

# The bottom-up and top-down processing of faces in the human occipitotemporal cortex

Xiaoxu Fan<sup>1,2</sup>, Fan Wang<sup>1,2</sup>, Hanyu Shao<sup>1</sup>, Peng Zhang<sup>1,2</sup>, Sheng He<sup>1,2,3\*</sup>

<sup>1</sup>State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China; <sup>2</sup>University of Chinese Academy of Sciences, Beijing, China; <sup>3</sup>Department of Psychology, University of Minnesota, Minneapolis, United States

**Abstract** Although face processing has been studied extensively, the dynamics of how face-selective cortical areas are engaged remains unclear. Here, we uncovered the timing of activation in core face-selective regions using functional Magnetic Resonance Imaging and Magnetoencephalography in humans. Processing of normal faces started in the posterior occipital areas and then proceeded to anterior regions. This bottom-up processing sequence was also observed even when internal facial features were misarranged. However, processing of two-tone Mooney faces lacking explicit prototypical facial features engaged top-down projection from the right posterior fusiform face area to right occipital face area. Further, face-specific responses elicited by contextual cues alone emerged simultaneously in the right ventral face-selective regions, suggesting parallel contextual facilitation. Together, our findings chronicle the precise timing of bottom-up, top-down, as well as context-facilitated processing sequences in the occipital-temporal face network, highlighting the importance of the top-down operations especially when faced with incomplete or ambiguous input.

\*For correspondence:  
sheng@umn.edu

**Competing interests:** The authors declare that no competing interests exist.

**Funding:** See page 15

**Received:** 24 May 2019

**Accepted:** 10 January 2020

**Published:** 14 January 2020

**Reviewing editor:** Ming Meng, South China Normal University, China

© Copyright Fan et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

## Introduction

There is ample evidence to show that the processing of face information involves a distributed neural network of face-sensitive areas in the occipitotemporal cortex and beyond (*Duchaine and Yovel, 2015; Haxby et al., 2000*). Three bilateral face-selective areas are considered as the core face-processing system, defined in functional Magnetic Resonance Imaging (fMRI) studies as regions showing significantly higher response to faces than objects, which are Occipital Face Area (OFA) in the inferior occipital gyrus (*Gauthier et al., 2000; Haxby et al., 1999*), Fusiform Face Area (FFA) in the fusiform gyrus (*Kanwisher et al., 1997; Grill-Spector et al., 2004*) and a face-sensitive area in the posterior superior temporal sulcus (pSTS) (*Hoffman and Haxby, 2000; Puce et al., 1998*). Similarly, a number of so-called face patches have been identified in macaque monkeys along the superior temporal sulcus (*Tsao et al., 2003; Tsao et al., 2006; Tsao et al., 2008*). Although the functional properties of these areas have been studied extensively, we do not yet have a comprehensive understanding of how the face-processing network functions in a dynamic manner. Hierarchical models postulate that face specific processes are initiated in the OFA based on local facial features, then the information is forwarded to higher level regions, such as FFA, for holistic processing (*Haxby et al., 2000; Fairhall and Ishai, 2007; Liu et al., 2002*). This model is supported by neuroimaging studies showing functional properties of face-selective areas and is consistent with generic local-to-global views of object processing. However, it has been challenged by results from studies in which patients with damaged OFA can still showed FFA activation to faces (*Rossion et al., 2003; Steeves et al., 2006*). Further, it was reported that during the perception of faces with minimal local facial features,

FFA could still show face-preferential activation without face-selective inputs from OFA (Rossion *et al.*, 2011). Thus a non-hierarchical model was proposed postulating that face detection is initiated at the FFA followed by a fine analysis in the OFA (Rossion *et al.*, 2011; Gentile *et al.*, 2017). These competing models may reflect different modes of operation of the face network under different demands. To reconcile these models, a comprehensive dynamic picture of face processing under different conditions with more detailed temporal information is needed.

In the current study, we investigated the dynamics of face processing in the 'core face processing system' using Magnetoencephalography (MEG) and fMRI. We designed the face-related stimuli specifically to reveal mechanisms for processing 1) normal faces, 2) Mooney faces with very little explicit facial features, 3) distorted faces with internal facial features spatially misarranged, and 4) contextually induced face representations with internal facial features completely missing. During the experiment, subjects were presented with various types of face pictures while MEG signals were recorded. The key effort in this study was in reconstructing the source signals from the MEG sensor data, to obtain a dynamic depiction of cortical responses to faces and other types of stimuli. With the timing of activation revealed in each face-selective area in the 'core face processing system', we could uncover when and where face information is processed in the human brain.

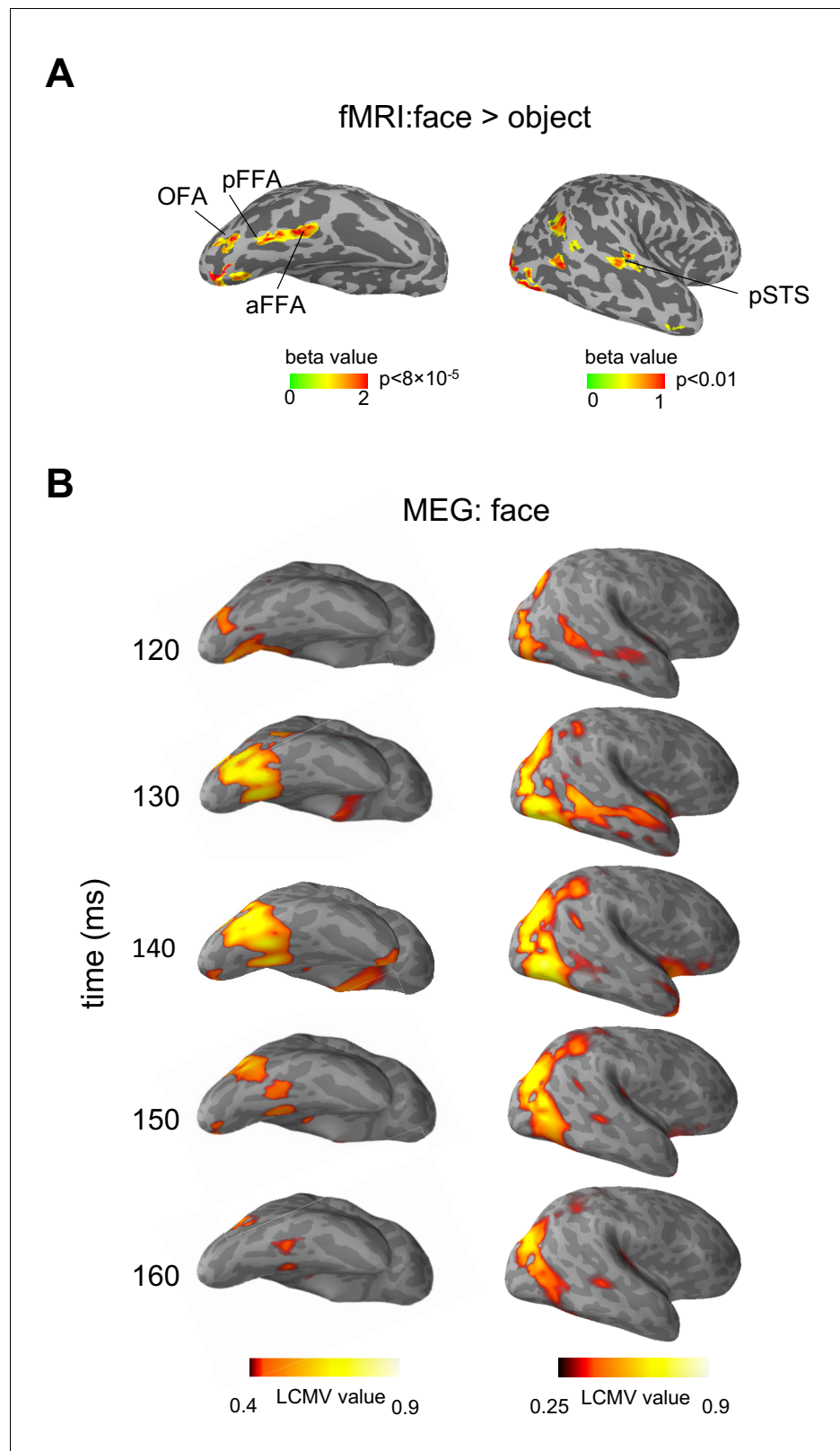
The main findings are briefly summarized here. First, we revealed the basic, mainly bottom-up, processing sequence along ventral temporal cortex by presenting face pictures of famous individuals to subjects. Face processing was initiated in the posterior areas and then proceeded forward to anterior regions. Right OFA (rOFA) and right posterior FFA (rpFFA) were activated very close in time, peaking around 120 ms, while right anterior FFA (raFFA) reached its peak at about 150 ms. The right pSTS (rpSTS) in the dorsal pathway showed a weaker and temporally more variable response, participating in face processing within a time window from 130 to 180 ms. Then, we highlighted the top-down operation in face processing by using two-tone Mooney face images (Mooney, 1957) lacking prototypical local facial features. According to the predictive coding theory (Rao and Ballard, 1999; Murray *et al.*, 2004; Mumford, 1992), face prediction created at FFA based on impoverished information of Mooney faces and prior knowledge is poorly matched with the input representation at OFA due to the lack of explicit local facial features. The activity in OFA, representing 'residual error' between top-down prediction and bottom-up input, is then expected to increase subsequently. Consistent with this model, rOFA was activated later than rpFFA, and rpFFA exerted extensive directional influence onto rOFA when processing Mooney faces, suggesting a cortical analysis dominated by rpFFA to rOFA projection. However, when explicit internal facial features were available but misarranged within a normal face contour, a temporal pattern similar to that of normal faces was observed. Finally, we further investigated the temporal dynamics when face-specific responses were driven by contextual cues alone with the internal face features entirely missing (Cox *et al.*, 2004). In this case, rOFA, rpFFA and raFFA were activated somewhat late and almost simultaneously, corresponding to contextual modulation that parallelly facilitated the processing of the core face-processing network.

## Results

### Face induced MEG signals in the source space

Subjects were presented with famous faces and familiar objects and instructed to perform a simple classification task (face or object) while their brain activity was recorded using MEG. After a rest period, each subject was scanned with fMRI viewing the same group of face and object images presented in separate blocks. Since each subject underwent both fMRI and MEG measurements, we could compare the face-selective regions defined by fMRI with the reconstructed MEG signals evoked by faces in the source space.

Subjects' face-selective regions in the occipitotemporal cortex were localized with fMRI contrasting responses to faces with that to objects. MEG signals at different time points were reconstructed in the source space by computing LCMV beamformer solution on evoked data after preprocessing (Van Veen *et al.*, 1997). The estimated activities for the whole cortical surface can be viewed as a 3D spatial distribution of LCMV value (power normalized with noise) at each time point (Sekihara and Nagarajan, 2008).



**Figure 1.** Face-selective areas identified by fMRI localizer and face-evoked MEG source activation displayed on an inflated right hemisphere of a typical subject. (A) Face-selective statistical map (faces>objects) showing four face-selective regions (rOFA, rpFFA, raFFA and rpSTS). (B) Face-evoked MEG source activation patterns represented as *Figure 1 continued on next page*

Figure 1 continued

LCMV value maps at different time points (120–160 ms) after the stimulus onset. LCMV values represent signal power normalized by noise.

The online version of this article includes the following figure supplement(s) for figure 1:

**Figure supplement 1.** Face-selective areas identified by fMRI localizer and face-evoked MEG source activation displayed on an inflated right hemisphere of typical subject one.

**Figure supplement 2.** Face-selective areas identified by fMRI localizer and face-evoked MEG source activation displayed on an inflated right hemisphere of typical subject two.

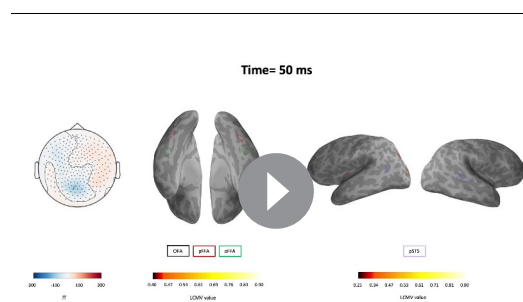
**Figure supplement 3.** Face-selective areas identified by fMRI localizer and face-evoked MEG source activation displayed on an inflated right hemisphere of typical subject three.

**Figure supplement 4.** Face-selective areas identified by fMRI localizer and face-evoked MEG source activation displayed on an inflated right hemisphere of typical subject four.

**Figure 1** shows the fMRI identified face regions and MEG measured face-evoked signals in a typical subject, displayed in ventral and lateral views of an inflated right hemisphere (Source localization results and fMRI localization results are shown in **Figure 1—figure supplement 1–4** for more individual subjects). Face-selective regions rOFA, rpFFA, raFFA and rpSTS were identified by fMRI localizer (**Figure 1A**). MEG responses evoked by faces are shown in 10 ms steps from 120 ms to 160 ms in source space (cortical surface) (**Figure 1B**). It could be seen in the MEG signal that the location of a cluster of activation in the right occipital cortex at about 120 ms after stimulus onset is consistent with rOFA. At about 150–160 ms, a cluster of activation was found in posterior part of superior temporal sulcus, overlapping with rpSTS. Two temporally separated clusters of MEG source activation were found in the right fusiform gyrus, one consistent with the location of pFFA (about 130 ms) and another with aFFA (about 150 ms) (see **Video 1**). Similar spatiotemporal patterns of activation could be seen across the 13 subjects tested. These results show that face response areas identified by MEG are highly consistent with that defined by fMRI, thus it is a reasonable approach to extract the MEG time courses based on fMRI-guided region of interest (ROI). In this paper, with the understanding that the sources of MEG signals were constrained by the fMRI defined ROIs, we use the fMRI terms (OFA, FFA and pSTS) to indicate the corresponding cortical area in MEG data.

## Bottom-up processing sequence induced by normal faces

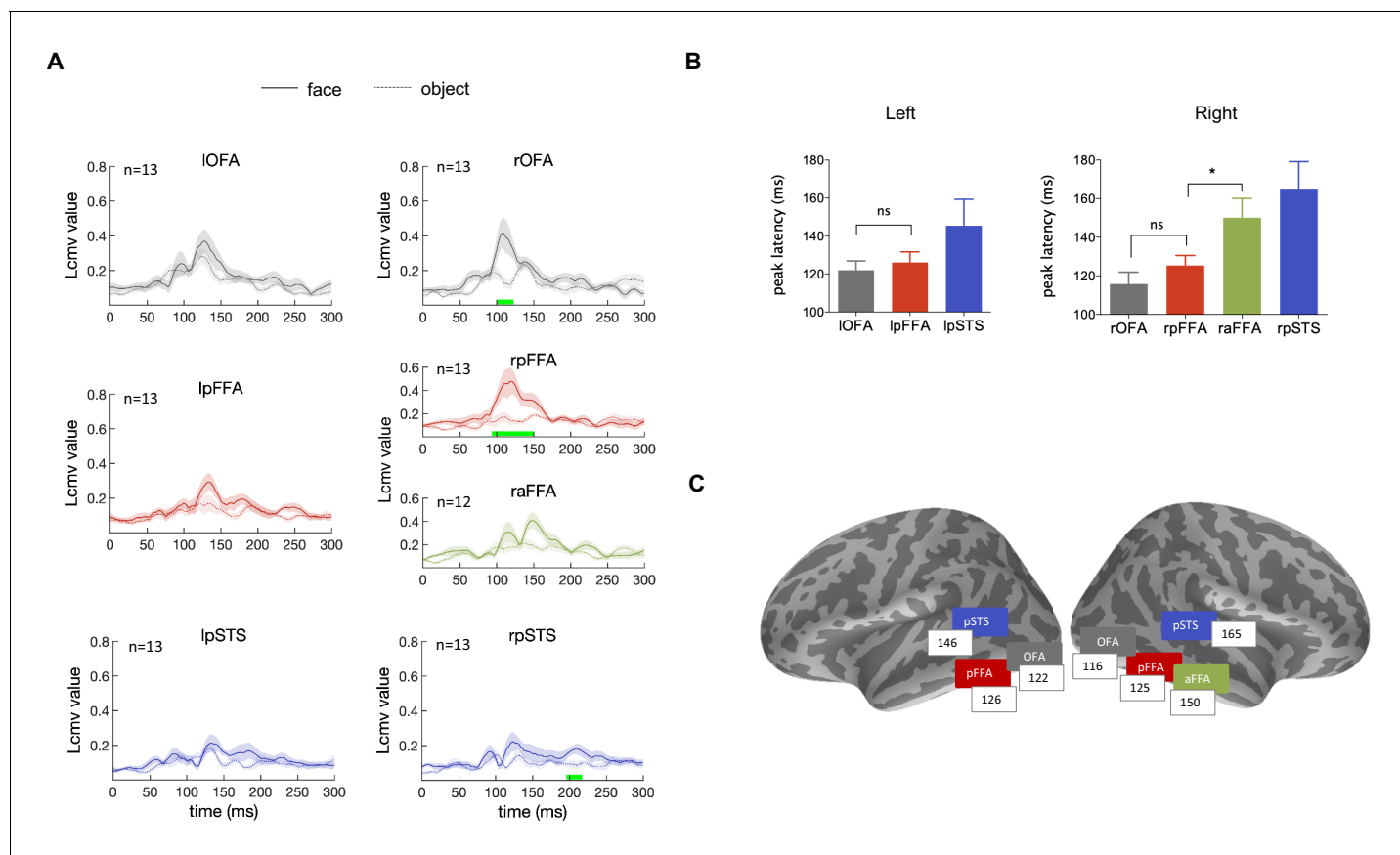
We investigated the typical dynamic sequence for processing faces in the ventral occipitotemporal cortex investigated by presenting subjects with face images of well-known individuals. We analyzed the time courses of face-selective areas identified in the source space. Seven face-selective areas (IOFA, rOFA, lpFFA, rpFFA, raFFA, lpSTS, rpSTS) were identified, guided by fMRI face localizer results from each individual subject, and they were used to extract the face-response time courses of the MEG source data. We averaged the resulting time courses across subjects and the waveforms are shown in **Figure 2A**. Face images induced stronger responses compared to objects in face-selective areas, especially for the right hemisphere. The timing of peak responses for individual ROIs are summarized in **Figure 2B and C**, revealing the fundamental temporal characteristics of the neural



**Video 1.** MEG activation of a typical subject.

<https://elifesciences.org/articles/48764#video1>

processing of faces. In the right hemisphere, face-evoked responses emerged earlier in the posterior areas than in the anterior areas, the peak responses occurred at  $116 \pm 6$  ms,  $125 \pm 5$  ms and  $150 \pm 10$  ms for rOFA, rpFFA and raFFA, respectively. Although there is no significant difference between rOFA and rpFFA ( $t_{12} = 1.57$ ,  $p=0.43$ , Bonferroni corrected), the peak response timing of raFFA is significantly delayed compared with rpFFA ( $t_{11} = 3.21$ ,  $p=0.025$ , Bonferroni corrected), suggesting a bottom-up process. Similarly, OFA reached its peak response earlier than pFFA in the left hemisphere (IOFA:  $122 \pm 5$  ms, lpFFA:  $126 \pm 6$  ms), although this trend is not statistically significant



**Figure 2.** Temporal response characteristics of face-selective ROIs. (A) The time courses of face (solid line) and object (dotted line) induced responses averaged across subjects, for the seven face-selective ROIs. Shaded area means SEM. The green bar indicates significant difference between face and object. Significance was assessed by cluster-based permutation test (cluster-defining threshold  $p < 0.05$ , significance level  $p < 0.05$ ) for each ROI. (B) The peak latency averaged across subjects for each ROI (mean  $\pm$  SEM). The peak latency of raFFA is significantly later than rpFFA ( $t_{11} = 3.21$ ,  $p = 0.025$ , Bonferroni corrected) (C) The mean peak latencies for the face-selective ROIs were shown on inflated cortical surfaces of both hemispheres at corresponding locations.

The online version of this article includes the following figure supplement(s) for figure 2:

**Figure supplement 1.** Temporal response characteristics of face-selective ROIs for unfamiliar faces.

( $t_{11} = 0.64$ ,  $p > 0.05$ , Bonferroni corrected). Responses from the left anterior FFA was not shown because the corresponding activation cluster was not observed clearly in most subjects. In addition, dorsal face-selective region pSTS showed weaker and temporally broader responses, involved in face processing roughly from 130 to 180 ms. The sequential progression from posterior to anterior regions along the ventral occipitotemporal cortex, especially the significantly delayed activation of raFFA, indicates a bottom-up hierarchical functional structure of the ventral face pathway.

In addition to famous faces, we also presented unfamiliar faces to subjects and analyzed the data in the same way. Results showed essentially similar hierarchical dynamic sequences of face processing regardless of face familiarity (Figure 2—figure supplement 1). Thus, unfamiliar face images were used in the next experiment reported below.

## Top-down operation in face processing highlighted by viewing Mooney faces

While the processing of normal (famous or unfamiliar) faces mainly followed the posterior to anterior (bottom-up) face processing sequence, we further investigated the possibility that under certain stimulus conditions, top-down modulation of face processing could become more prominent. According to the predictive coding theory, when the representation of sensory input in lower areas is poorly matched with the predictions generated from higher level areas, the activity in lower areas

representing residual error would be increased (Rao and Ballard, 1999; Murray et al., 2004; Mumford, 1992). Hence, we adopted the two-tone Mooney face images (Figure 3A), which could be recognized as faces but lack prototypical local facial features, as the main stimuli in this experiment. Our hypothesis was that when processing Mooney faces which could activate the FFA based on the global configuration, the top-down modulation from FFA to OFA (prediction of facial parts) would be more prominent.

In this experiment, subjects ( $n = 28$ ) were presented with normal unfamiliar faces and Mooney faces, they performed a one-back task, indicating the repetition of the same images. Verbal survey after the MEG experiment indicated that subjects could perceive at least 90% of the Mooney images as faces. In all face-selective areas except rOFA, similar peak latencies were observed during the perception of normal and Mooney faces. Strikingly, Mooney face elicited a response with significantly longer latency in rOFA than normal face (paired  $t$  test,  $t_{23} = 4.009$ ,  $p = 0.001$ ) (Figure 3B). The temporal relationship of signals in the face-selective areas was quite different during the perception of Mooney faces compared with that of normal ones (Figure 3C). Similar to Experiment 1, OFA were activated slightly earlier than pFFA in response to normal faces (IOFA:  $124 \pm 9$  ms, IpFFA:  $133 \pm 9$  ms, Paired permutation test  $p > 0.9$ ; rOFA:  $107 \pm 4$  ms, rpFFA:  $120 \pm 6$  ms, Paired permutation test  $p = 0.37$ , Bonferroni corrected for multiple comparisons). However, when processing Mooney faces, rOFA was engaged significantly later than rpFFA (rOFA:  $144 \pm 8$  ms, rpFFA:  $117 \pm 8$  ms, Paired permutation test  $p = 0.02$ , Bonferroni corrected). The response curve of rOFA was temporally shifted to a later point while the temporal characteristics of rpFFA was not much different from its response to normal faces (Figure 3C). The temporal relationship between OFA and pFFA in left hemisphere is similar to normal face condition (IOFA:  $127 \pm 10$  ms, IpFFA:  $133 \pm 10$  ms, Paired permutation test  $p > 0.9$ , Bonferroni corrected).

To further analyze the dynamic causal relationship between OFA and pFFA, we performed Granger causality analysis over sliding time windows of 50 ms duration from 75 to 230 ms after stimulus presentation which covers the periods of essential activation in OFA and pFFA. The significant directed connectivity in each time window is shown in Figure 3D. There were much more extensive directed influences from pFFA to OFA during the processing of Mooney than normal faces. In particular, rpFFA influenced rOFA in Mooney face condition continuously from 75 to 170 ms, which was more sparsely observed in normal face condition. Thus response time courses and Granger causality analysis together show that, compared with processing of normal faces, the cortical processing of Mooney faces is more dominated by the top-down rpFFA to rOFA projection.

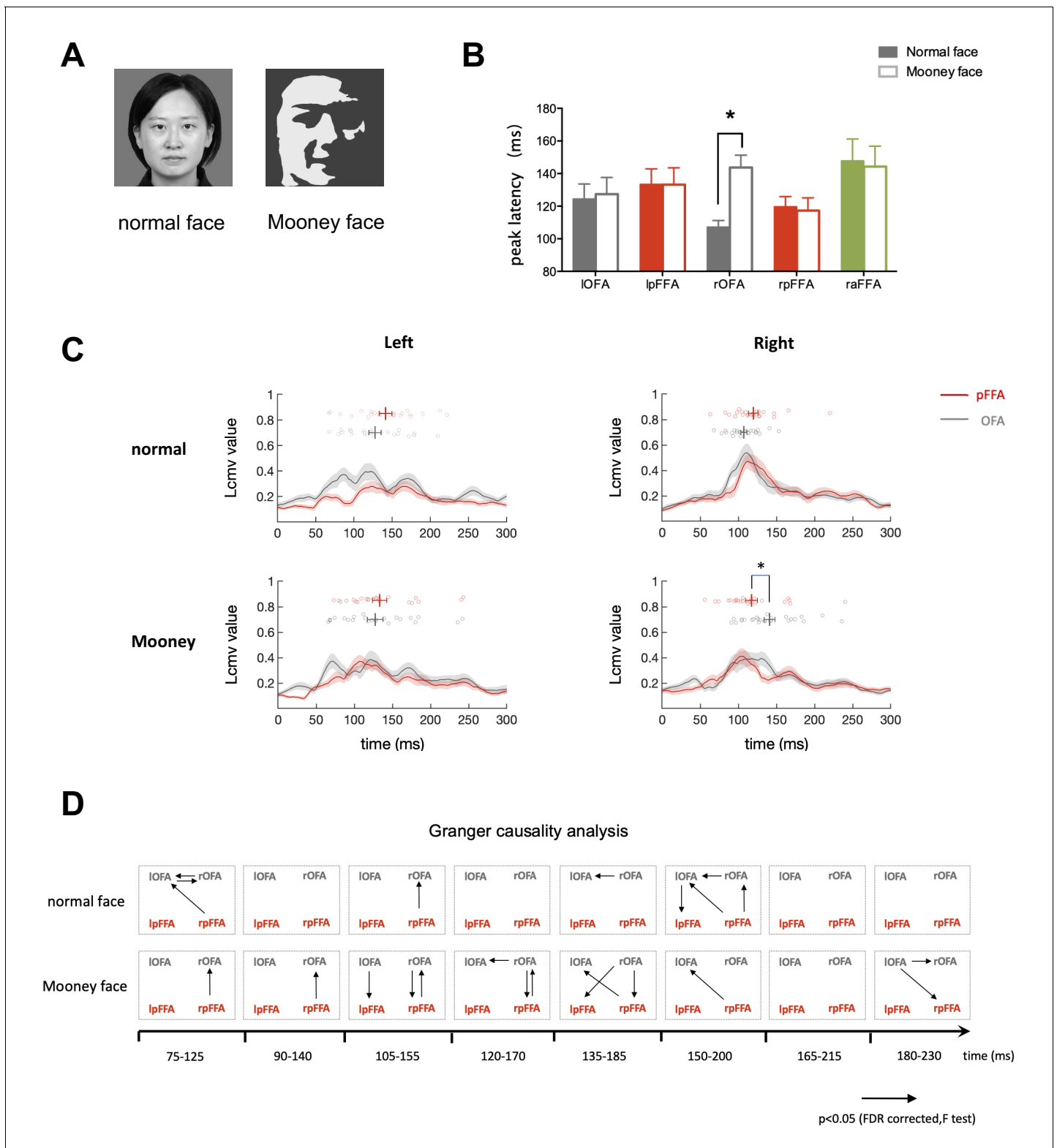
### Primarily feedforward processing of face-like stimuli with misarranged internal features

We also investigated the processing dynamics of face-like stimuli with internal features clearly available but spatially misarranged, to contrast with the processing of normal as well as Mooney faces. The normal external features (hair, chin, face outline) and the locally normal internal features led to the engagement of the face-sensitive areas. Results show that the rOFA, rpFFA and raFFA were activated sequentially (rOFA:  $132 \pm 7$  ms, rpFFA:  $133 \pm 5$  ms, raFFA:  $169 \pm 12$  ms, Figure 4B). Compared with the responses to normal faces, the activations in the rOFA and rpFFA were somewhat delayed in the case of the distorted faces. However, unlike the Mooney faces, the distorted faces still engaged the OFA earlier than the FFA, presumably because of the explicitly available local facial features. While the dominant signals are consistent with a feedforward processing from OFA to FFA, there was a hint of a predictive error signal, possibly related to the misarranged spatial configurations, that produced a low activity in rOFA at a later stage.

### Parallel facilitation of face-processing network from contextual cues alone

In real life, facial features are not always available. Previous studies showed that face-specific responses could be elicited by contextual body cues (Cox et al., 2004; Chen and Whitney, 2019; Martinez, 2019). Here we further investigated the dynamics of contextual facilitation of face processing when face perception was supported by contextual cues alone without explicit facial features using the same experimental paradigm and data analysis procedures as before.

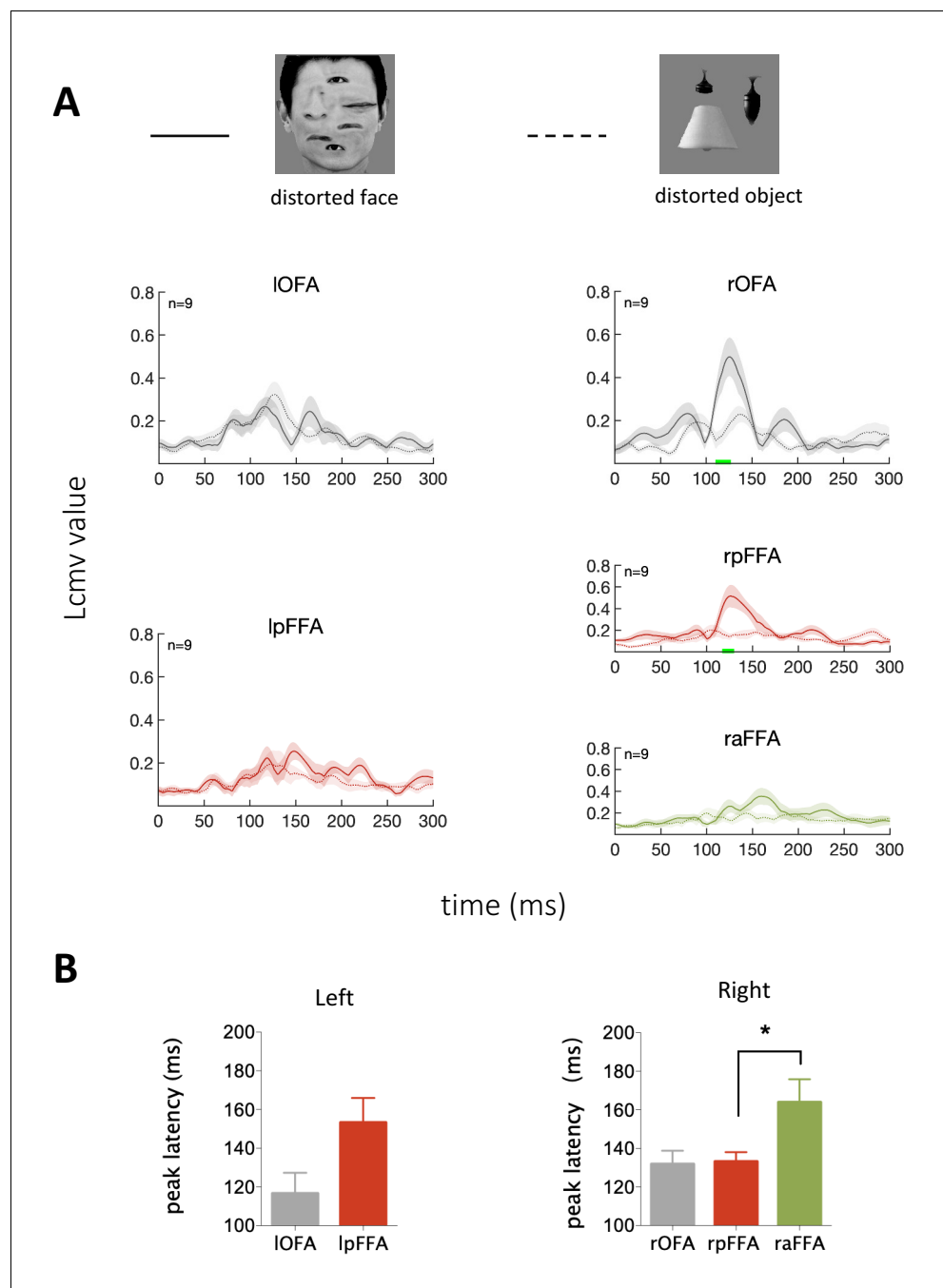




**Figure 3.** Temporal response characteristics and granger causality analysis for face-selective ROIs during perception of Mooney and normal faces. (A) Normal and Mooney face images. (B) The peak latency averaged across subjects for each face-selective ROI (mean  $\pm$  SEM). Mooney faces elicited a response with significantly longer latency in rOFA than normal faces (paired t test,  $t_{23} = 4.009$ ,  $p = 0.001$ ). (C) Time courses averaged across subjects for bilateral OFA and pFFA. Gray line is OFA and red line is pFFA. Shaded areas denote SEM. The circles above time courses represent peak latencies of individual subjects. rOFA was engaged significantly later than rpFFA when processing Mooney faces (Paired permutation test  $p = 0.02$ , Bonferroni). *Figure 3 continued on next page*

Figure 3 continued

corrected). (D) Granger causality analysis performed within a series of 50 ms time windows. Arrows represent statistically significant causal effects ( $p < 0.05$ , FDR corrected, F test. See Materials and methods for details).



**Figure 4.** Temporal response characteristics for face-selective ROIs in response to distorted face. (A) Example stimuli and averaged time courses for each face-selective ROI. The green horizontal bar indicates significant difference between distorted face and object (cluster-defining threshold  $p < 0.01$ , corrected significance level  $p < 0.05$ ). (B) Peak latency averaged across subjects for each ROI. The peak latency of raFFA is significant later than rpFFA (paired t test,  $p = 0.019$ ,  $t_8 = 2.92$ ).

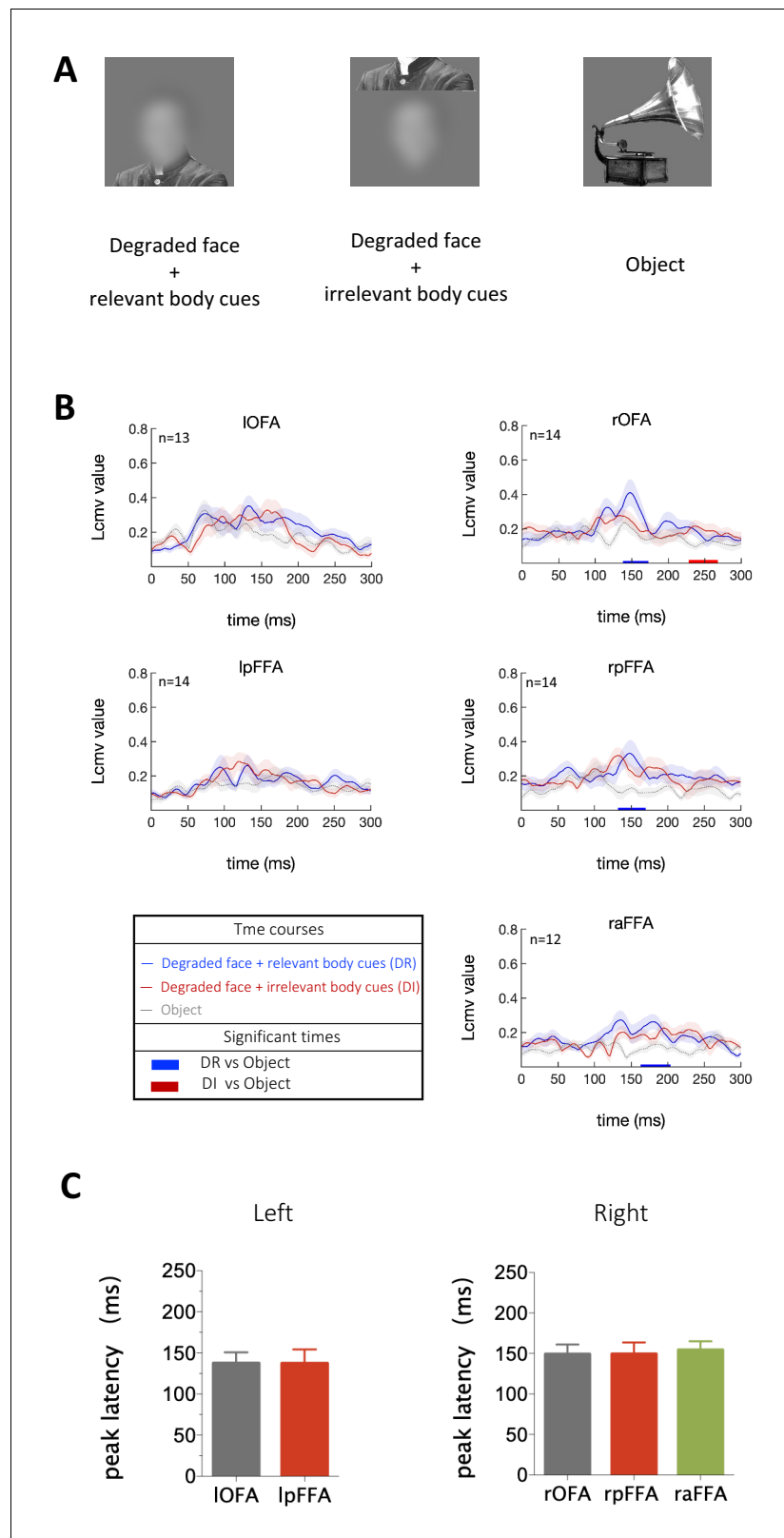


Three types of stimuli were presented to subjects: (i) images of highly degraded faces (no internal facial features) with contextual body cues that imply the presence of faces, (ii) similar to images in (i) but with body cues arranged in an incorrect configuration and thus do not imply the presence of faces, (iii) images of objects (**Figure 5A**). Activation in rOFA, rpFFA and raFFA were significantly higher for the condition in which faces were clearly implied due to the contextual cues compared to the condition when objects were presented (**Figure 5B**). However, when contextual cues were misarranged so that faces were not strongly implied, only the rOFA showed stronger activation than objects at a late stage (**Figure 5B**). Furthermore, peak latency analysis revealed that during the perception of 'faces' generated from contextual cues alone, the rOFA, rpFFA and raFFA were all engaged at about the same and relatively late time (rOFA:  $149 \pm 12$  ms, rpFFA:  $149 \pm 14$  ms, raFFA:  $155 \pm 11$  ms) rather than activated sequentially (**Figure 5C**). Thus when the presence of a face was facilitated by external cues alone, the evoked responses in the core face-processing network emerged slowly and almost simultaneously.

## Discussion

Using a combined fMRI and MEG source localization approach, our results systematically revealed an intricately detailed dynamic picture of face information processing. Within the ventral occipito-temporal face processing network, normal faces were processed mainly in a bottom-up manner through the hierarchical pathway where input information was processed sequentially from posterior to anterior ventral temporal cortex. This temporal order was also observed when processing face-like stimuli with misarranged internal facial features. In contrast, during the processing of Mooney faces in the absence of prototypical facial features, top-down modulation was more prominent in which the dominant information flow was from the rpFFA to rOFA. Moreover, face-specific responses from contextual cues alone were evoked late and simultaneously across the rOFA, rpFFA and raFFA, suggesting that contextual facilitation acted parallelly on the core face-processing network. These results advance our understanding of the hierarchical and non-hierarchical models of face perception, especially underscoring the stimulus- and context-dependent nature of the processing sequences.

During the perception of 2-tone Mooney faces, it is necessary to discount shadows and recover 3D surface structure from 2D images (**Grützner et al., 2010**). Interestingly, only familiar objects, like faces, can be interpreted to be volumetric easily from 2-tone representations (**Moore and Cavanagh, 1998; Hegde et al., 2007**). Thus it is supposed that prior knowledge should play an important role in the recovery of 3D shape from Mooney images (**Braje et al., 1998; Gerardin et al., 2010**). A top-down model emphasized the guidance of prior experience at higher levels (**Cavanagh, 1991**). This model is supported by evidence from experiments showing that early visual processing is affected by high-level attributes in both human and monkey (**Lee et al., 2002; Humphrey et al., 1997; Issa et al., 2018**). As briefly mentioned in the results section, the dynamics of MEG signals associated with processing Mooney faces, which highlights the top-down modulation, is consistent with the explanation based on predictive coding model. It proposed that hypotheses or predictions made at higher cortical areas are compared with, through feedback, representations at lower areas to generate residual error, which is then forwarded to higher stages as 'neural activity' (**Rao and Ballard, 1999; Murray et al., 2004; Friston, 2005; Friston, 2010**). Specifically, the face model/prediction is generated at the rpFFA based on the global configuration of Mooney faces using prior knowledge about 3D faces, illumination, and cast shadows. This prediction of expected facial features is then poorly matched with the input representation at the rOFA which lacks the explicit prototypical facial features due to the mixed illumination-invariant and illumination-dependent features, generating an increased signal at rOFA. Thus, the dominant signal at the rOFA (residual) necessarily lags behind the signal at the rpFFA (hypothesis). However, when processing normal faces or face with misarranged facial features, the prominent signal in the early stage of rOFA is mainly due to the strong feedforward input from early visual cortex as rOFA is robustly responsive to the clear facial components. The prediction feedback from rpFFA would be consistent with representation at the rOFA in the case of the normal faces, resulting in little error signal; with the misarranged facial features, there was a hint of a late increase of rOFA signal, possibly indicating that the feedback signal could contain some spatial information as well.



**Figure 5.** Temporal response characteristics for face-selective ROIs in response to contextual cues. (A) Example stimuli. (B) Time courses averaged across subjects for each condition. For each ROI, Blue horizontal bars indicate significant difference between degraded faces with relevant body cues and objects, and red horizontal bars indicate significant difference between degraded faces with irrelevant body cues and objects (cluster-defining

*Figure 5 continued on next page*

Figure 5 continued

threshold  $p < 0.05$ , corrected significance level  $p < 0.05$ ). (C) The peak latency averaged across subjects for each face-selective ROI (mean  $\pm$  SEM).

The timing of face induced neural activation has been studied for a long time with various techniques, such as the combination of MEG and fMRI using representational similarities (Cichy et al., 2014; Cichy et al., 2016), MEG source localization and intracranial EEG (Kadipasaoglu et al., 2017; Keller et al., 2017; Ghuman et al., 2014). An early MEG study suggested two stages (early categorization and late identification) were involved in face processing (Liu et al., 2002). Combined with the fMRI observation that OFA is responsible for identifying facial parts while FFA for holistic configuration (Rotshtein et al., 2005; Liu et al., 2010; Pitcher et al., 2011b; Arcurio et al., 2012; Pitcher et al., 2007; Schiltz, 2010), OFA is expected to respond earlier than FFA. An simultaneous electroencephalogram (EEG)-fMRI study also showed that OFA responded to faces earlier than FFA (OFA: 110 ms; FFA: 170 ms) (Sadeh et al., 2010). Using transient stimulation to temporally disrupt local neural processing, Transcranial Magnetic Stimulation (TMS) experiments suggested that OFA processes facial information at about 100/110 ms, while pSTS begins processing face at about 100/140 ms (Pitcher et al., 2012; Pitcher et al., 2014). However, the sources of N/M170 face selective component remain controversial, it is suggested to come from fusiform gyrus in some studies (Deffke et al., 2007; Kume et al., 2016; Perry and Singh, 2014). While some other studies emphasized the contribution of inferior occipital gyrus besides fusiform gyrus (Itier et al., 2006; Gao et al., 2013) or even of pSTS (Nguyen and Cunnington, 2014). Our results provide more precise and detailed timing information of the core face network under various stimulus and contextual conditions, especially the temporal relationship between rpFFA and raFFA. raFFA is engaged significantly later, about 20 ms after the rpFFA, suggesting that the raFFA likely plays a different functional role from rpFFA. This idea is supported by previous anatomical evidence showing that pFFA and aFFA have different cellular architectures (Weiner et al., 2017).

Our results also shed light on the role of internal and external features in face perception. Although when assembling into a whole face, facial features are processed holistically and the representation of internal features are influenced by external features (Andrews et al., 2010), eyes in isolation elicit a later but larger N170 (Bentin et al., 1996; Rossion and Jacques, 2011) and can drive face-selective neurons as well as full-face images (Issa and DiCarlo, 2012) in monkeys. In our results, the somewhat slower but still sequential progression of face responses elicited by face-like stimuli with clear but misarranged internal features in face outline further supports that facial features are sufficient to trigger the bottom-up face processing sequence. In addition, certain stimulus manipulations, such as face inversion (Bentin et al., 1996), contrast reversal, Mooney transformation or removal of facial features produced comparable (or even increased amplitude) but delayed N170 responses (Rossion and Jacques, 2011). Thus it is suggested that as long as the impoverished stimuli is perceived as a face, inferior temporal cortex areas would be activated (McKeeff and Tong, 2007; Grützner et al., 2010). Our results provide further more details for this explanation by showing the top-down rpFFA to rOFA projection when the prototypical facial features are lack.

Besides facial features, contextual information is also important for face interpretation (Chen and Whitney, 2019; Martinez, 2019). Interestingly, FFA can be activated by the perceived presence of faces from contextual body cues alone (Cox et al., 2004). Here our MEG data showed that the face-selective areas in ventral core face network were indeed activated by the contextual cues for faces, but they were not activated in any order, instead, they became active together at a late stage. This is similar to the temporal dynamics observed in visual imagery, a top-down process given the absence of visual inputs (Dijkstra et al., 2018). Future studies are needed to elucidate how core face network interacts with other brain regions to trigger the face perception. For example, according to a MEG study using fast periodic visual stimulation approach (Rossion et al., 2012; Rossion et al., 2015; de Heering and Rossion, 2015), top-down attention increase the response in FFA by gamma synchrony between the inferior frontal junction and FFA (Baldauf and Desimone, 2014).

Face perception is shaped by long-term visual experience, for example, familiar faces are processed more efficiently than unfamiliar ones (Landi and Freiwald, 2017; Schwartz and Yovel, 2016; Dobs et al., 2019; Gobbini and Haxby, 2006). In terms of the dynamics in the ventral

occipitotemporal areas, the present results showed little differences between processing famous and unfamiliar faces. This could be due to several reasons. First, many studies suggested that regions in the anterior temporal lobe rather than OFA and FFA represent face familiarity (Gobbini and Haxby, 2007; Pourtois et al., 2005; Sugiura et al., 2011). However, extended face system is beyond the scope of our current study because some areas in the extended system are too deep to obtain a good MEG source signal. Second, some subjects might not be familiar with all the famous faces we used. Third, familiarity may affect face recognition via high gamma frequency band activity (Anaki et al., 2007), which is not included in our data analysis.

Bilateral pSTS showed weak and multi-peaked responses during both famous and unfamiliar face processing despite the task differences. One possible reason for the multiple peaks of responses is that as a hub for integrating information from multiple sources (e.g., face, body, and voice), STS contains regions that respond to different types of information (Grossman et al., 2005; Bernstein and Yovel, 2015). A lot of studies have suggested diverse functional role of pSTS in representing changeable aspects of faces, such as expression, lip movement and eye-gaze (Baseler et al., 2014; Engell and Haxby, 2007). Specifically, pSTS is involved in the analysis of facial muscle articulations which are combined to produce facial expressions (Srinivasan et al., 2016; Martinez, 2017). In addition, pSTS may respond to dynamic motion information conveyed through faces (O'Toole et al., 2002).

Previous studies showed that left and right fusiform gyrus are differentially involved in face/non-face judgements (Meng et al., 2012; Goold and Meng, 2017), 'low-level' face semblance and perceptual learning of face (Bi et al., 2014; Feng et al., 2011; McGugin et al., 2018). Interestingly, in our results, the peak latency of the left pFFA was later than that of the right pFFA in all conditions except famous face. Responses evoked from distorted faces with misarranged features had the largest lateral difference (20 ms). One possible reason is that the signal attributed to the left pFFA is in fact a mixture of signals from pFFA and aFFA.

Although the exact correspondence between human and macaque face-selective areas are still unclear (Tsao et al., 2003; Tsao et al., 2006; Tsao et al., 2008), the dynamic picture of normal face processing revealed in our study is generally similar to that in macaques. Single-unit recording studies showed that activity begins slightly earlier in posterior face patches than anterior ones, reaching peak levels around 126, 133, and 145 ms for middle lateral (ML)/middle fundus (MF), anterior lateral (AL), and anterior medial (AM) (Freiwald and Tsao, 2010), respectively. Interestingly, there is a discrepancy in response to Mooney faces in high level face patch AM between two monkeys. One of them showed nearly the same peak latency as normal faces but with more sustained activation, while the other did not respond to Mooney faces (Moeller et al., 2017). This may imply that the processing of Mooney faces is related to individual face detection ability or life experience and face processing is not a simple feedforward process from low level to high level areas. Consistent with that, a more recent study showed a rapid and more sustained response in high level face area (aIT) and an early rising then quickly decreased activity in low level areas in monkeys, a signature of predictive coding model (Issa et al., 2018).

Our study is obviously limited in scope. There are many types of cues and tasks relevant for face perception that could be investigated. In addition to facial features and context, many low level cues contribute to face recognition, such as illumination direction, pigmentation (surface appearance) and contrast polarity (one region brighter than another) (Russell et al., 2007; Sinha et al., 2006). In particular, neurons tuned for contrast polarity were found in macaque inferotemporal cortex, supporting the notion that low-level image properties are encoded in face regions (Ohayon et al., 2012; Weibert et al., 2018). We purposely avoided the complication of color cues in this study by using gray-scale images, but we are aware the importance of color in face perception (Yip and Sinha, 2002; Benitez-Quiroz et al., 2018). Moreover, the temporal dynamics of face processing could very well be influenced by different tasks. In our results, there is little difference between the temporal patterns in response to unfamiliar faces under face category task (Figure 2—figure supplement 1) and image identity one-back task (Figure 3). Future studies are needed to more comprehensively investigate the role of behavioral tasks, especially during the relatively late stages of face processing.

In summary, our study delineated the precise timing of bottom-up, top-down, as well as context-facilitated processing sequences in the occipital-temporal face network. These results provide a way to understand and reconcile previous discrepant findings, revealing the dominant bottom-up

processing when explicit facial features were present, and highlighting the importance of the top-down feedback operations when faced with impoverished inputs with unclear or ambiguous facial features.

## Materials and methods

### Participants

All subjects (age range 19–31) provided written informed consent and consent to publish before the experiments, and experimental protocols were approved by the Institutional Review Board of the Institute of Biophysics, Chinese Academy of Sciences (#2017-IRB-004). The image used in **Figure 3** is a photograph of one of the authors and The Consent to Publish Form was obtained.

### Experiment 1 (normal famous and unfamiliar face)

Fifteen subjects were presented with famous faces (popular film actors, 50% female) and objects (houses, scenery and small manmade objects) and were instructed to perform a category classification task (face or object) while their brain activity was recorded using MEG. Two subjects with excessive head motion (>5 mm) were excluded from further analysis. Each type of image includes 50 exemplars and all faces are own race faces. All images used were equated for contrast and mean luminance using the SHINE toolbox (*Willenbockel et al., 2010*). Each trial was initiated with a fixation with a jittered duration (800–1000 ms), then a grayscale visual image (face or object,  $8 \times 6^\circ$ ) was presented at the center of screen for 500 ms, followed by a response period. Subjects were asked to maintain fixation and report whether the image was a face or an object using button press as soon as possible. There were 120 trials for each condition. Nine of the thirteen subjects participated in an additional experiment in which unfamiliar faces were used.

### Experiment 2 (normal unfamiliar face and Mooney face)

Experiment two was conducted similar to Experiment 1, except that unfamiliar faces and two-tone Mooney faces were presented to subjects ( $n = 28$ ) in separate blocks (15 trials each) during which subjects performed a one-back task. Two subjects with excessive head motion (>5 mm) were excluded from further analysis.

### Experiment 3 (face-like images with spatially misarranged internal features)

Experiment three was conducted similar to Experiment 1, except that distorted face and object images were presented to subjects ( $n = 9$ ). Distorted face images were created by rearranging the eyes, mouth and nose into a nonface configuration (*Liu et al., 2002*).

### Experiment 4 (contextual cues defined the presence of faces without internal features)

Experiment four was conducted similar to Experiment 2. Three types of stimuli (**Figure 5A**) were created as described in previous study (*Cox et al., 2004*): (i) images of highly degraded faces (no internal facial features) with contextual body cues that imply the presence of faces, (ii) similar to images in (i) but with body cues arranged in an incorrect configuration and thus do not imply the presence of faces, (iii) images of objects. Fifteen subjects participated in this experiment and one of them was excluded from further data analysis due to excessive head motion (>5 mm).

### MEG data acquisition and analysis

MEG data were recorded continuously using a 275-channel CTF system. Three coils were attached on the head, one close to nasion, and the other two close to left and right preauricular points respectively. fMRI scanning was performed shortly after MEG data collection, and the locations of coils were marked with vitamin E caplets to align with MEG frames. MEG data analysis was performed using MATLAB (RRID: [SCR\\_001622](https://doi.org/10.1006/scr.001622)) and Fieldtrip toolbox (*Oostenveld et al., 2011*) (RRID: [SCR\\_004849](https://doi.org/10.1006/scr.004849)) for artifact detection and MNE-python (RRID: [SCR\\_005972](https://doi.org/10.1006/scr.005972)) for source analysis (*Gramfort et al., 2013; Gramfort et al., 2014*).

## Preprocessing

After acquisition, we first conducted time correction as there was time delay (measured with a photodiode) between the stimulus onset on the screen and the trigger signal in the recorded MEG data. Then the data were bandpass filtered with a frequency range of 2–80 Hz and epoched from 250 ms before to 550 ms after the stimulus onset. Bad channels and trials contaminated by artifacts including eye blinks, muscle activities and SQUID jumps were removed before further analysis.

## Source localization

Source localization can be generally divided into two steps, forward solution and inverse solution. Boundary-element model (BEM) which describes the geometry of the head and conductivities of the different tissues, coregistration information between MEG and MRI, and volume source space which defines the position of the source locations (10242 sources per hemisphere and the source spacing is 3.1 mm) were used to calculate forward solution. For inverse solution, we first estimated noise and data covariance matrix from –250 to 0 ms epochs and 100 to 350 ms epochs respectively. Afterwards, the Linearly Constrained Minimum Variance (LCMV) beamformer was calculated using covariance matrix and forward solution (Van Veen et al., 1997). The regularization for the whitened data covariance is 0.01. The source orientation which maximizes output source power is selected.

## Time course analysis

To explore the time course, virtual sensors were computed on the 30 Hz low-pass filtered data using the LCMV beamformer at the grid points within individual face-selective areas. The time course of each face-selective area was extracted from the grid point showing max value of MEG response. Subjects who did not show corresponding face-selective areas in fMRI localizer were excluded from time course extraction (See Table 1 for details). To identify time-points of significant differences, we performed non-parametric statistical tests with cluster-based multiple comparison correction (Maris and Oostenveld, 2007).

## Peak latency analysis

For each ROI of each subject, peak latency was defined as the timing of the largest peak within the first 250 ms of averaged response. To avoid the influence of bad source data with weak signal, time course without any time points showing response 5 SDs above the baseline (time average from –250 to 0 ms) was eliminated from peak analysis. The numbers of subjects used in peak latency analysis are summarized in Table 2. Two-tailed paired t tests (subjects with missing values were excluded) were used to compare the peak latencies between ROIs. While in Experiment 2, a more rigorous statistical approach, two sample paired permutation test (10000 permutations), was used to compare the peak latencies between pFFA and OFA (See results for details).

## Granger causality analysis

To study the regional information flow between ROIs, we employed Granger causality analysis (Granger, 1969) which is a statistical technique that based on the prediction of one time series on another. Time courses used in this analysis were extracted from each ROI without low-passed

**Table 1.** Number of subjects showing fMRI defined face-selective areas.

|       | Experiment 1 |                 | Experiment 2 | Experiment 3 | Experiment 4 |
|-------|--------------|-----------------|--------------|--------------|--------------|
|       | famous face  | Unfamiliar face |              |              |              |
| IOFA  | 13/13        | 9/9             | 25/26        | 9/9          | 13/14        |
| lpFFA | 13/13        | 9/9             | 26/26        | 9/9          | 14/14        |
| lpSTS | 13/13        | 9/9             | 18/26        | 9/9          | 11/14        |
| rOFA  | 13/13        | 9/9             | 26/26        | 9/9          | 14/14        |
| rpFFA | 13/13        | 9/9             | 26/26        | 9/9          | 14/14        |
| raFFA | 12/13        | 9/9             | 18/26        | 9/9          | 12/14        |
| rpSTS | 13/13        | 9/9             | 23/26        | 9/9          | 14/14        |



**Table 2.** Number of subjects used in peak latency analysis.

|       | Experiment 1 |                 | Experiment 2 |             | Experiment 3   | Experiment 4                 |
|-------|--------------|-----------------|--------------|-------------|----------------|------------------------------|
|       | famous face  | unfamiliar face | normal face  | Mooney face | distorted face | contextual cues defined face |
| IOFA  | 13/13        | 9/9             | 24/26        | 24/26       | 9/9            | 13/14                        |
| lpFFA | 12/13        | 9/9             | 25/26        | 25/26       | 9/9            | 12/14                        |
| lpSTS | 11/13        | 8/9             | -            | -           | -              | -                            |
| rOFA  | 13/13        | 9/9             | 24/26        | 26/26       | 9/9            | 13/14                        |
| rpFFA | 13/13        | 9/9             | 24/26        | 25/26       | 9/9            | 13/14                        |
| raFFA | 12/13        | 8/9             | 18/26        | 15/26       | 9/9            | 10/14                        |
| rpSTS | 12/13        | 7/9             | -            | -           | -              | -                            |

filtering. Causality analysis was performed using Multivariate Granger Causality (MVGC) toolbox ([Barnett and Seth, 2014](#)). Evoked response was removed from the data by linear regression before further analysis because the time series is assumed to be stationary in Granger causality analysis and this assumption is challenged in evoked brain responses ([Wang et al., 2008](#)). We conducted separate analysis over a series of overlapping 50 ms time windows (based on a previous study [Ashrafulla et al., 2013](#)) from 75 to 230 ms, which covers the period of face-induced activation in both OFA and FFA. There is a trade-off between stationary, temporal resolution (shorter is better) and accuracy of model fit (longer is better) when considering the size of time window. Moreover, smaller window is not considered because activity beyond Beta-band is not strong according to the power spectrum. First, the best model order was selected according to Bayesian information criteria (BIC). Then the corresponding vector auto regressive (VAR) model parameters were estimated for the selected model order and the autocovariance sequence for the VAR model was calculated. Then the bidirectional Granger causality values for each pair ROI were obtained by calculating pairwise-conditional time-domain MVGCs based on autocovariance sequence. Finally, to evaluate whether causality values are significantly greater than zero (assume null hypothesis causality value = 0), we performed significance test using F null distribution with FDR correction for multiple comparisons ([Benjamini and Hochberg, 1995](#)).

### fMRI data acquisition and analysis

Scanning was performed on a 3T Siemens Prisma scanner in the Beijing MRI Center for Brain Research. We acquired high-resolution T1-weighted anatomical volumes first, and then performed a run of functional face localizer ([Pitcher et al., 2011a](#)) with interleaved face and object blocks using a gradient echo-planar sequence (20-channel head coil, TR = 2 s, TE = 30 ms, resolution 2.0 × 2.0 × 2.0 mm, 31 slices, matrix = 96 × 96). fMRI data were analyzed using FreeSurfer (RRID: [SCR\\_001847](#)) and AFNI (RRID: [SCR\\_005927](#)). Face-selective areas were defined as regions that responded more strongly to faces than to objects.

### Acknowledgements

We thank Daniel Kersten for helpful comments on the manuscript and Ling Liu for her help in MEG data analysis. This work was supported by the Beijing Science and Technology Project (Z181100001518002, Z171100000117003), the Ministry of Science and Technology of China grants (2015CB351701) and Bureau of International Cooperation, Chinese Academy of Sciences (153311KYSB20160030).

### Additional information

#### Funding

| Funder                                 | Grant reference number | Author   |
|--|------------------------|----------|
| Beijing Science and Technology Project | Z181100001518002       | Sheng He |

Ministry of Science and Technology of the People's Republic of China 2015CB351701 Fan Wang

Bureau of International Cooperation, Chinese Academy of Sciences 153311KYSB20160030 Peng Zhang

Beijing Science and Technology Project Z171100000117003 Sheng He

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

### Author contributions

Xiaoxu Fan, Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Methodology, Writing; Fan Wang, Methodology, Data acquisition, Data analysis, Funding acquisition; Hanyu Shao, Methodology, Data acquisition; Peng Zhang, Methodology, Data analysis, Funding acquisition; Sheng He, Conceptualization, Investigation, Methodology, Writing, Funding acquisition, Project administration

### Author ORCIDs

Xiaoxu Fan  <https://orcid.org/0000-0002-8115-8621>

Sheng He  <https://orcid.org/0000-0001-5547-923X>

### Ethics

Human subjects: All subjects (age range 19-31) provided written informed consent and consent to publish before the experiments, and experimental protocols were approved by the Institutional Review Board of the Institute of Biophysics, Chinese Academy of Sciences (# 2017-IRB-004). The image used in Figure 3 is a photograph of one of the authors and The Consent to Publish Form was obtained.

### Decision letter and Author response

Decision letter <https://doi.org/10.7554/eLife.48764.sa1>

Author response <https://doi.org/10.7554/eLife.48764.sa2>

## Additional files

### Supplementary files

- Source code 1. Preprocessing.
- Source code 2. Source localization.
- Source code 3. Extract timecourse.
- Transparent reporting form

### Data availability

The source data files have been provided for Figures 2, 3, 4, 5 and Figure 2—figure supplement 1. MEG source activation data (processed based on original fMRI and MEG datasets ) have been deposited in Open Science Framework and can be accessed at <https://osf.io/vhefz/>.

The following dataset was generated:

| Author(s) | Year | Dataset title        | Dataset URL   | Database and Identifier       |
|-----------|------|----------------------|---|-------------------------------|
| Fan X     | 2020 | MEG face experiments | <a href="https://osf.io/vhefz/">https://osf.io/vhefz/</a> | Open Science Framework, vhefz |

## References

- Anaki D, Zion-Golumbic E, Bentin S. 2007. Electrophysiological neural mechanisms for detection, configural analysis and recognition of faces. *NeuroImage* **37**:1407–1416. DOI: <https://doi.org/10.1016/j.neuroimage.2007.05.054>, PMID: 17689102
- Andrews TJ, Davies-Thompson J, Kingstone A, Young AW. 2010. Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *Journal of Neuroscience* **30**:3544–3552. DOI: <https://doi.org/10.1523/JNEUROSCI.4863-09.2010>, PMID: 20203214
- Arcurio LR, Gold JM, James TW. 2012. The response of face-selective cortex with single face parts and part combinations. *Neuropsychologia* **50**:2454–2459. DOI: <https://doi.org/10.1016/j.neuropsychologia.2012.06.016>, PMID: 22750118
- Ashrafulla S, Haldar JP, Joshi AA, Leahy RM. 2013. Canonical Granger causality between regions of interest. *NeuroImage* **83**:189–199. DOI: <https://doi.org/10.1016/j.neuroimage.2013.06.056>, PMID: 23811410
- Baldauf D, Desimone R. 2014. Neural mechanisms of object-based attention. *Science* **344**:424–427. DOI: <https://doi.org/10.1126/science.1247003>, PMID: 24763592
- Barnett L, Seth AK. 2014. The MVGC multivariate granger causality toolbox: a new approach to Granger-causal inference. *Journal of Neuroscience Methods* **223**:50–68. DOI: <https://doi.org/10.1016/j.jneumeth.2013.10.018>, PMID: 24200508
- Baseler HA, Harris RJ, Young AW, Andrews TJ. 2014. Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. *Cerebral Cortex* **24**:737–744. DOI: <https://doi.org/10.1093/cercor/bhs360>, PMID: 23172771
- Benitez-Quiroz CF, Srinivasan R, Martinez AM. 2018. Facial color is an efficient mechanism to visually transmit emotion. *PNAS*. **115**:3581–3586. DOI: <https://doi.org/10.1073/pnas.1716084115>, PMID: 29555780
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* **57**:289–300. DOI: <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bentin S, Allison T, Puce A, Perez E, McCarthy G. 1996. Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience* **8**:551–565. DOI: <https://doi.org/10.1162/jocn.1996.8.6.551>, PMID: 20740065
- Bernstein M, Yovel G. 2015. Two neural pathways of face processing: a critical evaluation of current models. *Neuroscience & Biobehavioral Reviews* **55**:536–546. DOI: <https://doi.org/10.1016/j.neubiorev.2015.06.010>, PMID: 26067903
- Bi T, Chen J, Zhou T, He Y, Fang F. 2014. Function and structure of human left fusiform cortex are closely associated with perceptual learning of faces. *Current Biology* **24**:222–227. DOI: <https://doi.org/10.1016/j.cub.2013.12.028>, PMID: 24412207
- Braje W, Kersten D, Tarr M, Troje N. 1998. Illumination effects in face recognition. *Psychobiology* **26**:371–380. DOI: <https://doi.org/10.3758/BF03330623>
- Cavanagh P. 1991. What's up in top-down processing. In: Gorea A (Ed). *Representations of Vision*. Cambridge University Press. p. 295–304.
- Chen Z, Whitney D. 2019. Tracking the affective state of unseen persons. *PNAS* **116**:7559–7564. DOI: <https://doi.org/10.1073/pnas.1812250116>, PMID: 30814221
- Cichy RM, Pantazis D, Oliva A. 2014. Resolving human object recognition in space and time. *Nature Neuroscience* **17**:455–462. DOI: <https://doi.org/10.1038/nn.3635>, PMID: 24464044
- Cichy RM, Pantazis D, Oliva A. 2016. Similarity-Based fusion of MEG and fMRI reveals Spatio-Temporal dynamics in human cortex during visual object recognition. *Cerebral Cortex* **26**:3563–3579. DOI: <https://doi.org/10.1093/cercor/bhw135>, PMID: 27235099
- Cox D, Meyers E, Sinha P. 2004. Contextually evoked object-specific responses in human visual cortex. *Science* **304**:115–117. DOI: <https://doi.org/10.1126/science.1093110>, PMID: 15001712
- de Heering A, Rossion B. 2015. Rapid categorization of natural face images in the infant right hemisphere. *eLife* **4**:e06564. DOI: <https://doi.org/10.7554/eLife.06564>, PMID: 26032564
- Deffke I, Sander T, Heidenreich J, Sommer W, Curio G, Trahms L, Lueschow A. 2007. MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *NeuroImage* **35**:1495–1501. DOI: <https://doi.org/10.1016/j.neuroimage.2007.01.034>, PMID: 17363282
- Dijkstra N, Mostert P, Lange FP, Bosch S, van Gerven MA. 2018. Differential temporal dynamics during visual imagery and perception. *eLife* **7**:e33904. DOI: <https://doi.org/10.7554/eLife.33904>, PMID: 29807570
- Dobs K, Isik L, Pantazis D, Kanwisher N. 2019. How face perception unfolds over time. *Nature Communications* **10**:1258. DOI: <https://doi.org/10.1038/s41467-019-09239-1>, PMID: 30890707
- Duchaine B, Yovel G. 2015. A revised neural framework for face processing. *Annual Review of Vision Science* **1**:393–416. DOI: <https://doi.org/10.1146/annurev-vision-082114-035518>, PMID: 28532371
- Engell AD, Haxby JV. 2007. Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia* **45**:3234–3241. DOI: <https://doi.org/10.1016/j.neuropsychologia.2007.06.022>, PMID: 17707444
- Fairhall SL, Ishai A. 2007. Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex* **17**:2400–2406. DOI: <https://doi.org/10.1093/cercor/bhl148>, PMID: 17190969

- Feng L, Liu J, Wang Z, Li J, Li L, Ge L, Tian J, Lee K. 2011. The other face of the other-race effect: an fMRI investigation of the other-race face categorization advantage. *Neuropsychologia* **49**:3739–3749. DOI: <https://doi.org/10.1016/j.neuropsychologia.2011.09.031>, PMID: 21971308
- Freiwald WA, Tsao DY. 2010. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* **330**:845–851. DOI: <https://doi.org/10.1126/science.1194908>, PMID: 21051642
- Friston K. 2005. A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**:815–836. DOI: <https://doi.org/10.1098/rstb.2005.1622>
- Friston K. 2010. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience* **11**:127–138. DOI: <https://doi.org/10.1038/nrn2787>, PMID: 20068583
- Gao Z, Goldstein A, Harpaz Y, Hansel M, Zion-Golombic E, Bentin S. 2013. A magnetoencephalographic study of face processing: m170, gamma-band oscillations and source localization. *Human Brain Mapping* **34**:1783–1795. DOI: <https://doi.org/10.1002/hbm.22028>, PMID: 22422432
- Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. 2000. The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience* **12**:495–504. DOI: <https://doi.org/10.1162/089892900562165>, PMID: 10931774
- Gentile F, Ales J, Rossion B. 2017. Being BOLD: the neural dynamics of face perception. *Human Brain Mapping* **38**:120–139. DOI: <https://doi.org/10.1002/hbm.23348>, PMID: 27585292
- Gerardin P, Kourtzi Z, Mamassian P. 2010. Prior knowledge of illumination for 3D perception in the human brain. *PNAS*. **107**:16309–16314. DOI: <https://doi.org/10.1073/pnas.1006285107>, PMID: 20805488
- Ghuman AS, Brunet NM, Li Y, Konecky RO, Pyles JA, Walls SA, Destefino V, Wang W, Richardson RM. 2014. Dynamic encoding of face information in the human fusiform gyrus. *Nature Communications* **5**:5672. DOI: <https://doi.org/10.1038/ncomms6672>, PMID: 25482825
- Gobbini MI, Haxby JV. 2006. Neural response to the visual familiarity of faces. *Brain Research Bulletin* **71**:76–82. DOI: <https://doi.org/10.1016/j.brainresbull.2006.08.003>, PMID: 17113931
- Gobbini MI, Haxby JV. 2007. Neural systems for recognition of familiar faces. *Neuropsychologia* **45**:32–41. DOI: <https://doi.org/10.1016/j.neuropsychologia.2006.04.015>, PMID: 16797608
- Goold JE, Meng M. 2017. Categorical learning revealed in activity pattern of left fusiform cortex. *Human Brain Mapping* **38**:3648–3658. DOI: <https://doi.org/10.1002/hbm.23620>, PMID: 28432767
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M. 2013. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience* **7**:267. DOI: <https://doi.org/10.3389/fnins.2013.00267>, PMID: 24431986
- Gramfort A, Luessi M, Larson E, Engemann DA, Strohmeier D, Brodbeck C, Parkkonen L, Hämäläinen M. 2014. MNE software for processing MEG and EEG data. *NeuroImage* **86**:446–460. DOI: <https://doi.org/10.1016/j.neuroimage.2013.10.027>, PMID: 24161808
- Granger CWJ. 1969. Investigating causal relations by econometric models and Cross-spectral methods. *Econometrica* **37**:424. DOI: <https://doi.org/10.2307/1912791>
- Grill-Spector K, Knouf N, Kanwisher N. 2004. The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience* **7**:555–562. DOI: <https://doi.org/10.1038/nn1224>, PMID: 15077112
- Grossman ED, Battelli L, Pascual-Leone A. 2005. Repetitive TMS over posterior STS disrupts perception of biological motion. *Vision Research* **45**:2847–2853. DOI: <https://doi.org/10.1016/j.visres.2005.05.027>
- Grützner C, Uhlhaas PJ, Genc E, Kohler A, Singer W, Wibral M. 2010. Neuroelectromagnetic correlates of perceptual closure processes. *Journal of Neuroscience* **30**:8342–8352. DOI: <https://doi.org/10.1523/JNEUROSCI.5434-09.2010>, PMID: 20554885
- Haxby JV, Ungerleider LG, Clark VP, Schouten JL, Hoffman EA, Martin A. 1999. The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* **22**:189–199. DOI: [https://doi.org/10.1016/S0896-6273\(00\)80690-X](https://doi.org/10.1016/S0896-6273(00)80690-X)
- Haxby JV, Hoffman EA, Gobbini MI. 2000. The distributed human neural system for face perception. *Trends in Cognitive Sciences* **4**:223–233. DOI: [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0), PMID: 10827445
- Hegde J, Thompson S, Kersten D. 2007. Identifying faces in two-tone ('Mooney') images: A psychophysical and fMRI study. *Journal of Vision* **7**:624. DOI: <https://doi.org/10.1167/7.9.624>
- Hoffman EA, Haxby JV. 2000. Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience* **3**:80–84. DOI: <https://doi.org/10.1038/71152>, PMID: 10607399
- Humphrey GK, Goodale MA, Bowen CV, Gati JS, Vilis T, Rutt BK, Menon RS. 1997. Differences in perceived shape from shading correlate with activity in early visual Areas. *Current Biology* **7**:144–147. DOI: [https://doi.org/10.1016/S0960-9822\(06\)00058-3](https://doi.org/10.1016/S0960-9822(06)00058-3), PMID: 9016702
- Issa EB, Cadieu CF, DiCarlo JJ. 2018. Neural dynamics at successive stages of the ventral visual stream are consistent with hierarchical error signals. *eLife* **7**:e42870. DOI: <https://doi.org/10.7554/eLife.42870>, PMID: 30484773
- Issa EB, DiCarlo JJ. 2012. Precedence of the eye region in neural processing of faces. *Journal of Neuroscience* **32**:16666–16682. DOI: <https://doi.org/10.1523/JNEUROSCI.2391-12.2012>, PMID: 23175821
- Itier RJ, Herdman AT, George N, Cheyne D, Taylor MJ. 2006. Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Research* **1115**:108–120. DOI: <https://doi.org/10.1016/j.brainres.2006.07.072>, PMID: 16930564

- Kadipasaoglu CM**, Conner CR, Baboyan VG, Rollo M, Pieters TA, Tandon N. 2017. Network dynamics of human face perception. *PLOS ONE* **12**:e0188834. DOI: <https://doi.org/10.1371/journal.pone.0188834>, PMID: 29190811
- Kanwisher N**, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience* **17**:4302–4311. DOI: <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>, PMID: 9151747
- Keller CJ**, Davidesco I, Megevand P, Lado FA, Malach R, Mehta AD. 2017. Tuning face perception with electrical stimulation of the fusiform gyrus. *Human Brain Mapping* **38**:2830–2842. DOI: <https://doi.org/10.1002/hbm.23543>, PMID: 28345189
- Kume Y**, Maekawa T, Urakawa T, Hironaga N, Ogata K, Shigyo M, Tobimatsu S. 2016. Neuromagnetic evidence that the right fusiform face area is essential for human face awareness: an intermittent binocular rivalry study. *Neuroscience Research* **109**:54–62. DOI: <https://doi.org/10.1016/j.neures.2016.02.004>, PMID: 26907522
- Landi SM**, Freiwald WA. 2017. Two Areas for familiar face recognition in the primate brain. *Science* **357**:591–595. DOI: <https://doi.org/10.1126/science.aan1139>, PMID: 28798130
- Lee TS**, Yang CF, Romero RD, Mumford D. 2002. Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nature Neuroscience* **5**:589–597. DOI: <https://doi.org/10.1038/nn0602-860>, PMID: 12021764
- Liu J**, Harris A, Kanwisher N. 2002. Stages of processing in face perception: an MEG study. *Nature Neuroscience* **5**:910–916. DOI: <https://doi.org/10.1038/nn909>, PMID: 12195430
- Liu J**, Harris A, Kanwisher N. 2010. Perception of face parts and face configurations: an fMRI study. *Journal of Cognitive Neuroscience* **22**:203–211. DOI: <https://doi.org/10.1162/jocn.2009.21203>, PMID: 19302006
- Maris E**, Oostenveld R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods* **164**:177–190. DOI: <https://doi.org/10.1016/j.jneumeth.2007.03.024>, PMID: 17517438
- Martinez AM**. 2017. Visual perception of facial expressions of emotion. *Current Opinion in Psychology* **17**:27–33. DOI: <https://doi.org/10.1016/j.copsyc.2017.06.009>, PMID: 28950969
- Martinez AM**. 2019. Context may reveal how you feel PNAS. **116**:7169–7171. DOI: <https://doi.org/10.1073/pnas.1902661116>, PMID: 30898883
- McGugin RW**, Ryan KF, Tamber-Rosenau BJ, Gauthier I. 2018. The role of experience in the Face-Selective response in right FFA. *Cerebral Cortex* **28**:2071–2084. DOI: <https://doi.org/10.1093/cercor/bhx113>, PMID: 28472436
- McKeeff TJ**, Tong F. 2007. The timing of perceptual decisions for ambiguous face stimuli in the human ventral visual cortex. *Cerebral Cortex* **17**:669–678. DOI: <https://doi.org/10.1093/cercor/bhk015>, PMID: 16648454
- Meng M**, Cherian T, Singal G, Sinha P. 2012. Lateralization of face processing in the human brain. *Proceedings of the Royal Society B: Biological Sciences* **279**:2052–2061. DOI: <https://doi.org/10.1098/rspb.2011.1784>
- Moeller S**, Crapse T, Chang L, Tsao DY. 2017. The effect of face patch microstimulation on perception of faces and objects. *Nature Neuroscience* **20**:743–752. DOI: <https://doi.org/10.1038/nn.4527>, PMID: 28288127
- Mooney CM**. 1957. Age in the development of closure ability in children. *Canadian Journal of Psychology/Revue Canadienne De Psychologie* **11**:219–226. DOI: <https://doi.org/10.1037/h0083717>
- Moore C**, Cavanagh P. 1998. Recovery of 3D volume from 2-tone images of novel objects. *Cognition* **67**:45–71. DOI: [https://doi.org/10.1016/S0010-0277\(98\)00014-6](https://doi.org/10.1016/S0010-0277(98)00014-6), PMID: 9735536
- Mumford D**. 1992. On the computational architecture of the neocortex. II. the role of cortico-cortical loops. *Biological Cybernetics* **66**:241–251. DOI: <https://doi.org/10.1007/bf00198477>, PMID: 1540675
- Murray SO**, Schrater P, Kersten D. 2004. Perceptual grouping and the interactions between visual cortical Areas. *Neural Networks* **17**:695–705. DOI: <https://doi.org/10.1016/j.neunet.2004.03.010>, PMID: 15288893
- Nguyen VT**, Cunnington R. 2014. The superior temporal sulcus and the N170 during face processing: single trial analysis of concurrent EEG-fMRI. *NeuroImage* **86**:492–502. DOI: <https://doi.org/10.1016/j.neuroimage.2013.10.047>, PMID: 24185024
- O'Toole AJ**, Roark DA, Abdi H. 2002. Recognizing moving faces: a psychological and neural synthesis. *Trends in Cognitive Sciences* **6**:261–266. DOI: [https://doi.org/10.1016/S1364-6613\(02\)01908-3](https://doi.org/10.1016/S1364-6613(02)01908-3), PMID: 12039608
- Ohayon S**, Freiwald WA, Tsao DY. 2012. What makes a cell face selective? the importance of contrast. *Neuron* **74**:567–581. DOI: <https://doi.org/10.1016/j.neuron.2012.03.024>, PMID: 22578507
- Oostenveld R**, Fries P, Maris E, Schoffelen JM. 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience* **2011**:1–9. DOI: <https://doi.org/10.1155/2011/156869>, PMID: 21253357
- Perry G**, Singh KD. 2014. Localizing evoked and induced responses to faces using magnetoencephalography. *European Journal of Neuroscience* **39**:1517–1527. DOI: <https://doi.org/10.1111/ejn.12520>, PMID: 24617643
- Pitcher D**, Walsh V, Yovel G, Duchaine B. 2007. TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology* **17**:1568–1573. DOI: <https://doi.org/10.1016/j.cub.2007.07.063>, PMID: 17764942
- Pitcher D**, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N. 2011a. Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage* **56**:2356–2363. DOI: <https://doi.org/10.1016/j.neuroimage.2011.03.067>, PMID: 21473921
- Pitcher D**, Walsh V, Duchaine B. 2011b. The role of the occipital face area in the cortical face perception network. *Experimental Brain Research* **209**:481–493. DOI: <https://doi.org/10.1007/s00221-011-2579-1>, PMID: 21318346
- Pitcher D**, Goldhaber T, Duchaine B, Walsh V, Kanwisher N. 2012. Two critical and functionally distinct stages of face and body perception. *Journal of Neuroscience* **32**:15877–15885. DOI: <https://doi.org/10.1523/JNEUROSCI.2624-12.2012>, PMID: 23136426



- Pitcher D**, Duchaine B, Walsh V. 2014. Combined TMS and fMRI reveal dissociable cortical pathways for dynamic and static face perception. *Current Biology* **24**:2066–2070. DOI: <https://doi.org/10.1016/j.cub.2014.07.060>, PMID: 25131678
- Pourtois G**, Schwartz S, Seghier ML, Lazeyras F, Vuilleumier P. 2005. View-independent coding of face identity in frontal and temporal cortices is modulated by familiarity: an event-related fMRI study. *NeuroImage* **24**:1214–1224. DOI: <https://doi.org/10.1016/j.neuroimage.2004.10.038>, PMID: 15670699
- Puce A**, Allison T, Bentin S, Gore JC, McCarthy G. 1998. Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience* **18**:2188–2199. DOI: <https://doi.org/10.1523/JNEUROSCI.18-06-02188.1998>, PMID: 9482803
- Rao R**, Ballard DH. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature* **2**:79–87. DOI: <https://doi.org/10.1038/4580>
- Rossion B**, Caldara R, Seghier M, Schuller AM, 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. DOI: <https://doi.org/10.1093/brain/awg241>, PMID: 12876150
- Rossion B**, Dricot L, Goebel R, Busigny T. 2011. Holistic face categorization in higher order visual Areas of the normal and prosopagnosic brain: toward a non-hierarchical view of face perception. *Frontiers in Human Neuroscience* **4**:225. DOI: <https://doi.org/10.3389/fnhum.2010.00225>, PMID: 21267432
- Rossion B**, Prieto EA, Boremanse A, Kuefner D, Van Belle G. 2012. A steady-state visual evoked potential approach to individual face perception: effect of Inversion, contrast-reversal and temporal dynamics. *NeuroImage* **63**:1585–1600. DOI: <https://doi.org/10.1016/j.neuroimage.2012.08.033>, PMID: 22917988
- Rossion B**, Torfs K, Jacques C, Liu-Shuang J. 2015. Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. *Journal of Vision* **15**:18. DOI: <https://doi.org/10.1167/15.1.18>, PMID: 25597037
- Rossion B**, Jacques C. 2011. The N170: understanding the time-course of face perception in the human brain. In: Luck S. J, Kappenman E. S (Eds). *The Oxford Handbook of Event-Related Potential Components*. Oxford University Press. p. 115–141. DOI: <https://doi.org/10.1093/oxfordhb/9780195374148.013.0064>
- Rotshtein P**, Henson RN, Treves A, Driver J, Dolan RJ. 2005. Morphing marilyn into Maggie dissociates physical and identity face representations in the brain. *Nature Neuroscience* **8**:107–113. DOI: <https://doi.org/10.1038/nn1370>, PMID: 15592463
- Russell R**, Biederman I, Nederhouser M, Sinha P. 2007. The utility of surface reflectance for the recognition of upright and inverted faces. *Vision Research* **47**:157–165. DOI: <https://doi.org/10.1016/j.visres.2006.11.002>, PMID: 17174375
- Sadeh B**, Podlipsky I, Zhdanov A, Yovel G. 2010. Event-related potential and functional MRI measures of face-selectivity are highly correlated: a simultaneous ERP-fMRI investigation. *Human Brain Mapping* **31**:1490–1501. DOI: <https://doi.org/10.1002/hbm.20952>, PMID: 20127870
- Schiltz C**. 2010. Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of Vision* **10**:1–16. DOI: <https://doi.org/10.1167/10.2.25>
- Schwartz L**, Yovel G. 2016. The roles of perceptual and conceptual information in face recognition. *Journal of Experimental Psychology: General* **145**:1493–1511. DOI: <https://doi.org/10.1037/xge0000220>
- Sekihara K**, Nagarajan SS. 2008. *Adaptive Spatial Filters for Electromagnetic Brain Imaging*. Berlin, Heidelberg: Springer. DOI: <https://doi.org/10.1007/978-3-540-79370-0>
- Sinha P**, Balas B, Ostrovsky Y, Russell R. 2006. Face recognition by humans: nineteen results all computer vision researchers should know about. *Proceedings of the IEEE* **94**:1948–1962. DOI: <https://doi.org/10.1109/JPROC.2006.884093>
- Srinivasan R**, Golomb JD, Martinez AM. 2016. A neural basis of facial action recognition in humans. *Journal of Neuroscience* **36**:4434–4442. DOI: <https://doi.org/10.1523/JNEUROSCI.1704-15.2016>, PMID: 27098688
- Steeves JK**, Culham JC, Duchaine BC, Pratesi CC, Valyear KF, Schindler I, Humphrey GK, Milner AD, Goodale MA. 2006. The fusiform face area is not sufficient for face recognition: evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia* **44**:594–609. DOI: <https://doi.org/10.1016/j.neuropsychologia.2005.06.013>, PMID: 16125741
- Sugiura M**, Mano Y, Sasaki A, Sadato N. 2011. Beyond the memory mechanism: person-selective and nonselective processes in recognition of personally familiar faces. *Journal of Cognitive Neuroscience* **23**:699–715. DOI: <https://doi.org/10.1162/jocn.2010.21469>, PMID: 20350171
- Tsao DY**, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB. 2003. Faces and objects in macaque cerebral cortex. *Nature Neuroscience* **6**:989–995. DOI: <https://doi.org/10.1038/nn1111>, PMID: 12925854
- Tsao DY**, Freiwald WA, Tootell RB, Livingstone MS. 2006. A cortical region consisting entirely of face-selective cells. *Science* **311**:670–674. DOI: <https://doi.org/10.1126/science.1119983>, PMID: 16456083
- Tsao DY**, Moeller S, Freiwald WA. 2008. Comparing face patch systems in macaques and humans *PNAS*. **105**:19514–19519. DOI: <https://doi.org/10.1073/pnas.0809662105>, PMID: 19033466
- Van Veen BD**, van Drongelen W, Yuchtman M, Suzuki A. 1997. Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering* **44**:867–880. DOI: <https://doi.org/10.1109/10.623056>, PMID: 9282479
- Wang X**, Chen Y, Ding M. 2008. Estimating Granger causality after stimulus onset: a cautionary note. *NeuroImage* **41**:767–776. DOI: <https://doi.org/10.1016/j.neuroimage.2008.03.025>, PMID: 18455441
- Weibert K**, Flack TR, Young AW, Andrews TJ. 2018. Patterns of neural response in face regions are predicted by low-level image properties. *Cortex* **103**:199–210. DOI: <https://doi.org/10.1016/j.cortex.2018.03.009>, PMID: 29655043



- Weiner KS**, Barnett MA, Lorenz S, Caspers J, Stigliani A, Amunts K, Zilles K, Fischl B, Grill-Spector K. 2017. The cytoarchitecture of Domain-specific regions in human High-level visual cortex. *Cerebral Cortex* **27**:146–161. DOI: <https://doi.org/10.1093/cercor/bhw361>, PMID: 27909003
- Willenbockel V**, Sadr J, Fiset D, Horne GO, Gosselin F, Tanaka JW. 2010. Controlling low-level image properties: the SHINE toolbox. *Behavior Research Methods* **42**:671–684. DOI: <https://doi.org/10.3758/BRM.42.3.671>, PMID: 20805589
- Yip AW**, Sinha P. 2002. Contribution of color to face recognition. *Perception* **31**:995–1003. DOI: <https://doi.org/10.1068/p3376>, PMID: 12269592