filter of 40 Hz was applied, and no additional filter was applied to the raw data for phase-locking analysis. Recordings were epoched from 100 msec before stimulus onset until 1,000 msec post-stimulus onset. For statistical analyses, however, only time points from 0 to 500 msec were entered into statistical analyses, given that later processing is more likely to reflect invariably widespread propagation of signal across visual pathways (Collins, Freud, Kainerstorfer, Cao, & Behrmann, 2019), making the interpretation of our results more difficult. None of the participants was excluded due to excessive numbers of rejected trials. On average, 7.56 among 64 trials (11.8 %) for crowd emotion condition and 7.22 among 64 trials (11.3 %) for individual emotion condition were removed from the further analyses. We confirmed that there were no significant

differences between the conditions for the number of removed trials ($p = 0.923$).

Source localization For source localization of MEG activities, anatomical brain images of individual participants were obtained with the structural MRIs on a 1.5T Siemens Avanto 32-channel “TIM” system. The structural MRI data were acquired using T1-weighted images for the reconstruction of each subject’s cortical surface (TR = 2,300 msec, TE = 2.28 msec, flip angle = 8°, FoV = 256 x 256 mm², slice thickness = 1 mm, sagittal orientation). Two source spaces were separately created for each participant for surface-based source space and volume-based source space of the bilateral amygdala (periamygdaloid cortex; PAC), as reconstructed and segmented by Freesurfer analysis package (Fischl et al., 2004). For the surface source space, a decimated dipole grid was fitted to the inflated white matter surface in the shape of an icosahedron recursively divided 5 times, generating a 20,480-point grid. Two forward solutions were then calculated, one for each source space, using the same geometry-dependent solution calculated from a single-compartment boundary-element model (BEM). Sources closer than 5 mm to the inner skull surface were excluded from the forward solution in all cases. The MRI-head coordinate transformation for each participant was then supplied to the forward model by aligning the digitizer data obtained at the beginning of the MEG recording session (e.g., HPI positions and approximately 200 points on the scalp and the face of each participant), by using a high-resolution head surface tessellation constructed from the MRI data. The inverse operator was computed with a loose orientation constraint (LOC) parameter of 0.2 to improve localization accuracy (Lin, Belliveau, Dale, & Hämäläinen, 2006). A depth-weighting coefficient of 0.8 also was set for the inverse operator to lessen the tendency of minimum-norm estimates to be localized to superficial currents in place of deep sources. MEG data were source localized onto the whole brain using a λ^2 regularization parameter based on Signal-to-Noise Ratio (SNR) equal to $1/(\text{SNR}^2)$, on a trial-by-trial basis using the minimum-norm estimate method (MNE, Hämäläinen & Ilmoniemi, 1994) to analyze time frequencies.

ROI selection and definition Our main regions of interest (ROIs) were chosen based on our previous fMRI study (Im et al., 2017) that showed differential reactivities to stimuli of emotional crowds of faces and individual faces: posterior parietal cortex (PPC), posterior superior temporal sulcus (pSTS), posterior fusiform cortex (pFusi). We also included periamygdaloid cortex (PAC) because of its tight connections with the amygdala and role in emotional processing, face perception, and threat detection (Morris et al., 1996; Pessoa, 2010). Because it has been well established that face processing is right-lateralized (Gazzaniga & Smylie, 1983; Kanwisher et al., 1997; McCarthy et al., 1997; Miller,

Kingstone, & Gazzaniga, 2002), our ROI analyses of the phase-locking maps only included the right hemisphere to keep our comparisons simple and straightforward.

The labels for the ROIs were first functionally derived in each individual’s anatomical space within a priori anatomical constraints (automatic cortical and subcortical parcellations) produced with the Freesurfer package (Fischl et al., 2004). For all the ROIs, the functional label within the anatomical parcellation was derived from an averaged activity from all conditions, so that the activity was independent of trial type. Functional labels were generated within the anatomical parcellation corresponding to the ROI by isolating the source-space vertex with the highest activation within the anatomical constraints as well as neighboring vertices in the source-space (also within the anatomical constraints) that reach at least 60% of the maximum activation (in dSPM values). The detailed procedure for extracting the label for anatomical constraint was slightly different across our ROIs, as explained below:

- Posterior parietal cortex (PPC): We used the two labels created from anatomical parcellation automatically generated by Freesurfer package (*G_pariet_inf-Supramar* and *S_intrapariet_and_P_trans*; Destrieux, Fischl, Dale, & Halgren, 2010) that covered the superior and inferior intraparietal sulcus (IPS), respectively.
- Periamygdaloid cortex (PAC): Because no suitable a priori parcellation of the amygdala and surrounding periamygdaloid cortex was available, a posteriori anatomical constraint was imposed in the form of user-drawn label in the Freesurfer software on the fsaverage inflated surface corresponding to the cortex surrounding and including the amygdala, which we will hence refer to simply as periamygdaloid cortex (PAC). The drawing of the PAC labels was tracked by linking the drawn points to be displayed on the fsaverage MRI volume in tkmedit (<https://surfer.nmr.mgh.harvard.edu/fswiki/tkmedit>) to ensure that only the cortical surface corresponding to the amygdalae was included in the label. These anatomical constraints were then morphed to each individual’s inflated surface and used to generate functional PAC labels according to the preceding procedure.
- Posterior superior temporal sulcus (pSTS): The posterior portion of STS as the a priori parcellation generated by Freesurfer extended beyond the true pSTS on many subjects’ cortical surfaces to inferior sulci. Thus, user-drawn labels were created: User-marked constraints on the fsaverage inflated surface were marked around STS and tracked in the fsaverage MRI volume. The same morphing procedure from above was used, and then the label was split into thirds (anterior, medial, and posterior). The most posterior third was then taken as each individual’s pSTS to be used as the anatomical constraint.

- Posterior fusiform cortex (pFusi): For pFusi labels, we first extracted anatomical labels from automatic parcellation (*G_oc-temp_lat-fusifor*; Destrieux et al., 2010) then split the label into three parts (anterior, medial, and posterior) using the python script, *mne.split_label* (https://mne.tools/stable/generated/mne.split_label.html). The most posterior third was then taken as each individual's pFusi to be used as the anatomical constraint when generating the functional pFusi labels.

MEG data analysis Time courses were produced for each ROI by averaging the activity (dSPM) from source-space vertices that fell within the label marked on the individual's inflated cortical surface, separately for the crowd emotion and individual emotion conditions. From each participant's time course data of source estimates, the time point of the peak amplitude also was obtained by using the mne-python function, *stc.get_peak*, within the time window of 0–500 msec after stimulus onset. The peak latency estimates in the crowd emotion and individual emotion conditions were then submitted to correlation coefficient analyses along with the participants' mean response times for the crowd emotion and individual emotion conditions, respectively.

For phase-locking analysis, the Phase-Locking Factor (PLF) across trials was calculated for each ROI, and Phase-Locking Value (PLV) was calculated to assess functional connectivity between PAC and the other ROIs using modified scripts from the MNE-Python package (Gramfort et al., 2013). The PLF is a number between 0 and 1 (1 representing perfect synchrony) that represents a magnitude-normalized measure of the phase angle consistency across trials for a particular time-point at a particular frequency (Lachaux et al., 1999). This number was obtained by source localizing each epoch into source space using the Minimum-Norm Estimate (MNE) method with the sign of the signal preserved. Source-space MNE epochs were then entered into spectral decomposition at each time point for each frequency of interest, using a continuous wavelet transformation with a family of complex morlet wavelets containing a number of cycles equal to $f/7$, where f denotes the frequency of interest. This keeps the time window at each frequency identical, resulting in stable temporal resolution across frequency ranges. We analyzed frequencies from 8 Hz (representing the lower limit of the α -band) to 30 Hz (representing the upper limit of the β -band). Similarly, interregional connectivity was assessed with PLVs, also a magnitude-normalized measure of phase-angle consistency across trials between two different ROIs. This was calculated with the same parameters on the same frequencies as above (8–30 Hz).

E. Statistical Analysis To test our main hypothesis for differential neurodynamics underlying crowd emotion and

individual emotion processing, we performed non-parametric comparisons based on *t*-tests for crowd emotion versus individual emotion conditions. Phase-locking maps for each participant (2-dimensional images with the x-axis indicating time after stimulus onset and the y-axis indicating frequency that corresponded to phase-locking values) were smoothed via a Gaussian image filter with a kernel size of 5 and a sigma of 2 before being submitted to permutations and statistical analyses. All statistics were computed using non-parametric cluster-level permutation tests based on 3,000 permutations with a critical α -value of 0.05. Condition labels for each participant were randomly shuffled to split the data into two halves to ensure that the condition label for phase-locking map of each participant was randomized, but no one subject ended up having both phase-locking maps falling under the same condition. The permuted statistical maps were thresholded at an alpha-level of 0.05 with 45 degrees of freedom to identify clusters. Among observed clusters, those with contiguous supra-threshold time-frequency points whose masses were exceeded by 5% or less of clusters from the null distribution were considered significant. All the time points between 0 and 500 msec were submitted to analysis, with the range of frequencies from 8 to 30 Hz included for the frequency domain.

Results

Behavioral results: fast processing of crowd emotion Fig. 1c and d summarize the 46 participants' behavioral accuracies and responses times for the crowd emotion (red bar graphs) and individual emotion (blue bar graphs) conditions. The overall accuracies for both crowd emotion and individual emotion conditions were significantly higher than chance (crowd vs. chance: 72.9% vs. 50%; individual vs. chance: 81.1% vs. 50%; all p values < 0.001 ; Fig. 1c). Unlike previous studies that have reported the equivalent (or greater) accuracy for crowd perception compared to individual face perception (Im et al., 2017; Li et al., 2016; Leib et al., 2014), we found a slight drop in the behavioral accuracy for crowd emotion compared with individual emotion conditions. The accuracy for individual emotion condition was significantly higher than that for crowd emotion condition ($t(90) = 6.12$, $p < 0.001$, $\eta^2 = 0.294$). One possible explanation for this inconsistency that we observed in participants' response accuracies could be related to an intriguing perceptual asymmetry in the upper and lower visual fields (Kahn & Lawrence, 2005; Thomas & Elias, 2011). Specifically, it has been reported that the upper visual field (UVF) advantage occurs for local processing and visual inspection or search of an individual object (Previc & Blume, 1993), whereas the lower visual field (LVF) advantage occurs for global processing, motion perception, and spatial judgements (Christman, 1993; Danckert & Goodale,

2001; Niebauer & Christman, 1998; Previc, 1990). Due to the specific settings of the MEG scanner in our facility (Athinoula A. Martinos Center, Boston, MA) with the lowered position of a chair to fit in the sensor helmet attached to the MEG scanner and the floor standing display that was taller than the height of the sensor helmet, the participants' viewpoints were always lower than the center fixation of the screen, although they were always in the middle in the horizontal axis. This is a unique constraint we had in the current study, as opposed to most of the previous studies on ensemble perception in which participants' viewpoints were strictly controlled to be centered in both x- and y-dimensions during behavioral testing (Chong & Treisman, 2003; Im et al., 2016; Li et al., 2016; Leib et al., 2014) and fMRI scanning with the supine position (Cant & Xu, 2012, 2014; Im et al., 2017). Thus, the participants were looking upward to see the stimuli during the entire MEG recording in the current study with the whole stimuli presented in their upper visual field, which possibly facilitated individual emotion processing relative to crowd emotion processing. To validate this speculation, future research will be needed to examine whether the upper visual field versus lower visual field presentation can selectively facilitate the perception of crowd emotions and individual emotion, respectively.

As shown in Fig. 1d, the response times (RTs) for the crowd emotion and individual emotion conditions did not significantly differ from each other (1.076 sec vs. 1.023 sec: $t(90) = 1.531$, $p = 0.129$, $\eta^2 = 0.025$). This is consistent with the previous reports that ensemble coding can occur as fast as, or somewhat faster than, processing of a single item in visual stimuli comprised of faces (Haberman et al., 2009; Im et al., 2017; Leib et al., 2014) and dots (Ariely, 2001). The relative lack of a decrease in RT for deciding between two individual faces (one in each visual field) in the individual emotion condition and eight faces (four in each visual field) in the crowd emotion condition suggest that crowd emotion from eight faces can be extracted as a whole in a parallel manner, not necessarily relying on serial processing of individuals (Im et al., 2017).

MEG results: Temporal profiles of dorsal and ventral areas during crowd and individual emotion processing Figure 2 shows the group average ($N = 46$) of noise-normalized dynamic statistical parametric maps (dSPMs) of the whole brain during the crowd emotion and individual emotion processing (thresholded at $dSPM > 4.6$) at the two different time points: 70 msec (Fig. 2a) and 370 msec (Fig. 2b) after stimulus onset. As early as 70 msec after stimulus onset, dSPM activations of the posterior parietal cortex (PPC) was predominant in the crowd emotion condition, but not in the individual emotion condition, suggesting relatively early engagement of the PPC in the processing of crowd emotion. Later at 370 msec after stimulus onset, the dSPM of the periamygdaloid cortex (PAC) and the posterior fusiform cortex (pFusi), as well as the medial

frontal areas, increased in the crowd emotion condition, while the PPC activation decreased. Earlier activation of PPC in the crowd emotion condition (but not in the individual condition) suggests that rapid extraction of crowd emotion can be initially mediated by the PPC area, followed by the involvement of the ventral stream areas in relatively later stages of crowd emotion processing. Unlike the crowd emotion processing, the individual emotion processing did not show the early onset of the PPC area at 70 msec after stimulus onset. The later activation of the pFusi and other ventral stream areas (including the inferior and middle temporal regions) at 370 msec after stimulus onset, however, was more salient and extensive in the individual emotion condition compared with the crowd emotion condition. Relatively greater engagement of the PPC areas and pFusi areas in crowd emotion processing and individual emotion processing, respectively, also has been observed in our prior work using fMRI (Im et al., 2017). The current findings, however, describe the novel findings on temporal dynamics of activity in PPC, pFusi, and PAC during the processing of crowd emotion and individual emotion.

We also computed correlations between the participants' response times and peak latencies of the dSPM values for crowd emotion and individual emotion conditions, separately. The correlation coefficients (Pearson's r) were calculated between the 46 participants' mean response times and the peak latency estimates from the three of our ROI labels – pFusi and the two subregions of the PPC, superior and inferior intraparietal sulcus (IPS). For the crowd emotion condition, we found statistically significant, positive correlations (corrected for multiple comparisons) between the participants' response times and the peak latencies of the superior IPS ($r = 0.567$, $p < 0.0001$) and between their response times and the peak latencies of the inferior IPS ($r = 0.466$, $p = 0.0013$), but not the peak latencies of the pFusi activation ($r = 0.014$, $p = 0.926$). For the individual emotion condition, however, none of the correlation coefficients was statistically significant (all $r < 0.320$, all $p > 0.032$). Thus, early onset of activity in the PPC appears to be associated with participants' response speed but only in the crowd emotion condition.

Although this was not the primary focus of our investigation, we also observed different patterns of hemispheric lateralization at the two time points. At 70 msec after stimulus onset, the dSPMs were quite symmetrical in the left and right hemispheres, with suprathreshold activity visible mostly in the crowd emotion condition (Fig. 2a, left panel). At the later time point (370 msec), we observed right hemisphere dominance in the dSPM of the inferior temporal and visual areas, particularly in the individual emotion condition.

Phase-locking results We next examined the time courses of our ROIs (PPC, pFusi, pSTS, and PAC) in more detail by looking at the phase-locking patterns when these ROIs were engaged in the processing of crowd emotion and individual

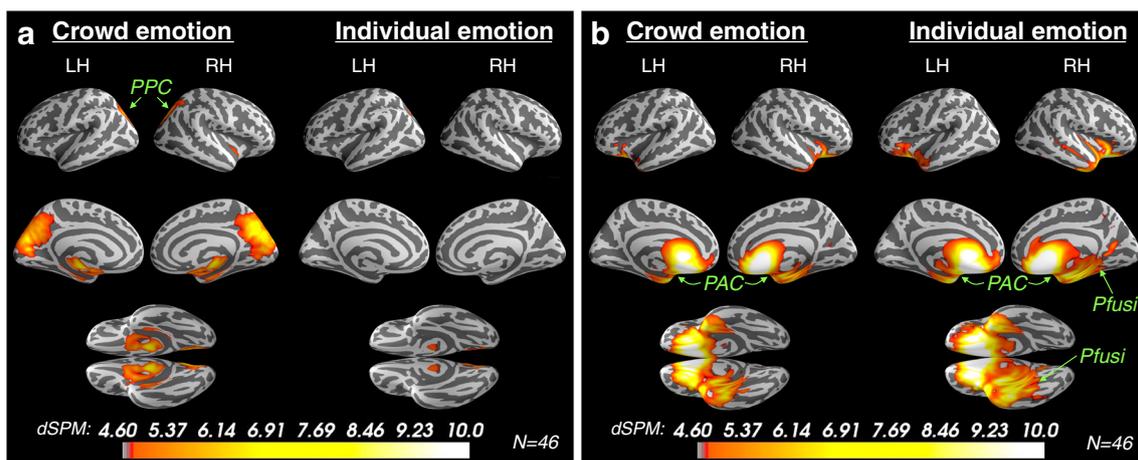


Fig. 2 Dynamic statistical parametric map (dSPM) at two different time points (70 msec and 370 msec) after stimulus onset. Group averages ($N = 46$) are overlaid on the surface of FSAverage template. **a** dSPM at 70 msec after stimulus onset in the crowd emotion condition (the left

column) and individual emotion condition (the right column). **b** dSPM at 370 msec after stimulus onset in the crowd emotion condition (the left column) and individual emotion condition (the right column)

emotion. Although there is direct communication between the dorsal and ventral visual pathways at multiple stages and locations (Distler, Boussaoud, Desimone, & Ungerleider, 1993; Nassi & Callaway, 2009; Rosa et al., 2009; Webster, Bachevalier, & Ungerleider, 1994; Zhong & Rockland, 2003), information processing after a widespread cascade of information and long-range feedback modulations between the processing streams in later stages (Givre, Schroeder, & Arezzo, 1994; Oram & Perrett, 1996; Schroeder, Mehta, & Givre, 1998) is particularly ambiguous and challenging to interpret.

Because our main goal was to establish the temporal precedence of activations in the dorsal stream to mediate the rapid extraction of crowd emotion, we focused on the earlier segments of the time course, between 0 and 500 msec after stimulus onset to examine early, stimulus-evoked activity, rather than later components. To identify differential temporal dynamics of crowd emotion processing and individual emotion processing, we directly compared the time courses in the alpha and beta frequency bands from the phase-locking data (PLF; phase-locking factor) across trials for each of our ROIs during crowd emotion versus individual emotion processing. The PLF is a number between 0 and 1 (1 representing perfect synchrony) that represents a magnitude-normalized measure of the phase angle consistency across trials for a particular time-point at a particular frequency (Lachaux et al., 1999).

Figure 3 shows differences in amplitude of PLFs in both the alpha and beta-bands during the perception of crowd emotion (red clusters) versus individual emotion (blue clusters) for each of the ROIs. We only plot statistically significant clusters at nonparametric $p < 0.05$, corrected for multiple comparisons. Compared with individual emotion condition, crowd emotion condition showed earlier and greater PLFs mostly in the beta-band, as shown in the

periamygdaloid cortex (PAC, Fig. 3a), the posterior superior temporal sulcus (pSTS, Fig. 4b), and the two subregions (superior intraparietal sulcus and inferior intraparietal sulcus) of the posterior parietal cortex (PPC, Fig. 3c), and the posterior fusiform cortex (pFusi, Fig. 3d). Specifically, the right PAC (Fig. 3a) showed greater PLFs for crowd emotion condition with a peak at about 170 msec (cluster around 155–218 msec) after stimulus onset in the beta-band (between 13 Hz and 30 Hz), followed by greater and sustained PLFs for individual emotion processing with an initial peak at approximately 250 msec (cluster between 205–295 msec) mostly in the alpha-band. Similarly, the right pSTS (Fig. 3b) showed greater PLFs for crowd emotion condition which reached a peak at approximately 157 msec (cluster around 137–196 msec) in both alpha- and beta-bands, whereas greater PLFs for individual emotion condition reached a peak much later, at approximately 351 msec (cluster around 285–385 msec), predominantly in the alpha-band. The two Freesurfer labels that split the posterior parietal cortex (PPC; Fig. 3c and d) into inferior and superior intraparietal sulcus showed earlier peaks for greater PLFs during crowd emotion processing compared with individual emotion processing than any other ROIs. In the inferior intraparietal sulcus (IPS; Fig. 3c), we observed that the significantly greater PLFs for crowd emotion condition reached a peak at approximately 82 msec (cluster between 68–118 msec) across both alpha and beta-bands, followed by a later peak greater for individual emotion condition at approximately 272 msec (cluster around 230–295 msec) in the alpha-band. In the other subregion of PPC, superior IPS (Fig. 3d), we only found significantly greater PLFs for crowd emotion condition, which reached a peak at approximately 121 msec (cluster around 103–130 msec) in the beta-band and at approximately 168 msec (cluster around 160–188 msec) in the

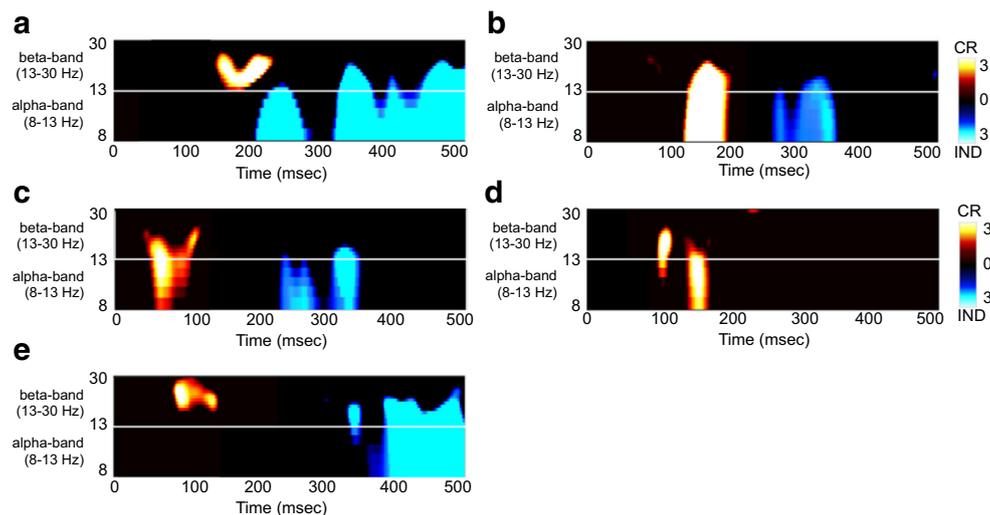


Fig. 3 Statistical maps of phase-locking across trials (0-500 msec after stimulus onset), showing significant contrasts (nonparametric $p < 0.05$, corrected for multiple comparisons) between crowd emotion (red clusters) versus individual emotion (blue clusters) conditions in each of the

ROIs: **a** Right periamygdaloid cortex (PAC). **b** Right posterior superior temporal sulcus (pSTS). **c** Right inferior intraparietal sulcus (inferior IPS, a subset of PPC). **d** Right superior intraparietal sulcus (superior IPS, a subset of PPC). **e** Right posterior fusiform cortex (pFusi)

alpha-band, suggesting that this area was mostly engaged in crowd emotion processing and less so in individual emotion processing. Finally, the posterior fusiform cortex (pFusi; Fig. 3e) showed an early, brief peak at approximately 113 msec (cluster around 106-143 msec) in the beta-band for crowd emotion processing, followed by later processing beginning at approximately 365 msec (cluster starting from 245 msec) for individual emotion processing, mostly in the alpha-band. Table 1 shows a comprehensive list of the timing of the statistically significant clusters in each of these ROIs separately for crowd emotion and individual emotion processing. Earlier emergence of greater PLFs in our ROIs during crowd emotion condition than individual emotion condition suggests that crowd emotion processing in the dorsal stream is not simply a consequence of activity propagating from sustained individual emotion processing in the ventral stream. Rather, the dorsal stream areas appear to be engaged predominantly in crowd emotion processing, particularly early on, and this activity occurs before or concurrently with the involvement of face-selective areas in the ventral stream (e.g., pFusi). While we did not find clear spectral separation of activity evoked by crowd and individual emotion processing, there appears to be a tendency for the former to peak in a slightly higher frequency band, at least in some of our ROIs (PAC, posterior fusiform cortex).

Connectivity between periamygdaloid cortex and cortical regions in the dorsal and ventral pathways Any complex processing requires the integration of numerous functional areas widely distributed over the brain, usually via phase-locked oscillatory interactions (Friston, Stephan, & Frackowiak, 1997; Tonini & Edelman, 1998). To better understand how the brain processes a given task, it is necessary not only to

look at the temporal profiles of activation in each region individually but also to examine how these regions interact over time. Therefore, we next examined the phase-locking between our ROIs as a measure of interregional functional connectivity to build a more complete picture of the neurodynamics at play between PAC and the brain regions in the dorsal stream (e.g., PPC) and the ventral stream (e.g., pFusi) to gain insights into how the amygdaloid complex interacts with the dorsal and ventral stream regions during processing of crowd emotion and individual emotion.

Figure 4 shows the PLVs (phase-locking values) between the periamygdaloid cortex and the superior intraparietal sulcus in the dorsal stream (e.g., PAC and superior IPS; Fig. 4a) and the periamygdaloid cortex and the posterior fusiform cortex in the ventral stream (e.g., PAC and pFusi; Fig. 4b). Between the PAC and the PPC (Fig. 4a), we found earlier and greater peaks in PLVs in the alpha-band around 233-262 msec with a peak at approximately 246 msec for crowd emotion condition, followed by later increase in the PLVs for individual emotion condition, both at the alpha-band (earlier peak at around 328-357 msec) and the beta-band (later peak at around 367-402 msec). On the other hand, between the PAC and the pFusi (Fig. 4b), we found that the synchronization order related to crowd versus individual emotion processing was reversed: namely, earlier, greater PLVs in the alpha-band were observed for individual emotion condition (at around 105-129 msec), followed by later, greater PLVs in the beta-band for crowd emotion condition (at around 395-464 msec). Together, phase-locking in the amygdala-dorsal stream network (e.g., PAC and PPC) and the amygdala-ventral stream network (e.g., PAC-pFusi) reflects differential temporal dynamics underlying crowd emotion and individual emotion processing. Finally, crowd emotion processing showed greater beta-band

Table 1 Timing of significant differences in activation within ROIs from 0 msec and 500 msec after stimulus onset. Reported *p*-values are nonparametric and corrected based on cluster permutations, thresholded at $p < 0.05$

ROI	Contrast	Peak (msec)	Time range (msec)	<i>p</i> value
Periamygdaloid cortex (PAC)	Crowd > Individual	170	155-218	$p < 0.001$
	Individual > Crowd	250	205-295	$p < 0.0001$
	Individual > Crowd		318-500 msec	$p < 0.0001$
Posterior superior temporal sulcus (pSTS)	Crowd > Individual	157 msec	137-196 msec	$p < 0.0001$
	Individual > Crowd	351 msec	285-385 msec	$p < 0.01$
Posterior parietal cortex (PPC)				
Inferior intraparietal sulcus (IPS)	Crowd > Individual	82 msec	68-118 msec	$p < 0.0005$
	Individual > Crowd	272 msec	230-295 msec	$p < 0.01$
	Individual > Crowd	342 msec	310-355 msec	$p < 0.001$
Superior intraparietal sulcus (IPS)	Crowd > Individual	121 msec	103-130 msec	$p < 0.0001$
	Individual > Crowd	168 msec	160-188 msec	$p < 0.0001$
	Individual > Crowd	None		n.s.
Posterior fusiform cortex (pFusi)	Crowd > Individual	113 msec	106-143 msec	$p < 0.001$
	Individual > Crowd	360 msec	340-370 msec	$p < 0.001$
	Individual > Crowd		380-500 msec	$p < 0.0001$

(frequency range of 13-30 Hz) involvement during the neural synchronies between PAC and pFusi, just as observed in single-region phase-locking patterns of PAC and pFusi (see red clusters in the beta-bands of Fig. 3a and e).

Discussion

We used magnetoencephalography (MEG) to characterize temporal dynamics of brain regions in the dorsal stream (e.g., posterior parietal cortex [PPC]) and the ventral stream (e.g., posterior fusiform cortex [pFusi]), as well as the

periamygdaloid complex (PAC) during the processing of crowd emotion from groups of faces and individual emotion of single faces. To summarize our main findings, we reported that 1) MEG (dSPM) measurements showed early onset of activity in the posterior parietal cortex, predominantly in the crowd emotion condition, followed by a later activation of posterior fusiform cortex in the individual emotion condition; 2) the behavioral response times for crowd emotion condition were positively correlated with the peak latency of the PPC areas; 3) all of our ROIs (PAC, pSTS, PPC, and pFusi) showed earlier phase-locking measurements (PLFs) for crowd emotion processing compared with individual emotion processing; 4) the interaction between PAC and PPC (amygdala-dorsal stream connectivity) revealed earlier phase-locking for crowd emotion processing, followed by significant phase-locking for individual emotion processing later; and 5) the interaction between PAC and pFusi (amygdala-ventral stream connectivity) revealed early phase synchrony for individual emotion processing, followed by the later beta-band phase synchrony for crowd emotion processing. Our findings suggest that the crowd emotion perception exhibits distinct neurodynamics patterns from those mediating the processing of single emotional faces to support a rapid social affective judgment during a brief exposure.

This paper reported that the dorsal and ventral ROIs (PPC, pSTS, and pFusi) that we examined showed different patterns of activity while subjects were comparing crowd emotion and individual emotional face stimuli, consistent with our prior fMRI study (Im et al., 2017). Importantly, the current MEG findings add new information about the timing and dynamic interactions of activity in these regions during the perception of crowd emotion and individual emotion. These findings

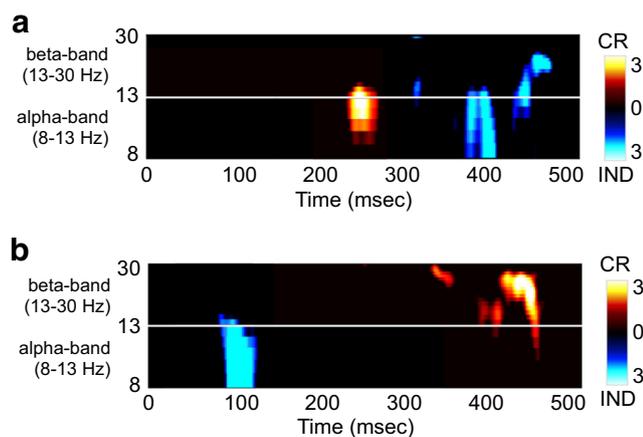


Fig. 4 Significant phase-locking between ROIs in the crowd emotion condition (red clusters) and individual emotion condition (blue clusters), from 0 msec to 500 msec after stimulus onset. **a** Significant phase-locking between the right periamygdaloid cortex (PAC) and the right superior intraparietal sulcus (PPC). **b** Significant phase-locking between the right periamygdaloid cortex (PAC) and the right posterior fusiform cortex (pFusi)

together suggest that there exist two distinct processing routes at play: a faster processing route via the dorsal stream and a slower processing route via the ventral stream. The two processing routes—faster and slower—in the dorsal and ventral streams appear to be engaged in both processing of crowd emotion and individual emotion but in qualitatively distinct temporal profiles and different functional connectivities to a subcortical region (e.g., PAC). Our finding that the onset of activations and neural synchronies in our ROIs during the processing of crowd emotion preceded those during the processing of individual emotion also is well in line with the notion that perceiving crowd emotion from multiple faces relies on fast and efficient ensemble coding. Just as in many previous studies (Haberman et al., 2009; Im et al., 2017; Leib et al., 2014), our behavioral data showed that participants' responses were not significantly delayed with more faces to be processed (up to 8 faces in the current study). However, how such efficient, fast coding of crowd emotion as an ensemble can occur has been an unanswered question. By examining fine-scale temporal profiles of neural processes underlying crowd emotion perception, the current MEG study can suggest that the efficiency of crowd emotion processing can be achieved by earlier engagement of multiple brain areas and faster dynamic interactions between them. Although past research studies on ensemble coding have focused mostly on identifying types of visual features that can be extracted as ensembles (e.g., size, orientation, numerosity, color, facial emotion, gender, identity, and age, etc.) and characterizing compositions of ensemble representations (e.g., average, variance, median, or range, etc.), our current findings emphasize the importance of examining temporal profiles (e.g., processing speeds) in which ensemble processing contrasts sharply with the processing of individual objects. Therefore, the current findings provide novel neuroimaging evidence for the long-standing framework that perceptual averaging is accomplished in a qualitatively different way, possibly as parallel, holistic processing (Ariely, 2001; Baek & Chong, 2020; Chong et al., 2008), instead of relying on serial processing of a few individual object representations by subsampling heuristics (Maule & Franklin, 2016; Myczek & Simons, 2008). Ensemble coding of facial crowds may use qualitatively and functionally different visual representations, instead of the mere sum or a part of singular representations of each emotional face.

It should be noted that our phase-locking results represent direct contrasts of spectrotemporal profiles between crowd emotion and individual emotion perception focused on significant differences between the time courses for these conditions. Thus, our results do not imply that there was no dorsal stream activity at all during individual face comparisons. Instead, our results show that the dorsal visual stream could differentiate a facial crowd and an individual as early as 68 msec, preferentially responding to facial crowds over

individual faces. Previous neuroimaging studies have shown that emotional stimuli (e.g., fear) of facial expressions or body movements preferentially activate the dorsal visual stream, suggesting a tighter functional link between the emotion and action, compared with neutral emotion (de Gelder, Snyder, Greve, Gerard, & Hadjikhani, 2004; Huis In't Veld & de Gelder, 2015). It also was reported that fearful body expressions activated the dorsal stream more strongly, whereas happy expressions preferentially activated the ventral stream (de Borst & de Gelder, 2016). Another MEG study also reported that the dorsal visual stream regions showed greater activation for a fearful body expression than neutral, emerging as early as 80 msec (Meeren et al., 2016). Together, our study presents further empirical evidence that the dorsal visual stream is highly attuned to visual representations that are emotionally expressive and action-relevant. The functional distinction between the dorsal and ventral streams during social-emotional visual processing appears to be modulated flexibly depending on the valence and type of stimuli, rather than serving as specialized modular systems for a particular type of visual inputs.

In a classical framework, distinct functions for the “vision for action” and “vision for perception” were associated with the dorsal and ventral streams, respectively (Goodale et al., 1994; Goodale & Milner, 1992; Milner & Goodale, 2008). According to this framework, different social-emotional visual stimuli can engage either visual stream depending on its principal behavioral motivation of the processing. For example, some social-emotional stimuli have relatively greater ecological values for fast, action-oriented processing, such as fearful stimuli, collective and coherent social signals, or clear threat cues (Cushing et al., 2018; de Borst & de Gelder, 2016; de Gelder et al., 2004) than others (e.g., neutral or happy expressions and ambiguous facial cues). A crowd's overall emotion that we tested in the current study also appears to convey such behavioral significance to observers, driving stronger action-oriented signals than an individual's emotion does. Indeed, the average direction of eye gaze in a crowd of people is more effective than an individual's eye gaze for directing one's attention (Gallup et al., 2012), suggesting that social information conveyed by a crowd amplifies our social reactions (Sweeny & Whitney, 2014). Along the same line, greater involvement of the dorsal stream in ensemble coding of multiple emotional faces may reflect that visual representation of crowd emotion provides stronger cues for context-relevant and time-sensitive action preparation and execution, as suggested in a previous study on perception of interactive body movements in a crowd (Huis In't Veld & de Gelder, 2015). Given that the visual system can process only a limited number of objects (no more than four; Luck & Vogel, 1997), however, stimuli containing multiple, different faces are not as suitable as a single face for real-time

person recognition and identification, which is a dominant function of the ventral visual stream (Tsao & Livingstone, 2008). Our findings suggest that the fast and efficient coding of the average emotion of facial crowds can preferentially activate action-oriented signals (e.g., avoiding an angry mob) by projecting the “quick and dirty” global representation to the dorsal processing stream.

By investigating different temporal dynamics of the ROIs in the dorsal (PPC and pSTS) and ventral (pFusi) streams, the current findings also highlight the roles of the dorsal pathway in visual information processing. It has been unclear whether visual representations of objects, such as faces, in the dorsal pathway result from independent computation or whether they simply reflect the projection of information processed in the ventral pathway. For example, one notion is that visual representations of the shape of objects in the dorsal regions are the output of computations in the occipitotemporal cortex, which the posterior parietal cortex only “uploads” from the ventral pathway depending on the current task goal (see Xu, 2018a and 2018b for reviews). Alternatively, others suggest that visual shape information is at least partially computed independently in each of the two major pathways (see Freud, Plaut, & Behrmann, 2016, for review). The current findings are consistent with the notion that the dorsal visual pathway plays more active and independent roles in supporting and mediating different units of visual representations (e.g., ensembles and individuals) than it has been traditionally proposed. For example, we observed that the inferior intraparietal sulcus (IPS), one of the two subregions of the PPC in the dorsal stream, showed early phase-locking related to crowd emotion processing compared with individual emotion processing, starting from 68 msec after stimulus onset and preceding any other ROIs we tested. This is impossible if the dorsal stream were merely the recipient of information represented in the ventral stream.

What roles do the regions in the dorsal stream network play during the crowd emotion processing, in parallel to the ventral stream? One possibility is that magnocellular input, which is more dominant in the dorsal stream than the ventral stream (although the dorsal pathway also receives parvocellular and koniocellular inputs; Nassi, Lyon, & Callaway, 2006; Sincich, Park, Wohlgenuth, & Horton, 2004), benefits crowd emotion processing by being better suited for global information processing with a fast readout time. The magnocellular pathway is known to be highly sensitive to low-spatial frequency information (Derrington & Lennie, 1984; Tootell, Silverman, Hamilton, Switkes, & De Valois, 1988), transmits faster (Breitmeyer, 1975; Lupp, Hauske, & Wolf, 1976), and is critical for global processing (Hughes, Fendrich, & Reuter-Lorenz, 1990; Hughes, Nozawa, & Kitterle, 1996). Rapid global representation via magnocellular input to the dorsal

stream can contribute to the rapid extraction of ensemble features, including those involved in perceiving crowd emotion. Besides, rapid transmission of magnocellular signals has also been suggested to facilitate slower processing for object recognition in the ventral stream (e.g., fusiform gyrus) by enabling early predictions or “initial guess” via top-down guidance or modulation (Bar et al., 2006; Kveraga, Boshyan, & Bar, 2007; Tapia & Breitmeyer, 2011). Our findings in the current study suggest that the dorsal stream is engaged in an early stage of perceptual processing, perhaps relying on rapid magnocellular input to form a global representation of the stimulus emotion.

At the level of attentional selection, the dorsal stream can also play a critical role in the processing of crowd emotion and individual emotion by rapidly activating different parietal attentional mechanisms prior to object processing of the ventral stream. The dorsal visual stream is known to play major roles in the rapid and efficient visuospatial orienting of attention and the deployment of attentional resources (Marrett et al., 2011; Sciberras-Lim & Lambert, 2017; Siegel, Donner, Oostenveld, Fries, & Engel, 2008). Therefore, it would be important that the dorsal stream is actively involved in the visual processing of perception of crowd emotion and individual emotion for efficient deployment of attention. It has been suggested that ensemble perception and individual object processing rely on different attentional modes: non-selective, distributed attention versus focused attention for a selective mechanism, respectively (for review, see Baek & Chong, 2020; Chong & Evans, 2011; Wolfe et al., 2011). The early engagement of the dorsal visual stream may signify that different attentional modes—distributed or focused—can be selectively activated by its attentional network that precedes and facilitates visual processing of crowd emotion and individual emotion.

Finally, another important aspect of visual processing of the dorsal stream regions (e.g., parietal cortex) is that it is closely linked to the ability to track task demand and intention (Snyder, Batista, & Andersen, 2000; Toth & Assad, 2002). For example, fMRI response amplitude in the superior IPS in the dorsal stream tracks behavioral visual working memory capacity for various task-relevant visual features (e.g., color and shape, etc.: Todd & Marois, 2004, 2005; Xu & Chun, 2006; Jeong & Xu, 2013; see also Xu & Jeong, 2015). Previous fMRI data also suggest that the human parietal cortex participates in the moment-to-moment goal-directed visual representation of faces, invariant to changes in low-level visual features and viewpoints (Jeong & Xu, 2016).

Conclusions

Our new findings on differential temporal dynamics in the dorsal and ventral stream activations provide a new

framework encompassing the two visual pathways that differentially contribute to complementary visual functions: the ventral stream, including the fusiform cortex that mediates the detailed visual perception of individual emotional faces and the dorsal stream including the posterior parietal cortex that allows for rapid, global, and goal-driven representations of crowd emotion.\

Code availability All the MATLAB codes for the behavioral and fMRI analyses presented in this article are available from the first author (heeyeon.im@ubc.ca) or the corresponding author (keastas@nmr.mgh.harvard.edu) upon the request.

Author Contributions H. Y. Im and K. Kveraga developed the study concept and designed the experiment. Testing and data collection were performed by H. Y. Im, C. A. Cushing, and N. Ward. H. Y. Im analyzed the data and wrote the first draft of the manuscript; and H. Y. Im and K. Kveraga edited the manuscript.

Data availability The datasets generated during and/or analyzed during the current study are available from the first author (heeyeon.im@ubc.ca) or the corresponding author (keastas@nmr.mgh.harvard.edu) on reasonable request.

Declarations

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

References

- Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, *19*(4), 392–398.
- Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(18), 7345–7350.
- Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science*, *12*(2), 157–162.
- Baek, J., & Chong, S. C. (2020). Ensemble perception and focused attention: Two different modes of visual processing to cope with limited capacity. *Psychonomic Bulletin & Review*, Advance online publication, 1–5.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmidt, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 449–54.
- Borra, E., Belmalih, A., Calzavara, R., Gerbella, M., Murata, A., Rozzi, S., & Luppino, G. (2008). Cortical connections of the macaque anterior intraparietal (AIP) area. *Cerebral Cortex*, *18*, 1094–1111.
- Brainard, D. H. The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436 (1997).
- Breitmeyer B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, *15*, 1411–1412.
- Budisavljevic, S., Dell'Acqua, F., & Castiell, U. (2018). Cross-talk connections underlying dorsal and ventral stream integration during hand actions. *Cortex*, *103*, 224–239.
- Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*, 425–428.
- Cant, J. S. & Xu, Y. (2012). Object ensemble processing in human anterior -medial ventral visual cortex. *Journal of Neuroscience*, *32*, 7885–7700.
- Catani M., Jones D. K., Ffytche D. H. (2005). Perisylvian language networks of the human brain. *Annual Neurology*, *57*, 8–16.
- Chong, S. C. & Evans, K. K. (2011). Distributed versus focused attention (count vs estimate). *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(6), 634–638.
- Chong, S. C. & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*, 393–404.
- Christman, S.D. (1993). Local-global processing in the upper versus lower visual fields. *Bulletin of the Psychonomic Society*, *31*, 275–278.
- Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences*, *20*, 324–335.
- Collins, E., Freud, E., Kainerstorfer, J. M., Cao, J., & Behrmann, M. (2019). Temporal Dynamics of Shape Processing Differentiate Contributions of Dorsal and Ventral Visual Pathways. *Journal of Cognitive Neuroscience*, *31*, 821–836.
- Cushing, C., Im, H.Y., Adams, R.B. Jr., Ward, N., Albohn, N.D., Steiner, T.G., & Kveraga, K. (2018). Neurodynamics and connectivity during facial fear perception: The role of threat exposure and signal congruity. *Scientific Reports*, *8*, 2776.
- Cushing, C., Im, H.Y., Adams, R.B. Jr., Ward, N., & Kveraga, K. (2019). Magnocellular and parvocellular pathway contributions to facial threat cue processing. *Social Cognitive and Affective Neuroscience*, *14*, 151–162.
- Dakin, S. C. & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, *37*, 3181–3192.
- Danckert, J., & Goodale, M. A., (2001). Superior performance for visually guided pointing in the lower visual field. *Experimental Brain Research*, *137*(3), 303–308.
- de Borst, A.W., de Gelder, B. (2016). Clear signals or mixed messages: Inter-individual emotion congruency modulates brain activity underlying affective body perception. *Social Cognitive and Affective Neuroscience*, *11*(8), 1299–309.
- de Gelder, B., Snyder, J., Greve, D., Gerard, G., & Hadjikhani, N. (2004). Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(47), 16701–16706.
- Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque. *Journal of Physiology*, *357*, 219–240.
- Destrieux, C., Fischl, B., Dale, A., & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage*, *53*, 1–15.
- Distler, C., Boussaoud, D., Desimone, R., & Ungerleider, L. G. (1993). Cortical connections of inferior temporal area TEO in macaque monkeys. *Journal of Comparative Neurology*, *334*(1), 125–150.
- Dolcos, F., Iordan, A. D., & Dolcos, S. (2011). Neural correlates of emotion-cognition interactions: A review of evidence from brain imaging investigations. *Journal of cognitive psychology (Hove, England)*, *23*(6), 669–694.
- Dumas, T., Dubal, S., Attal, Y., et al. (2013). MEG evidence for dynamic amygdala modulations by gaze and facial emotions. *PLoS ONE*, *8*(9), 1–11.
- Ehlers, C. L., Wills, D. N., Desikan, A., Phillips, E., & Havstad, J. (2014). Decreases in energy and increases in phase locking of event-related oscillations to auditory stimuli occur during adolescence in human and rodent brain. *Developmental neuroscience*, *36*, 175–195.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of Facial Affect* (Consulting Psychologists Press, Palo Alto, CA, 1976).

- Elias, E., Dyer, M., & Sweeny, T. D. (2017). Ensemble perception of dynamic emotional groups. *Psychological Sciences*, *28*(2), 193–203.
- Fischl, B., Van Der Kouwe, A., Destrieux, C., et al. (2004). Automatically parcellating the human cerebral cortex. *Cerebral Cortex*, *14*(1), 11–22.
- Fitts, P. M. & Seeger, C. M. (1953). S-R compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, *46*, 199–201.
- Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. (2009). The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study. *NeuroImage*, *44*(2), 569–580.
- Freud, E., Plaut, D. C., & Behrmann, M. (2016). "What" is happening in the dorsal visual pathway *Trends in Cognitive Sciences*, *20*, 773–784.
- Friston, K. J., Stephan, K. M., & Frackowiak, R. S. (1997). Transient phase-locking and dynamic correlations: Are they the same thing? *Human Brain Mapping*, *5*(1), 48–57.
- Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, *4*, 59–65.
- Gallup, A. C., Hale, J. J., Sumpter, D. J., Garnier, S., Kacelnik, A., Krebs, J. R., & Couzin, I. D. (2012). Visual attention and the acquisition of information in human crowds. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(19), 7245–7250.
- Gazzaniga, M. S., & Smylie, C. S. (1983). Facial recognition and brain asymmetries: clues to underlying mechanisms. *Annals of Neurology*, *13*, 536–540.
- Givre, S. J., Schroeder, C. E., & Arezzo, J. C. (1994). Contribution of extrastriate area V4 to the surface-recorded flash VEP in the awake macaque. *Vision Research*, *34*(4), 415–428.
- Goodale M.A., Meenan J.P., Bühlhoff H.H., Nicolle D.A., Murphy K.J., Racicot C.I. Separate neural pathways for the visual analysis of object shape in perception and prehension. *Curr Biol* 1994;4:604–610.
- Goodale M.A., Milner A.D. Separate visual pathways for perception and action. *Trends Neurosci*. 1992;15:20–25.
- Gramfort, A., Luessi, M., Larson, E., et al. (2013). MEG and EEG data analysis with MNE-Python. *Front Neurosci*. *7*, 1–13.
- Gramfort, A., Luessi, M., Larson, E., et al. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, *86*, 446–60.
- Haberman, J., Harp, T., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision*, *9*(11), 1–13.
- Haberman, J. & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751–753.
- Haberman, J. & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception & Performance*, *35*(3), 718–734.
- Halberda, J., Sires, S. F., & Feigenson, L. (2016). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science*, *17*(7), 572–576.
- Hämäläinen, M.S., Ilmoniemi, R.J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & Biological Engineering & Computing*, *32*(1), 35–42.
- Haxby, J. V., Grady, C. L., Horowitz, B., Ungerleider, L. G., Mishkin, M., Carson RE, et al. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(5), 1621–1625.
- Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*, *69*(2), 387–396.
- Hughes, H. C., Fendrich, R., & Reuter-Lorenz, P. A. (1990). Global versus local processing in the absence of low spatial frequencies. *Journal of Cognitive Neuroscience*, *2*, 272–282.
- Hughes, H. C., Nozawa, G., & Kitterle, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience*, *8*(3), 197–230.
- Huis In't Veld, E. M. J., & de Gelder, B. (2015). From individual fear to mass panic. The neurological basis of crowd perception. *Human Brain Mapping*, *36*, 2338–1351.
- Im, H.Y., Albohn, N.D., Steiner, T.G., Cushing, C., Adams, R.B.Jr., & Kveraga, K. (2017). Ensemble coding of crowd emotion: Differential hemispheric and visual stream contributions. *Nature Human Behaviour*, *1*, 828–842.
- Im, H.Y., & Halberda, J. (2013). The effects of sampling and internal noise on the representation of ensemble average size. *Attention, Perception, & Psychophysics*, *75*, 278–286.
- Im, H. Y., Park, W., & Chong, S. C. (2015). Ensemble statistics as a unit of selection. *Journal of Cognitive Psychology*, *27*, 114–127.
- Im, H.Y., Tiurina, N.A., & Utochkin, I.S. (2020). An explicit investigation of the roles that feature distributions play in rapid visual categorization. *Attention, Perception, & Psychophysics*
- Im, H.Y., Zhong, S., & Halberda, J. (2016). Perceptual groups as a unit for rapid extraction of approximate number of elements in random dot arrays. *Vision Research*, *126*, 291–307.
- Jeong, S. K. and Xu, Y. (2013). Neural representation of targets and distractors during visual object individuation and identification. *Journal of Cognitive Neuroscience*, *25*, 117–126.
- Jeong, S. K. and Xu, Y. (2016). Behaviorally relevant abstract object identity representation in the human parietal cortex. *Journal of Neuroscience*, *36*, 1607–1619.
- Ji, L., & Pourtois, G. (2018). Capacity limitations to extract the mean emotion from multiple facial expressions depend on emotion variance. *Vision Research*, *145*, 39–48.
- Kahn, M. A., & Lawrence, G. P. (2005). Differences in visuomotor control between the upper and lower visual fields. *Experimental Brain Research*, *164*(3), 395–398.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of The Royal Society B Biological Sciences*, *361*(1476), 2109–2128.
- Kveraga K., Boshyan J. & Bar M. (2007) The magnocellular trigger of top-down facilitation in object recognition. *Journal of Neuroscience*, *27*(48), 13232–13240.
- Lachaux, J.P., Rodriguez, E., Martinerie, J., Varela, F.J. (1999). Measuring phase synchrony in brain signals. *Human Brain Mapping*, *8*(4), 194–208.
- Leib, A. Y., Fischer, J., Liu, Y., Qui, S., Robertson, L., & Whitney, D. (2014). Ensemble crowd perception: a viewpoint-invariant mechanism to represent average crowd identity. *Journal of Vision*, *14*, 1–13.
- Leib, A. Y., Puri, A. M., Fischer, J., Bentin, S., Whitney, D., & Robertson, L. (2012). Crowd perception in prosopagnosia. *Neuropsychologia*, *50*(7), 1698–1707.
- Li, H., Ji, L., Tong, K., Ren, N., Chen, W., et al. (2016). Processing of individual items during ensemble coding of facial expressions. *Frontiers in Psychology*, *7*, 1332.
- Lin, F.H., Belliveau, J.W., Dale, A.M., Hämäläinen, M.S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, *27*(1), 1–13.
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *The Behavioral and brain sciences*, *35*(3), 121–143.
- Luck, S., & Vogel, E. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Lupp, U., Hauske, G., & Wolf, W. (1976). Perceptual latencies to sinusoidal gratings. *Vision Research*, *16*, 969–972.

- Marrett, N. E., de-Wit, L. H., Roser, M., Kenridge, R. W., Milner, A. D., & Lambert, A. J. (2011). Testing the dorsal stream attention hypothesis: electrophysiological correlates and the effects of ventral stream damage. *Visual Cognition*, *19*, 1089–1121.
- Maule, J., & Franklin, A. (2015). Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of Vision*, *15*(4):6.
- Maule, J., & Franklin, A. (2016). Accurate rapid averaging of multi hue ensembles is due to a limited capacity subsampling mechanism. *Journal of the Optical Society of America A*, *33*(3), A22-29.
- McCarthy, G., Puce, A., Gore, J.C., Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *9*(5), 605–10.
- McClelland, T., & Bayne, T. (2016). Ensemble coding and two conceptions of perceptual sparsity. *Trends in Cognitive Sciences*, *20*(9), 641-642.
- Miller, M. B., Kingstone, A., & Gazzaniga, M. S. (2002). Hemispheric Encoding Asymmetry is More Apparent Than Real. *Journal of Cognitive Neuroscience*, *14*(5), 702-708
- Milner A.D., Goodale M.A. Two visual systems re-viewed. *Neuropsychologia* 2008;46:774–785.
- Morgan, M., Chubb, C., & Solomon, J. A. (2008). A ‘dipper’ function for texture discrimination based on orientation variance. *Journal of Vision*, *8*(11), 1–8.
- Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., et al. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, *383*, 812–815.
- Nassi, J. J., Lyon, D. C., & Callaway, E. M. (2006). The parvocellular LGN provides a robust disynaptic input to the visual motion area MT. *Neuron*, *50*, 319–327.
- Nassi, J. J., & Callaway, E. M. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience*, *10*(5), 360-372.
- Niebauer, C. L., & Christman, S. D. (1998). Upper and lower visual field differences in categorical and coordinate judgments. *Psychonomic Bulletin & Review*, *5*, 147–151.
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, *76*(1), 109-129.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Peyrin C., Michel C. M., Schwartz S., Thut G., Seghier M., Landis T., et al. (2010). The neural substrates and timing of top-down processes during coarse-to-fine categorization of visual scenes: a combined fMRI and ERP study. *Journal of Cognitive Neuroscience*, *22*, 2768–2780.
- Pessoa, L. (2010). Emotion and cognition and the amygdala: from “what is it?” to “what’s to be done?” *Neuropsychologia*, *48*, 3416–3429.
- Previc, F. H. (1990). Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral & Brain Sciences*, *13*, 519–575.
- Previc, F. H., & Blume, J. L. (1993). Visual search asymmetries in three-dimensional space. *Vision Research*, *33*(18), 2697–2704
- Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition*, *7*, 17-42.
- Rosa, M. G., Palmer, S. M., Gamberini, M., Burman, K. J., Yu, H. H., Reser, D. H., et al. (2009). Connections of the dorsomedial visual area: pathways for early integration of dorsal and ventral streams in extrastriate cortex. *Journal of Neuroscience*, *29*(14), 4548-4563.
- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, *126*, 2381-2395.
- Sauseng, P., & Klimesch, W. (2008). What does phase information of oscillatory brain activity tell us about cognitive processes? *Neuroscience & Biobehavioral Reviews*, *32*(5), 1001-1013.
- Schroeder, C. E., Mehta, A. D., & Givre, S. J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cerebral Cortex*, *8*(7), 575–592.
- Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: evidence for time-and spatial-scale-dependent scene recognition. *Psychological Science*, *5*, 195–200.
- Sciberras-Lim, E. T., & Lambert, A. J. (2017). Attentional Orienting and Dorsal Visual Stream Decline: Review of Behavioral and EEG Studies. *Frontiers in Aging Neuroscience*, *9*, 246.
- Siegel, M., Donner, T. H., Oostenveld, R., Fries, P., & Engel, A. K. (2008). Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron*, *60*, 709–719.
- Simões, C., Jensen, O., Parkkonen, L., & Hari, R. (2003). Phase locking between human primary and secondary somatosensory cortices. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 2691–2694.
- Sincich, L. C., Park, K. F., Wohlgenuth, M. J., & Horton, J. C. (2004). Bypassing V1: a direct geniculate input to area MT. *Nature Neuroscience*, *7*, 1123–1128.
- Snyder, L. H., Batista, A. P., & Andersen, R. A. (2000). Intention-related activity in the posterior parietal cortex: a review. *Vision Research*, *40*, 1433-1441.
- Styliadis, C., Ioannides, A.A., Bamidis, P.D., Papadelis, C. (2014). Amygdala responses to valence and its interaction by arousal revealed by MEG. *International Journal of Psychophysiology*, *93*(1), 121–33.
- Sweeny, T. D. & Whitney, D. (2014). Perceiving crowd attention: ensemble perception of a crowd’s gaze. *Psychological Science*, *25*, 1903–1913.
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience*, *16*(13), 4240-4249.
- Tapia, E., & Breitmeyer, B. G. (2011). Visual consciousness revisited: magnocellular and parvocellular contributions to conscious and non-conscious vision. *Psychological Science*, *22*, 934–942.
- Tesche, C.D., Uusitalo, M.A., Ilmoniemi, R.J., Huotilainen, M., Kajola, M., Salonen, O. (1995). Signal-space projections of MEG data characterize both distributed and well-localized neuronal sources. *Electroencephalography and Clinical Neurophysiology*, *95*(3), 189–200
- Thomas, C., Kveraga, K., Huberle, E., Karnath, H. O., & Bar, M. (2012). Enabling global processing in simultanagnosia by psychophysical biasing of visual pathways. *Brain*, *135*, 1578-1585.
- Thomas, N. A., & Elias, L. J. (2011). Upper and lower visual field differences in perceptual asymmetries. *Brain Research*, *1387*, 108-115.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, *428*(6984), 751-754.
- Todd, J. J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cognitive, Affective, & Behavioral Neuroscience*, *5*, 144–155.
- Toth, L. J., & Assad, J. A. (2002). Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature*, *415*(6868), 165-168.
- Tootell, R. B., Silverman, M. S., Hamilton, S. L., Switkes, E., & De Valois, R. L. (1988). Functional anatomy of macaque striate cortex: V. Spatial frequency. *Journal of Neuroscience*, *8*, 1610–1624.
- Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual review of neuroscience*, *31*, 411–437.
- Utochkin, I. S. & Tiurina, N. A. (2014). Parallel averaging of size is possible but range-limited: a reply to Marchant, Simons, and De Fockert. *Acta Psychologica*, *146*, 7–18.

- Utochkin, I. S., & Vostrikov, K. O. (2017). The numerosity and mean size of multiple objects are perceived independently and in parallel. *PLoS ONE*, *12*(9), e0185452.
- Uusitalo, M.A., & Ilmoniemi, R.J. (1997). Signal-space projection method for separating MEG or EEG into components. *Medical & Biological Engineering & Computing*, *35*, 135–140.
- Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cerebral Cortex*, *4*, 470–483.
- Wolfe, J. M., Võ, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*, 77–84.
- Xu, Y. (2018a). The posterior parietal cortex in adaptive visual processing. *Trends in Neuroscience*, *41*, 806–822.
- Xu, Y. (2018b). A tale of two visual systems: Invariant and adaptive visual information representations in the primate brain. *Annual Review in Vision Science*, *4*, 311–336.
- Xu, Y. & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, *440*, 91–95.
- Xu, Y. & Jeong, S. (2015). The contribution of human superior intraparietal sulcus to visual short-term memory and perception. In *Mechanisms of Sensory Working Memory: Attention and Performance XXV*, Jolicoeur, P. and Martinez-Trujillo, J.
- Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli. *Brain Research Reviews*, *41*(1), 88–123.
- Zhong, Y. M., & Rockland, K. S. (2003) Inferior Parietal Lobule Projections to Anterior Inferotemporal Cortex (Area TE) in Macaque Monkey. *Cerebral Cortex*, *13*(5), 527–540.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.