

## Face categorization in visual scenes may start in a higher order area of the right fusiform gyrus: evidence from dynamic visual stimulation in neuroimaging

Fang Jiang,<sup>1</sup> Laurence Dricot,<sup>1</sup> Jochen Weber,<sup>2</sup> Giulia Righi,<sup>3</sup> Michael J. Tarr,<sup>4</sup> Rainer Goebel,<sup>5</sup> and Bruno Rossion<sup>1</sup>

<sup>1</sup>*Institute of Psychology and Institute of Neuroscience, University of Louvain, Belgium;* <sup>2</sup>*Department of Psychology, Columbia University, New York, New York;* <sup>3</sup>*Division of Developmental Medicine, Children's Hospital Boston, Boston, Massachusetts;*

<sup>4</sup>*Center for the Neural Basis of Cognition and Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania;* and <sup>5</sup>*Maastricht Brain Imaging Center, Maastricht University, Maastricht, The Netherlands*

Submitted 30 July 2010; accepted in final form 5 July 2011

**Jiang F, Dricot L, Weber J, Righi G, Tarr MJ, Goebel R, Rossion B.** Face categorization in visual scenes may start in a higher order area of the right fusiform gyrus: evidence from dynamic visual stimulation in neuroimaging. *J Neurophysiol* 106: 2720–2736, 2011. First published July 6, 2011; doi:10.1152/jn.00672.2010.—How a visual stimulus is initially categorized as a face by the cortical face-processing network remains largely unclear. In this study we used functional MRI to study the dynamics of face detection in visual scenes by using a paradigm in which scenes containing faces or cars are revealed progressively as they emerge from visual noise. Participants were asked to respond as soon as they detected a face or car during the noise sequence. Among the face-sensitive regions identified based on a standard localizer, a high-level face-sensitive area, the right fusiform face area (FFA), showed the earliest difference between face and car activation. Critically, differential activation in FFA was observed before differential activation in the more posteriorly located occipital face area (OFA). A whole brain analysis confirmed these findings, with a face-sensitive cluster in the right fusiform gyrus being the only cluster showing face preference before successful behavioral detection. Overall, these findings indicate that following generic low-level visual analysis, a face stimulus presented in a gradually revealed visual scene is first detected in the right middle fusiform gyrus, only after which further processing spreads to a network of cortical and subcortical face-sensitive areas (including the posteriorly located OFA). These results provide further evidence for a nonhierarchical organization of the cortical face-processing network.

face perception; functional magnetic resonance imaging

THE HUMAN BRAIN CAN DETECT a complex pattern such as a face in a visual scene in a fraction of a second (e.g., Crouzet et al. 2010; Fei-Fei et al. 2007; Lewis and Edmonds 2003; Rousselet et al. 2003). However, the specific neural mechanisms supporting the initial categorization of a visual stimulus as a face remain largely unclear. To effectively categorize a visual stimulus as a face, it must be segmented from the background scene and matched to an internal representation of a face rather than to competing non-face object shapes. In a complex visual scene, shapes may be extracted based on the pattern of similarity of textures inside coherent objects and the discontinuity

of textures across borders, a process that is thought to be initiated in early visual areas such as V2 (e.g., Appelbaum et al. 2006; Leventhal et al. 1998; Zhan and Baker 2006). In addition, higher areas of the visual system such as the lateral occipital complex (LOC; Malach et al. 1995), which show general shape selectivity that is mostly independent of local segmentation cues and spatial position (e.g. Grill-Spector and Malach 2001), are thought to play a critical role in object categorization in visual scenes (Appelbaum et al. 2006; Peelen et al. 2009).

With respect to faces in particular, neuroimaging studies have identified a set of high-level visual areas that respond significantly more to pictures of segmented faces than to other object shapes (Fox et al. 2009; Haxby et al. 2000; Ishai 2008; Sergent et al. 1992; Tsao et al. 2008; Weiner and Grill-Spector 2010), thus potentially playing an important role in the fast and accurate categorization of a visual stimulus as a face. These areas, as identified in functional magnetic resonance imaging (fMRI) studies, are of few square millimeters in size and are located outside of anatomically well-defined retinotopic visual areas (Halgren et al. 1999; Weiner and Grill-Spector 2010), in the middle fusiform gyrus (fusiform face area, FFA; e.g., Kanwisher et al. 1997; Puce et al. 1995), and more posteriorly in the lateral part of the inferior occipital lobe (occipital face area, OFA; e.g., Gauthier et al. 2000) as well as in the posterior part of the superior temporal sulcus (pSTS; e.g., Puce et al. 1998). These three areas are believed to form a core system of an extensive network of brain areas that are preferentially or exclusively sensitive to faces (Haxby et al. 2000), including areas in the ventral temporal cortex (anterior fusiform gyrus and temporal pole), the amygdala, and the dorsolateral prefrontal cortex (e.g., Haxby et al. 2000; Gobbini and Haxby 2007; Ishai 2008; Rajimehr et al. 2009; Sergent et al. 1992; Tsao et al. 2008; Weiner and Grill-Spector 2010). These areas are larger in size and generally show stronger and more consistent face-preferential responses in the right than in the left hemisphere (e.g., Fox et al. 2009; Kanwisher et al. 1997; Sergent et al. 1992), in agreement with the well-known dominant role of the right hemisphere in face perception (e.g., Hécaen and Anguelergues 1962; Parkin and Williamson 1987; Sergent 1988). A similar set of areas devoted to face processing has also been identified in the nonhuman primate brain in fMRI

Address for reprint requests and other correspondence: B. Rossion, Institute of Research in Psychology, Université de Louvain, 10, Place Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium (e-mail: bruno.rossion@uclouvain.be).

(Pinsk et al. 2009; Tsao et al. 2006, 2008), with recent investigations showing that these areas form a tightly and specifically interconnected network (Moeller et al. 2008).

An important question to clarify is how this network of face-preferential areas is functionally organized, that is, what is the dynamic flow of information between these areas, in our case particularly focusing on when detection of a face stimulus in a visual scene occurs. Most neurofunctional models of face perception, derived from neuroimaging studies and consistent with more generic, feedforward hierarchical object processing models, postulate that face-related processes are initiated in the most posteriorly located face-sensitive area of the network, namely, the inferior occipital cortex (OFA). Specific information about faces would then be forwarded to the more anteriorly located middle fusiform gyrus (FFA) and pSTS, and then to the anterior temporal and prefrontal cortices (e.g., Fairhall and Ishai 2007; Haxby et al. 2000; Ishai 2008). Thus, although some of these models also incorporate the possibility of feedback connections between face-sensitive areas (Haxby et al. 2000; Moeller et al. 2008), the initial processing of the stimulus as a face is thought to occur in a feedforward and hierarchical manner.

The hierarchical view of face processing is inspired by both the hierarchical organization of the visual system (Hubel and Wiesel 1962; Felleman and Van Essen 1991; Van Essen and Maunsell 1983) and influential computational models of object and scene recognition (Biederman 1987; Marr 1982; Riesenhuber and Poggio 1999; Ullman 2007). According to the hierarchical view of the visual system, feedforward connections carry the information from low-order areas to higher order areas (e.g., V1 to IT) in sequential processing steps. Feedback connections may transfer information in the reverse direction. From a computational point of view, the visual stimulus would be initially decomposed into small parts/features (or fragments; Ullman 2007) that would be subsequently combined into more complex object representations (Biederman 1987; Marr 1982). A similar “local-to-global” hierarchical view has generally been endorsed by computational and theoretical accounts of face perception (Burton 1994; Jiang et al. 2006).

When considering the functional neuroanatomy of face processing, this hierarchical view suggests that populations of neurons located in the inferior occipital cortex (OFA) would initially code for specific features of the face (e.g., mouth, nose, etc.). These features would then be integrated in a global face representation at later stages of the face processing network, such as the FFA (Fairhall and Ishai 2007; Haxby et al. 2000; Ishai 2008; Pitcher et al. 2007; see also Lerner et al. 2001) and then the anterior inferotemporal cortex (e.g., Kriegeskorte et al. 2007; Nestor et al. 2008; Sergent et al. 1992).

However, there are at least two sets of data that are incompatible with a simple hierarchical view of face processing, or at least with the idea that the earliest stage of face sensitivity arises in the most posteriorly located face-sensitive area, the inferior occipital cortex (OFA). First, in the normal human brain, sensitivity to faces can be observed in the FFA and pSTS in the absence of such face sensitivity in the inferior occipital cortex. For instance, in a two-tone (degraded black and white) “Mooney” face image, the local features often become too ambiguous to be recognized individually and must be disam-

biguated based on their organization within a global configuration (Mooney 1957; Moore and Cavanagh 1998). Interestingly, such Mooney faces can activate the right FFA without any evidence of face sensitivity in the posterior visual areas (Dolan et al. 1997), including a prelocalized right OFA (Rossion et al. 2011). Second, face sensitivity can be observed in high-level visual areas of the right middle fusiform gyrus such as the right FFA despite structural damage to the neural region containing the posteriorly located right OFA (Rossion et al. 2003; see also Steeves et al. 2006).

Together, these observations suggest that, in the intact brain, a preferential activation to faces in higher level visual areas such as the FFA may possibly arise independently of putative face-sensitive inputs from the inferior occipital cortex (OFA), perhaps through direct connections from early (non-face sensitive) visual cortices. On the basis of these observations, we hypothesized that the initial categorization of a visual stimulus as a face, rather than being carried out in the inferior occipital cortex according to a hierarchical scheme, may be initiated instead in higher visual areas of the face-processing network, especially in the right middle fusiform gyrus (Rossion et al. 2003; Rossion 2008). According to this view, an initial face categorization in a higher visual area of the right hemisphere cortical face network might be based on a global and coarse face representation, which would then be refined through a cortical reentrant loop to lower areas such as the OFA (see Mumford 1992 for a proposed reverse hierarchy in the visual system; also Hochstein and Ahissar 2002).

In the present fMRI study, we aimed to further investigate this nonhierarchical view of face processing. That is, we tested the hypothesis that during face categorization in visual scenes, preferential responses to faces in a higher order area of the ventral visual stream, namely, the right middle fusiform gyrus (right FFA), precede preferential responses to faces in the lower visual area, namely, the ipsilateral inferior occipital cortex (right OFA).

By using fMRI, we were cognizant that if we used a simple speeded detection task with briefly presented faces, we would not be able to directly compare the onset of face-related activation in the FFA, OFA, and other regions of the cortical face network. Indeed, the low temporal resolution of fMRI (e.g., Menon and Kim 1999) would not allow separation of the time of activation in two visual areas such as the FFA and OFA, separated by about 2 cm of cortex, and whose earliest face-related responses might be isolated by a few tens of milliseconds at most during fast face categorization.

Thus, to maximize chances to observe any difference in the time onset and duration of face sensitivity between high-level visual areas, we slowed down the perception of faces in visual scenes while maintaining the low-level visual stimulation constant for a sustained period of time. To accomplish this, we used a paradigm of gradually revealing information in successive continuous steps, that is, a dynamic sequence. Such dynamic visual stimulation paradigms have been used in previous fMRI studies to reveal aspects of visual priming (James et al. 2000), top-down facilitation (Eger et al. 2007), cortical areas contributing to different stages of recognition (Carlson et al. 2006), and perceptual hysteresis (Kleinschmidt et al. 2002). It has been shown that with these slow dynamic stimulation paradigms, the time course of activation and sensitivity to stimulus manipulation as observed in fMRI blood oxygen

level-dependent (BOLD) responses may reveal timing differences between visual areas (e.g., Carlson et al. 2006). In this study, as in these previous studies, the meaningful picture was not present at the onset of stimulation and was instead gradually revealed throughout a long dynamic sequence. However, in contrast to these previous studies, we controlled the low-level image properties such as luminance and frequency spectra by progressively denoising only the phase spectrum of the visual stimulus (Sadr and Sinha 2004). To our knowledge, this procedure has only been used so far on segmented face and object stimuli, not visual scenes, in a handful of fMRI studies (Esterman and Yantis 2010; Reinders et al. 2005, 2006; Philastides and Sajda 2007) that had different objectives from the present study.

## MATERIALS AND METHODS

**Participants.** Twelve participants (8 females and 4 males; mean age 24 yr, 9 right-handed) were included in the study. All had normal or corrected vision. Written informed consent was obtained from all participants before the experiment, following procedures approved by University of Maastricht, where all imaging took place.

**Stimuli.** Image sequences used in this experiment were generated with Random Image Structure Evolution (RISE) methods (Sadr and Sinha 2004) that implemented a manipulation of the spatial structure of the original images in which the original power spectrum as well as the overall luminance and contrast were kept constant. Specifically, based on an original image, a RISE sequence was generated by combining progressively degraded/randomized phase spectrum and the original intact amplitude spectrum. This resulted in a sequence of images in which a recognizable object gradually evolved from randomness. Note all the images that belong to a sequence have identical amplitude spectrum and overall luminance/contrast (see Fig. 3 in Sadr and Sinha 2004).

A total of 24 face and 24 car grayscale images from the Corel CD-ROM libraries and a few additional digitized pictures were selected as the original images [stimuli from a larger set as used by VanRullen (2006); available as figures online ([http://www.nefy.ucl.ac.be/Face\\_Categorisation\\_Lab.htm](http://www.nefy.ucl.ac.be/Face_Categorisation_Lab.htm))] to create the stimuli for this study. These images contained faces and cars that are highly variable in their visual appearance, size, and spatial location. These highly

variable visual scenes were used rather than well-segmented full-front pictures of faces and cars for two reasons. First, in a visual face detection task, it is more ecological to use variable visual scenes than full-front faces that could be rapidly and easily detected on the basis of a few cues (e.g., contour) if they always appeared at the same location. Second, using visual scenes greatly minimizes the predictability of the overall structure of the target stimuli, and thus the potential contribution of perceptual expectations on the BOLD signal in the areas of interest (Puri et al. 2009; Summerfield et al. 2006). At the same time, we should acknowledge that using faces embedded in complex scenes does introduce the possibility that some of our effect is driven more by “person perception” (i.e., face + body), rather than by faces per se, a point we will return to in the discussion of our results.

Visual scenes containing cars were chosen as contrast stimuli to faces, for several reasons. First, cars constitute complex visual stimuli, which, like faces, can be categorized on the basis of their particular outer shape and some of their internal elements (e.g., wheels, grid, etc.). Second, both faces and cars form a particularly visually homogeneous category. Third, images of cars, similar to images of faces, are highly familiar in visual scenes for our observers. Fourth, pictures of cars are from one of the categories commonly compared with faces in behavioral and neural studies aimed at testing for face-specific effects (e.g., Gauthier et al. 2000; Grill-Spector et al. 2004; Lerner et al. 2001; Yin 1969). Categorization of visual scenes based on the presence of cars has also been commonly used and shown to be quite efficient and fast (e.g., Peelen et al. 2009; VanRullen and Thorpe 2001). Finally, to define the face-sensitive regions of interest (ROIs), we have been using pictures of (segmented) cars as control stimuli in face localizer scans (e.g., Jiang et al. 2009; Rossion et al. 2011), including the localizer used in the present study. Thus both the definition of face-sensitive areas performed in the localizer and the test of face sensitivity of these areas to gradually revealed visual scenes rely on a comparison between the same categories of stimuli (faces vs. cars), although different images were used in the localizer and experimental runs.

Before implementing RISE, we equalized the luminance across all 48 original images. Each RISE sequence included 15 frames, ranging from 50 to 10% interpolation of the original and random phase spectra, in steps of 2.86% (Fig. 1). An interpolation level of 0% would correspond to unaltered phase spectrum of the original image, and an interpolation level of 100% would correspond to a random phase

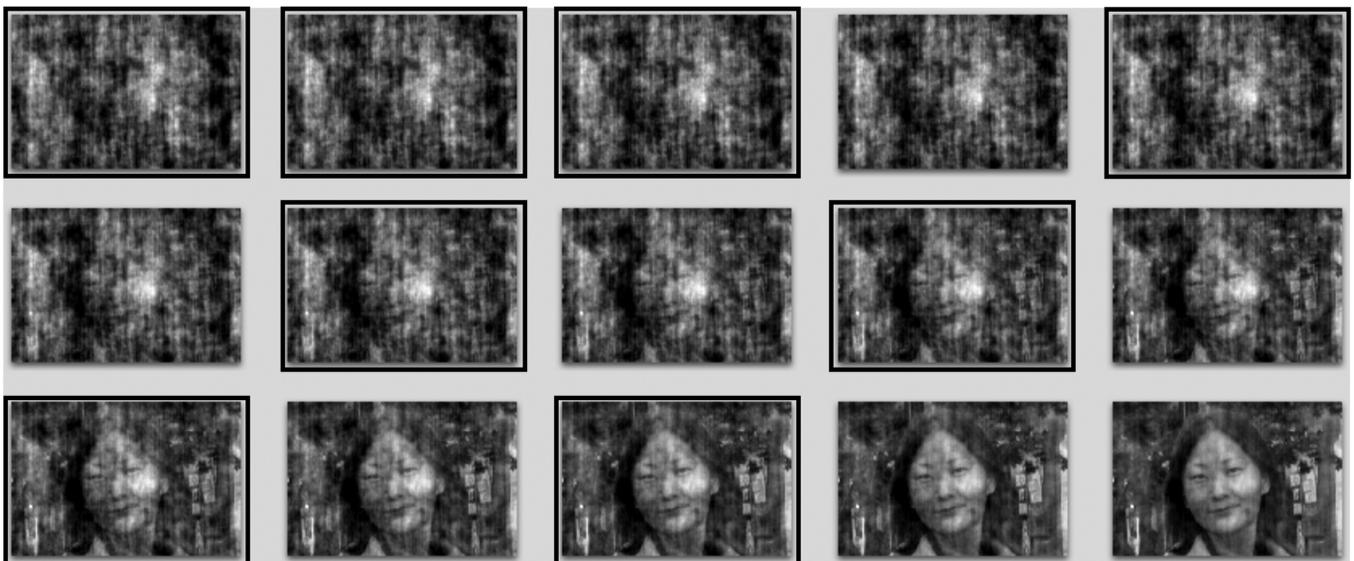


Fig. 1. Examples of a Random Image Structure Evolution (RISE) sequence, consisting a total of 15 frames, with interpolation level starting from 50% to 10% in steps of 2.86% (face sequence). The 8 frames used in the current study are framed in black.

spectrum. We have indicated the eight frames that were used in the experiment by framing them in black (Fig. 1).

**Design and procedure.** The experiment consisted of presentation of eight-frame RISE sequences, in which a recognizable object (face or car) gradually evolved from degraded images. Participants were asked to press one of the two buttons as soon as they were sure that they detected a face or a car during the presentation of the sequences. Participants were also requested to maintain a constant level of confidence in their judgment across trials. They were informed that only one target object (face or car) was present in each trial and that target objects could vary in size, appearance, and their spatial location within the image. Note that after the participant indicated having detected a face or a car, the presentation of the sequence continued. Participants were allowed to correct themselves, if they realized that they had made a mistake, by pressing the correct button before the next trial started.

Each participant performed two experimental runs. Each run lasted ~12 min. For each run, a total of 24 trials were presented (12 face trials and 12 car trials). Trials were presented randomly for each participant so that they could not anticipate whether the next trial would contain a face or car, also minimizing the potential contribution of perceptual expectation factors to the observation of category-sensitive responses. Each trial contained a unique 20-s sequence of 8 images, each for 2.5 s (2 TRs). The sequence was followed by 2.5 s of blank screen and a long fixation before the start of the next sequence. The duration of fixation varied between 5,000, 6,250, and 7,500 ms. The onsets of any two subsequent trials, therefore, were separated by an average of 8,750 ms (7,500–10,000 ms/6–8 TRs).

Stimuli were back-projected onto a screen located over the participant's head. A personal computer running E-prime 1.1 (PST) was used to present stimuli and collect behavioral responses. All stimuli were presented in grayscale. The images were always presented in the center of the screen and subtended approximately a visual angle of 8.54°.

**Localizer scans.** Independent localizer scans were performed to localize areas responding preferentially to faces. Each participant conducted two runs, in which they viewed blocks of faces, cars, phase-scrambled faces, and phase-scrambled cars, and performed a one-back matching task. Each run lasted 11 min and consisted of 24 alternating blocks (18 s each) with 9 s of fixation in between. During each block, 18 images were presented for 750 ms, followed for 250 ms by a blank screen. All images of faces and cars were presented in color with equalized luminance, and the scrambled versions were created with Fourier phase randomization (Sadr and Sinha 2004). Note that images of faces and cars used in the localizer scans, unlike those used in the experimental scans, were all segmented (see Jiang et al. 2009 and Rossion et al. 2011 for details).

**Image acquisition.** All participants were scanned at the Maastricht Brain Imaging Center. Data were collected using a 3T Allegra head scanner (Siemens, Erlangen, Germany). Functional data in the localizer scans were obtained from 36 transverse slices with a spatial resolution of 3.5 × 3.5 × 3.5 mm (acquisition matrix 64 × 64) using a repeated single-shot echo-planar imaging sequence [echo time (TE) = 50 ms, repetition time (TR) = 2,250 ms, flip angle (FA) = 90°, field of view (FOV) = 224 mm]. Functional data in the experimental scans were obtained from 20 transverse slices, with a spatial resolution of 3.5 × 3.5 × 3.5 mm for 4 participants and a spatial resolution of 3.5 × 3.5 × 5 mm for the remaining 8 participants. To have a relatively good temporal resolution, TR was set to 1,250 ms for experimental scans. High-resolution structural images were obtained with 1 × 1 × 1-mm spatial resolution (acquisition matrix 256 × 256), using ADNI sequence (TE = 2.6 ms, TR = 2,250 ms, FA = 9°, FOV = 256 mm). These T1-weighted images provided detailed anatomical information. A 25° angle perpendicular to the main magnetic field B<sub>0</sub> was used to reduce magnetic artifacts and signal dropout, allowing us to record up to the anterior inferior temporal lobe (Deichmann et al. 2003).

**fMRI data preprocessing.** Data were analyzed using Brain Voyager QX (version 1.10; Brain Innovation, Maastricht, The Netherlands). The first four volumes of each functional dataset were discarded to cope with T2\* contrast saturation effect. Before statistical analysis, the functional data underwent a series of preprocessing steps, namely, slice scan time correction, three-dimensional motion correction (with realignment to the first volume), linear trend removal, and high-pass filtering (removing frequencies lower than 3 cycles/session, ~0.004 Hz for experimental runs and 0.005 Hz for localizer runs). Functional data were further smoothed with a 5-mm full-width half-maximum Gaussian kernel so that the spatial resolution was comparable not only between the localizer scans and the experimental scans but also among participants. Both anatomical and functional data were transformed into Talairach space (Talairach transformation; Talairach and Tournoux 1988).

**Regions of interest.** Areas responding preferentially to faces were defined independently for each individual participant from the localizer scans by using the contrast [faces – cars] in conjunction with the contrast [faces – scrambled faces]. This conjunction analysis ensured that the activation in face-sensitive regions was not related to low-level differences between faces and non-face object categories. For each participant, all contiguous voxels in the middle fusiform gyrus and inferior/middle occipital gyrus, with a minimum significance of  $q(\text{false discovery rate, FDR}) < 0.001$ , were selected (FFA and OFA, respectively). We raised statistical threshold for four participants to separate their overlapping fusiform and inferior occipital activation. We also lowered the  $q(\text{FDR})$  to 0.01 and 0.05, respectively, for two participants to localize OFA so that the smallest cluster size is ~60 mm<sup>3</sup> (Fox et al. 2009).

Using the same method, we also selected several additional ROIs that we could identify on most of our participant, including the right pSTS, bilateral amygdala, and the right inferior frontal gyrus (IFG). These additional face-preferential ROIs were defined using a minimum statistical threshold of  $q(\text{FDR}) < 0.05$ . The mean Talairach coordinates, cluster size, and their standard deviation of all face-preferential ROIs selected in this study are reported in Table 1.

**Time-course analysis.** To test our hypothesis, we examined the onset of significant difference between responses to face and car sequences in individually defined ROIs. Specifically, for each subject and each ROI, event-related averages were computed across all correct trials for face and car conditions. Using planned pairwise *t*-tests, BOLD responses to face and car sequences were then compared at group level to determine the time point at which responses became face sensitive (i.e., significantly higher to face sequences than to car sequences). A time point with a significant difference at the  $P < 0.05$  level was determined as the onset only when at least two subsequent time points were also significant at this level. This method of determining onset of difference between two time courses ensured both reliability and sensitivity. In addition, to demonstrate that the onset of difference was closely linked to participants' percepts, we repeated the analysis on a subset of trials in which participants responded the fastest (32%) and slowest (32%).

**Whole brain response-locked analysis.** We also conducted a whole brain-based response-locked analysis to reveal brain areas that were activated more for face sequences than for car sequences upon successful detection. The condition-dependent behavioral response was modeled as a simple two-gamma hemodynamic response function with a duration of 200 ms in a multiple-subject random-effect general linear model (GLM). Clusters that showed a significant difference ( $P < 0.005$  uncorrected, 1-tailed *t*-test, postcorrected by cluster size threshold >16) were reported. The face sensitivity of reported clusters was examined based on data extracted from localizer scans. A cluster was considered face sensitive when both contrasts (i.e., [faces – cars] and [faces – scrambled faces]) reached statistical significance.

To obtain a more accurate estimate of the fMRI activity within clusters from whole brain analysis, we ran an additional deconvolution analysis time-locked to the behavioral response. The deconvolu-

Table 1. Average Talairach coordinates, cluster size, and their SD of face-preferential ROIs

ROI	n	Talairach Coordinates			Mean Size, mm <sup>3</sup>	SD x	SD y	SD z	SD Size, mm <sup>3</sup>
		Mean x	Mean y	Mean z					
<i>Right hemisphere</i>									
FFA	12	39	-49	-15	403	2.7	6.1	2.8	204
OFA	11	33	-77	-10	245	8.8	8.9	5.3	167
Amygdala	7	20	-6	-10	285	2.7	1.9	2.3	195
pSTS	8	51	-45	10	658	3.8	10.4	5.7	370
IFG	10	43	11	25	540	5.2	7.2	7	282
<i>Left hemisphere</i>									
FFA	12	-39	-47	-10	429	2.8	7.1	3.9	281
OFA	9	-36	-77	-11	213	6.8	8.2	6	225
Amygdala	7	-19	-7	-9	216	1.9	1.7	1.5	115

Values are means and SD ( $n$  = no. of subjects) of Talairach coordinates ( $x, y, z$ ) and cluster size of face-preferential regions of interest (ROIs). FFA, fusiform face area; OFA, occipital face area; pSTS, posterior superior temporal sulcus; IFG, inferior frontal gyrus.

tion analysis included a total of 16 predictors, 4 before the response, 1 at the response, and 11 after the response. Each predictor had a duration of 1 TR (1,250 ms). To avoid overlap between subsequent trials, we further discarded trials in which behavioral response occurred too early (i.e., during *frame 3*) and too late (i.e., during wait period after the sequence, see Fig. 2C for response distribution). This procedure allowed us to estimate beta weights associated with experimental conditions (i.e., the coefficients of predictors) at each time point and compare them between conditions. For each cluster, we determined the time point at which responses became face sensitive (i.e., significantly higher beta weight in face condition than in car condition) using the same method we described in the ROI-based analysis section.

## RESULTS

**Behavioral results.** The paradigm with dynamic RISE sequences effectively slowed down participants' perceptions of faces (and cars): average correct response time was 14,084 ms for car sequences and 14,698 ms for face sequences (Fig. 2). Participants successfully detected 93% of the car trials and 89% of the face trials. The accuracy was very high, given that to have a precise estimate of the accuracy and correct response time, participants' self-corrected trials were not considered as correct trials. The difference between car trials and face trials just failed to reach significance for accuracy [ $F(1,11) = 4.12, P < 0.07$ ] but was significant for response times [ $F(1,11) =$

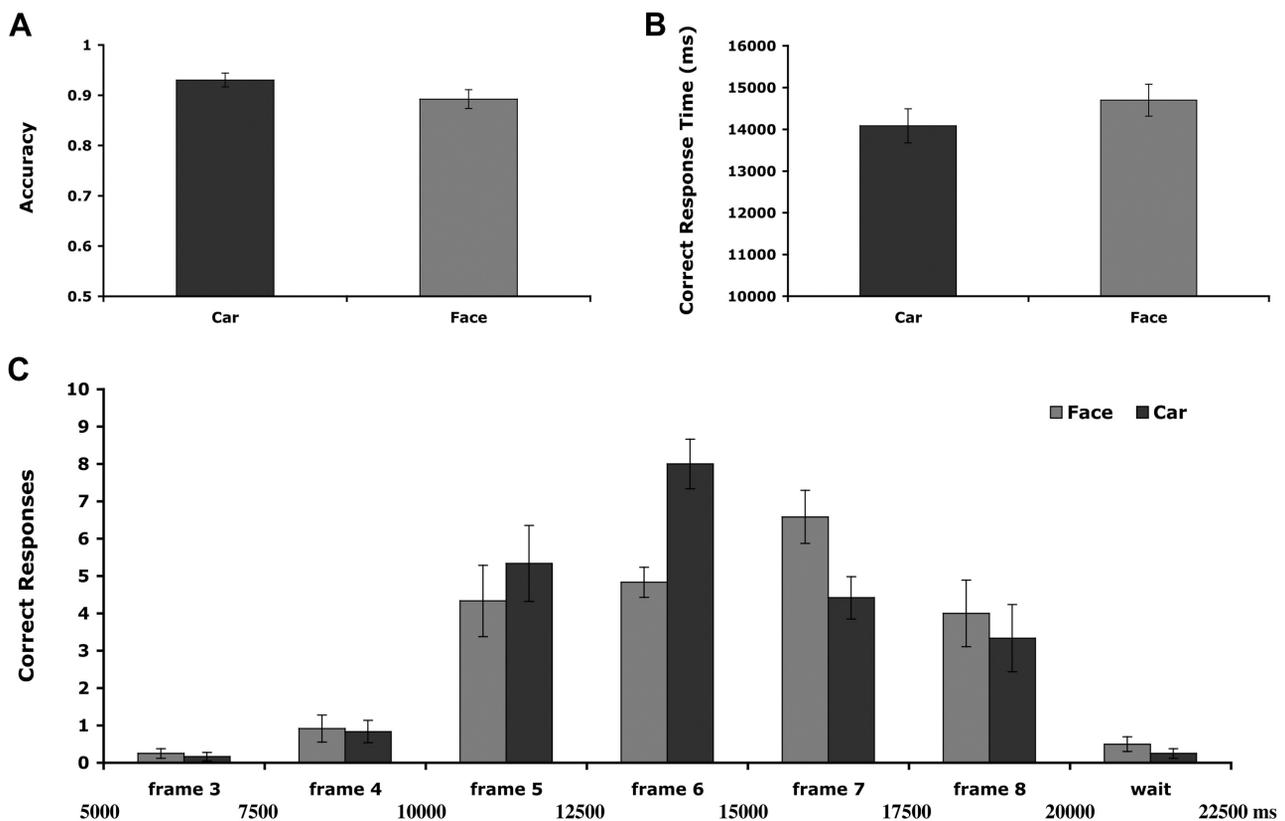


Fig. 2. Behavioral results. A: accuracy. B: correct response time. C: distribution of the number of correct responses as a function of frame (time). Note that responses were slightly faster for cars than for faces, most probably due to the particular limited set of stimuli used.

25.98,  $P < 0.001$ ], with faster response times for car trials than for face trials.<sup>1</sup> The distribution of correct responses as a function of frames (time) is shown in Fig. 2C.

For the 32% fastest trials, participants required on average  $< 12$  s to decide whether there was a face or a car present in the visual scenes ( $11,546 \pm 1,817$  ms for cars and  $11,742 \pm 2,065$  ms for faces, mean  $\pm$  SD response time). For slow trials, the same task required  $\sim 17$  s ( $16,862 \pm 2,025$  ms for cars and  $17,540 \pm 1,644$  ms for faces).

**ROI-based imaging results.** We analyzed time course data in each individually localized ROI (Table 1). Analyses were concentrated on the right FFA and the right OFA. The right FFA was identified for all 12 participants (mean Talairach coordinates: 39,  $-49$ ,  $-15$ ), and the right OFA was found for 11/12 participants (mean Talairach coordinates: 33,  $-77$ ,  $-10$ ). An example of individually localized right FFA and right OFA is shown in Fig. 3.

The time courses of fMRI activity to face and car sequences in the right FFA and right OFA are shown in Fig. 4. In both ROIs, the fMRI activity rose slowly as the face and car in the visual scene gradually emerged from randomness and peaked 2–3 s after the presentation of the last frame of the 20-s-long sequence.

First, we ran a three-way repeated-measures ANOVA with data from 11 participants who had both the right FFA and right OFA localized, using the factors category (faces vs. cars), area (right FFA vs. right OFA), and time (25 time points, starting from the 0-ms sequence time and with each time point taking 1,250 ms). All the effects were significant (all  $P < 0.002$ ), with the exception of the two-way interaction between area and category [ $F(1,10) = 1.47$ ,  $P = 0.25$ ], suggesting that, overall, the right FFA did not show a stronger face-preferential response than the right OFA. However, the three-way interaction between category, area, and time was highly significant [ $F(24,240) = 2.17$ ,  $P = 0.0018$ ]. Based on this highly significant three-way interaction, we analyzed the effect of category in each area and time point separately for the 12 participants in the right FFA and the 11 participants in the right OFA.

Importantly, fMRI activity in response to face sequences in the right FFA rose significantly above the activity in response to car sequences at 13.75 s after sequence onset [ $t(11) = 2.92$ ,  $P < 0.03$ , 1-tailed], whereas this difference emerged much later in the right OFA, at 22.5 s after sequence onset [ $t(10) = 3.31$ ,  $P < 0.004$ , 1-tailed]. These findings support the prediction that face sensitivity emerges earlier in the right FFA than in the right OFA. It is important to note that if we were to consider the time point at which fMRI activity in response to face sequences rose significantly above the baseline level, the right OFA would be considered to be earlier [at 8,750 ms,

<sup>1</sup> The slower response to faces presented in visual scenes than to cars does not agree with the recent observations of faster saccadic responses to faces (Crouzet et al. 2010). This is not a fundamental issue for the goal of the present study, and it might be due to several factors, for instance, the kind of paradigm (flash vs. gradual revealing from scrambling of phase information) or the particular limited set of stimuli used in the present study, or the mode of response (saccadic response times in Crouzet et al. 2010 vs. manual response time). Also, Crouzet et al. (2010) compared faces with images in more broadly defined categories (animals or vehicles) than used presently (cars), so the responses to such categories with less well-predictable structures might have been slowed down in their study. As a matter of fact, manual response times to animal and human detection in visual scenes did not differ in a previous study (Rousselet et al. 2003).

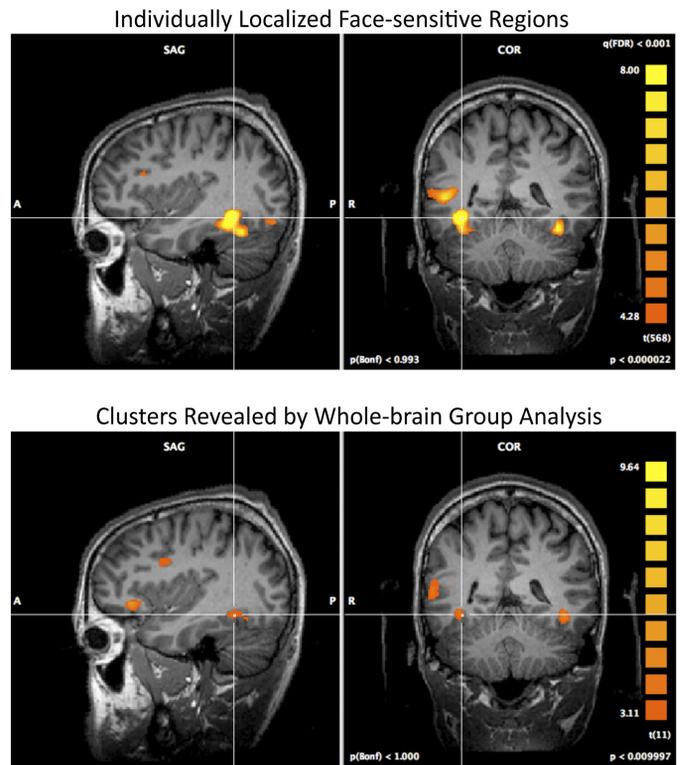


Fig. 3. *Top*: an example of individually localized right fusiform face area (FFA) and right occipital face area (OFA). The right FFA of a representative participant (cross coordinates: 37,  $-45$ ,  $-11$ ) and her more posteriorly located right OFA are shown, along with the right inferior frontal gyrus [IFG; sagittal section (SAG)] and the right posterior superior temporal sulcus [pSTS; coronal section (COR)]. *Bottom*: clusters revealed by response-locked whole brain group analysis displayed on the brain of the same representative participant. The same cross as at *top* indicates the cluster in the right fusiform gyrus. The sagittal section also displays the cluster in the right IFG and the cluster in the right orbital gyri, whereas the coronal section shows the cluster in the left fusiform gyrus and the cluster centered at the middle occipital gyrus.

$t(10) = 2.26$ ,  $P < 0.004$ , 1-tailed] than the right FFA [at 12,500 ms,  $t(11) = 3.28$ ,  $P < 0.024$ , 1-tailed]. However, the early responses in the right OFA were not face sensitive, i.e., they were not larger for faces than for cars in the visual scenes.

To demonstrate that the rise time of face sensitivity is closely linked to participants' percepts (e.g., McKeef and Tong 2007), we repeated the time-course analysis on the subsets of trials for which participants gave the fastest and slowest responses (see *Behavioral results* for trial selection). As shown in Fig. 5, in the right FFA, fMRI activation in response to face sequences became significantly higher than that to car sequences at 12.5 s in fast trials [ $t(11) = 1.86$ ,  $P < 0.045$ , 1-tailed] and at 18.75 s in slow trials [ $t(11) = 1.99$ ,  $P < 0.036$ ], consistent with the average behavioral response time required for participants to reach a decision in these trials. In the right OFA, a significant difference between activation in response to face and car sequences did not occur until several seconds later for both fast trials [18.75 s,  $t(10) = 1.92$ ,  $P < 0.042$ ] and slow trials [22.5 s,  $t(10) = 2.07$ ,  $P < 0.033$ ]. These results strengthen the finding of an earlier face sensitivity in the right FFA than in the right OFA and also indicate that the time onset of these differences is relevant with respect to behavior.

We also analyzed the time course in other face-sensitive ROIs, including the left FFA and OFA, the right pSTS and IFG, and the bilateral amygdala (Fig. 6). In both the left FFA

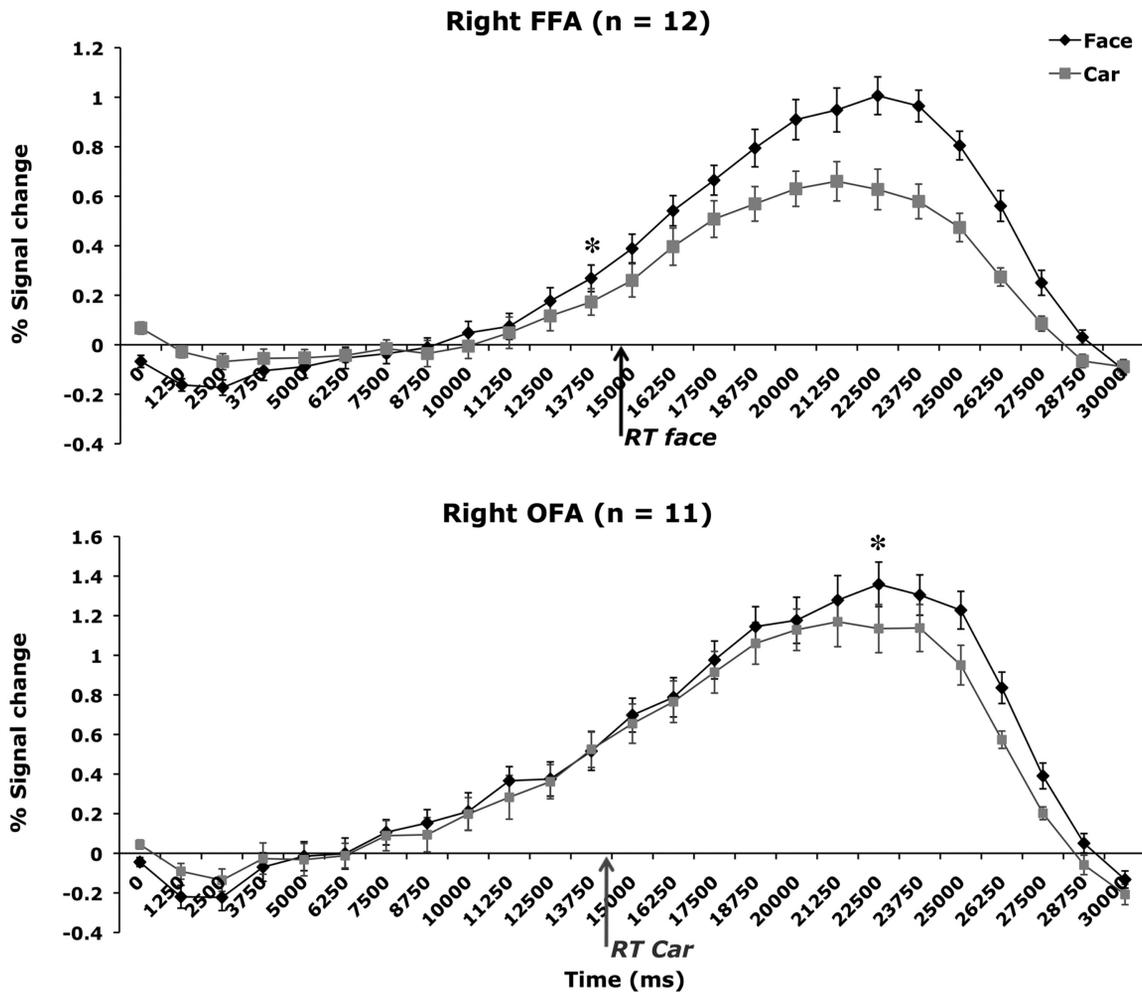


Fig. 4. Averaged time-course data in individually localized right FFA and right OFA. Asterisks indicate the starting time points at which a face sensitivity (i.e., significantly higher activation for visual scenes containing a face than for scenes containing a car) emerged. Note that the onset time of face sensitivity in the FFA slightly precedes the average behavioral response times (RT) to detect a face. Although this difference could reflect the earlier discrimination sensitivity of the area with respect to behavior, this early sensitivity in the signal could be driven by a significant proportion of trials that were responded to below 12 s (see Fig. 5).

(mean Talairach coordinates:  $-39, -47, -15$ ; 12/12 participants) and the left OFA (mean Talairach coordinates:  $-36, -77, -11$ ; 9/12 participants), face-preferential responses occurred at 15 s after the onset of sequence [ $t(11) = 2.55, P < 0.013$ , 1-tailed, and  $t(8) = 2.35, P < 0.023$ , respectively]. In both the right pSTS (mean Talairach coordinates:  $51, -45, 10$ ; 8/12 participants) and the right IFG (mean Talairach coordinates:  $43, 11, 25$ ; 10/12 participants), face-preferential activation emerged at 18.75 s after sequence onset [ $t(7) = 3.34, P < 0.006$ , and  $t(9) = 2.19, P < 0.028$ , respectively]. Because of the relatively low number of participants ( $n = 7/12$ ) on which we could localize bilateral amygdala, their time course data showed large variability, resulting in no significant starting point of face-preferential activity based on our criteria (i.e., at least 3 consecutive time points with higher activation in response to face sequences than to car sequences at the  $P < 0.05$  level). Overall, these results indicate that among the face-sensitive areas identified in the current study, the right FFA was the first to show a face-preferential response.

**Whole brain-based imaging results.** A whole brain random-effect GLM analysis revealed seven clusters that showed significantly higher activation when a face, compared with a car,

was successfully detected (Table 2). Beta weights in these clusters were further estimated through a deconvolution operation, which consisted of 4 predictors before and 11 predictors after the behavioral response (Fig. 7, left). Among these seven clusters, four were determined to respond preferentially to faces based on data extracted from localizer scans, including clusters in bilateral fusiform gyrus, right IFG, and right middle occipital gyrus extending to STS and the middle temporal gyrus. Note the consistency between the face-sensitive cluster in the right fusiform gyrus revealed by the whole brain group analysis (Fig. 3, bottom) and the right FFA localized on one individual participant based on independent localizer scans (Fig. 3, top).

Consistent with ROI-based results, the face-sensitive cluster in the right fusiform gyrus (*cluster 2*, center Talairach coordinates:  $39, -46, -11$ , see Fig. 3, bottom) responded significantly higher for face than car sequences, shortly before successful detection [ $t(11) = 2.12, P < 0.034$ , Fig. 7]. This condition-specific difference emerged only after the behavioral response in the other three face-sensitive clusters. Specifically, *cluster 1* in the right IFG (center Talairach coordinates:  $39, 8, 29$ ) showed a significant difference at 5,000 ms (i.e., 4 TRs)

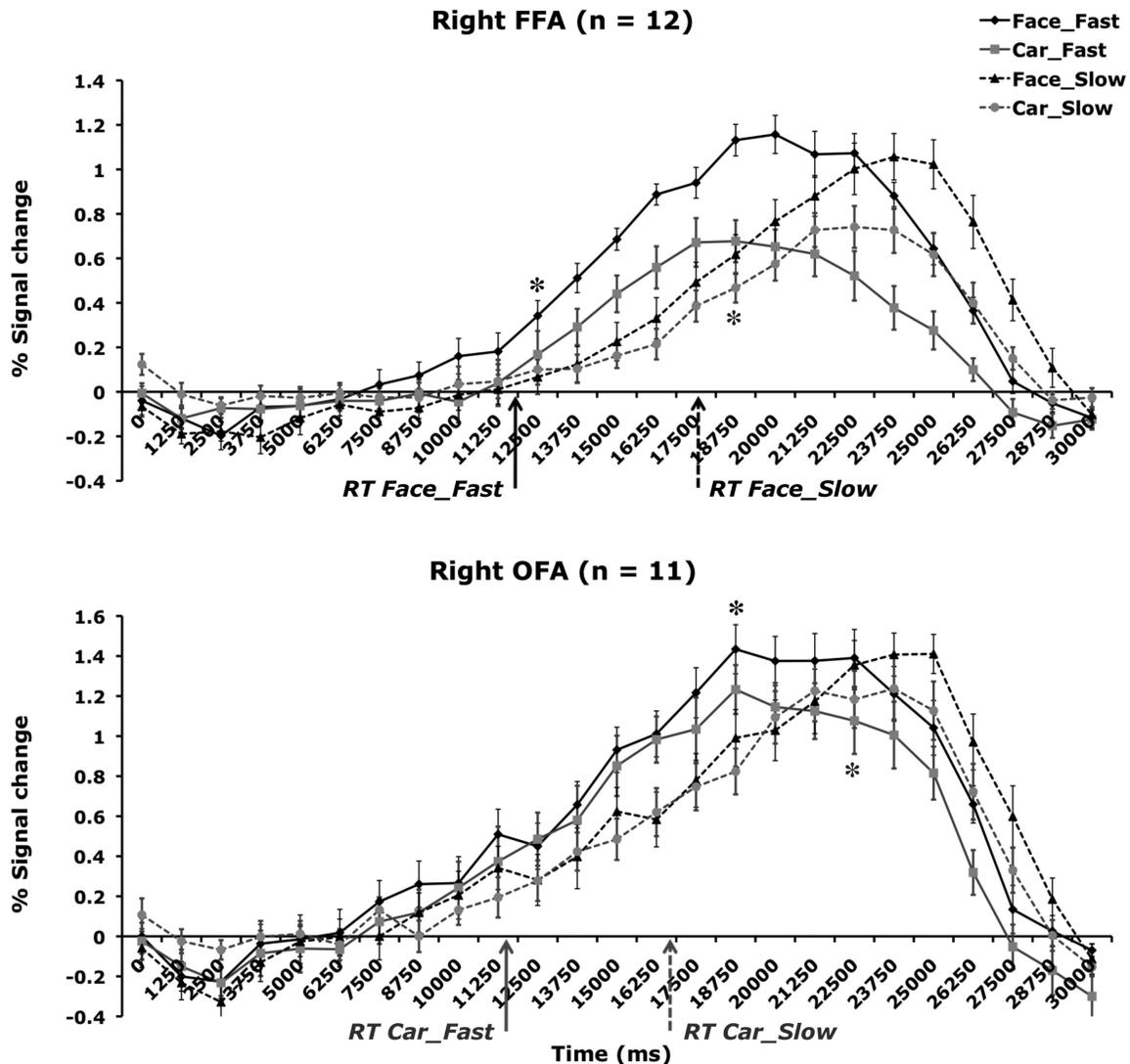


Fig. 5. Averaged time-course data separated for fast and slow trials (i.e., 32% trials in which participants give the fastest or slowest correct responses) in individually localized right FFA and right OFA. Asterisks indicate the starting time points at which a face sensitivity (i.e., significantly higher activation for visual scenes containing a face than for scenes containing a car) emerged. The mean RTs are indicated.

after behavioral response [ $t(11) = 2.29$ ,  $P < 0.022$ ]. For *cluster 3* in the left fusiform gyrus (center Talairach coordinates:  $-41, -44, -12$ ), this difference did not become significant until 3,750 ms (3 TRs) after behavioral response [ $t(11) = 2.28$ ,  $P < 0.023$ ]. Finally a significant difference was shown for *cluster 7* centered in middle occipital gyrus (center Talairach coordinates:  $52, -60, 7$ ) at 2,500 ms (2 TRs) after behavioral response [ $t(11) = 3.34$ ,  $P < 0.001$ ]. Critically, these results indicate that among face-sensitive clusters revealed by time-locked whole brain analysis, the cluster in the right fusiform gyrus was the earliest to show face preference.

Interestingly, a non-face-sensitive cluster located in the left inferior occipital gyrus (*cluster 4*, center Talairach coordinates:  $-46, -75, 3$ ) illustrated similar prior-response face sensitivity [ $t(11) = 2.66$ ,  $P < 0.008$ , Fig. 7] to the right middle fusiform gyrus/right FFA. Its center Talairach coordinates correspond to the dorsal posterior portion of LOC, which has been shown to be important for object perception and recognition (e.g., Grill-Spector et al. 1999; Malach et al. 1995), and more specifically to a left extrastriate area that responds preferentially to body

parts (extrastriate body area, EBA; Downing et al. 2001; e.g., Peelen et al. 2006:  $-45, -71, -1$ ). Although this cluster showed a stronger activation to both faces and cars compared with its scrambled counterparts [ $t(11) = 3.66$ ,  $P < 0.004$ ], its level of activation was similar to faces and cars in the localizer scans [ $t(11) = 0.272$ , not significant]. Thus this cluster did not show face sensitivity in the face localizer (consisting of isolated images of faces without body parts), consistent with the response profile of EBA.

To summarize our results, we observed the earliest category-related response in a high-order visual area, namely, the middle fusiform gyrus of the right hemisphere, corresponding to the FFA. Other areas defined in the independent localizer by their preferential response to faces, most notably the lower order right OFA, also showed category-related responses during categorization of visual scenes. However, this larger response to faces than cars in visual scenes took place significantly later than in the right FFA.

These observations cannot be accounted for by the slightly faster behavioral responses to cars than faces overall, since the

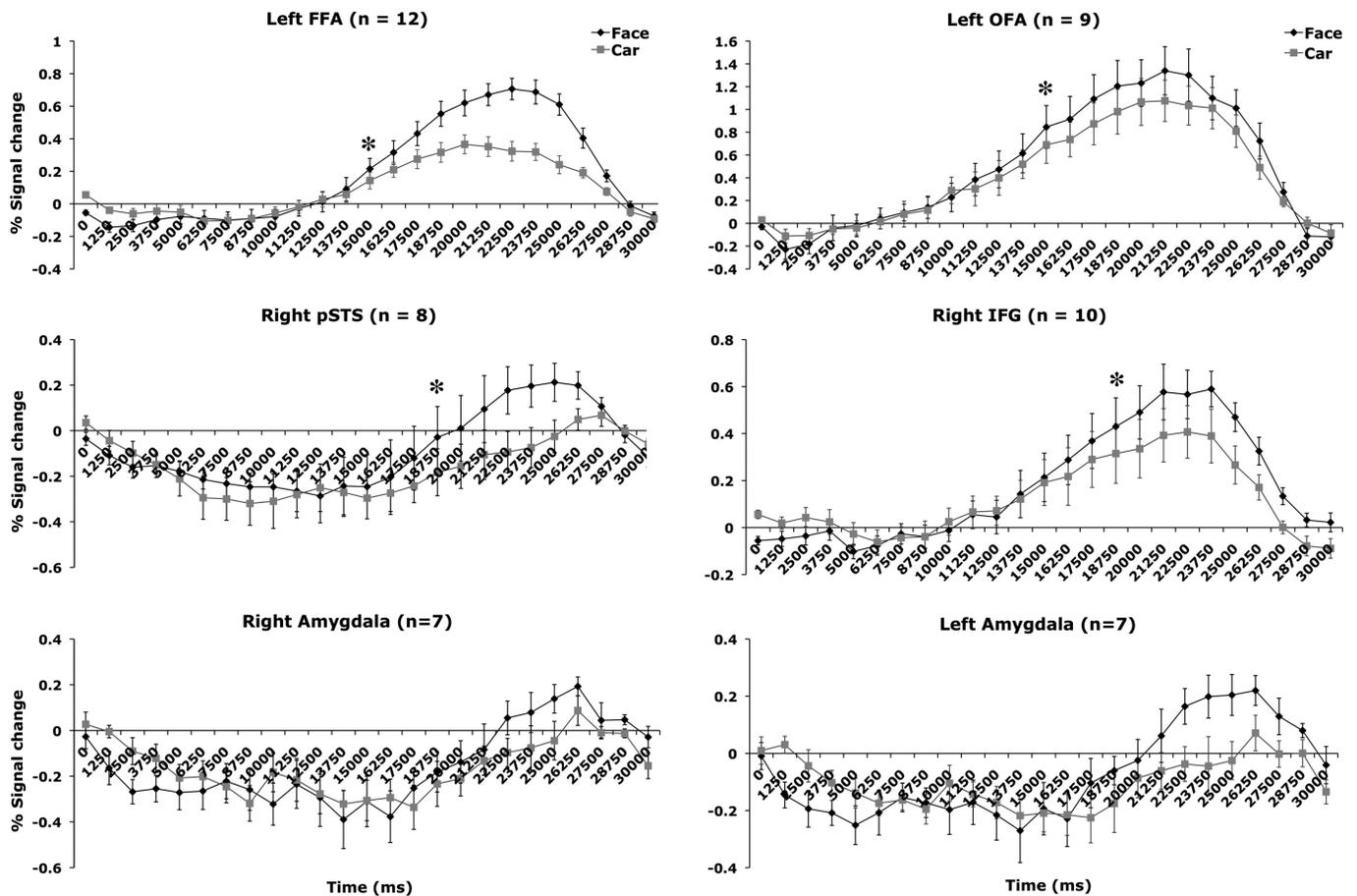


Fig. 6. Averaged time-course data in individually localized regions of interest (ROIs) other than the right FFA and right OFA. *n* indicates the number of participants for which the ROI was defined in the localizer runs. Asterisks indicate the starting time points at which a face sensitivity (i.e., significantly higher activation for visual scenes containing a face than scenes containing a car) emerged.

analyses of the time course and sensitivity of different brain areas were based on the exact same trials. Also, the size of the different face-sensitive ROIs, as well as the number of participants in which individual ROIs could be identified in the face localizer (Table 1), did not appear to be related to the time course of face-sensitive responses observed in the dynamic sequence stimulation. Indeed, both the right FFA and OFA could be identified in almost all individual participants (12 and

11, respectively). Moreover, the left FFA was equally large, and the right pSTS even larger, in size than the right FFA, yet both of these areas showed a delayed sensitivity to faces presented in dynamic sequences compared with the right FFA.

Also worth noting is that the earlier face-preferential activation in the right FFA than in the right OFA cannot be accounted for by the absolute magnitude of the BOLD response or by the absolute time onset of activation in the respective areas of interest. In fact, the hemodynamic response was larger and rose earlier in the right inferior occipital cortex (right OFA) than in the right middle fusiform gyrus (right FFA). However, the significantly larger response to faces than cars emerged later in right OFA than in the right FFA (Fig. 4).

Finally, the significance of the earlier face-preferential response found in the right FFA is also supported by the relation found between the speed of behavioral categorization and the time course of face-preferential activation in the right FFA (and other areas, Fig. 5). Specifically, in both the right FFA and right OFA, visual scenes that were categorized relatively rapidly as containing faces led to an earlier preferential response to faces than visual scenes that were categorized more slowly. Importantly, for both fast and slow trials, preferential activation to faces arose earlier in the right FFA than in the right OFA.

*Additional object-sensitive ROI-based imaging results.* The finding of a non-face-sensitive cluster in the whole brain

Table 2. Significant clusters revealed by whole brain random-effect GLM analysis for contrast face response – car response

Cluster	Talairach coordinates			Size, mm <sup>3</sup>	Location	Face Preference
	x	Y	z			
1	39	8	29	446	Right inferior frontal gyrus	Yes
2	39	-46	-11	467	Right fusiform	Yes
3	-41	-44	-12	556	Left fusiform	Yes
4	-46	-75	3	581	Left inferior occipital gyrus	No
5	36	31	-4	899	Right orbital gyri/sulcus	No
6	-29	30	-5	529	Left orbital gyri/sulcus	No
7	52	-60	7	6478	Right STS/MTG/MOG	Yes

Values are Talairach coordinates and cluster sizes for clusters revealed by whole brain random-effect general linear model (GLM) analysis to be significantly face sensitive [*P* < 0.005, uncorrected, 1-tailed *t*-test: *t*(11) = 3.11], postcorrected by cluster size threshold of 16]. MTG, middle temporal gyrus; MOG, middle occipital gyrus.

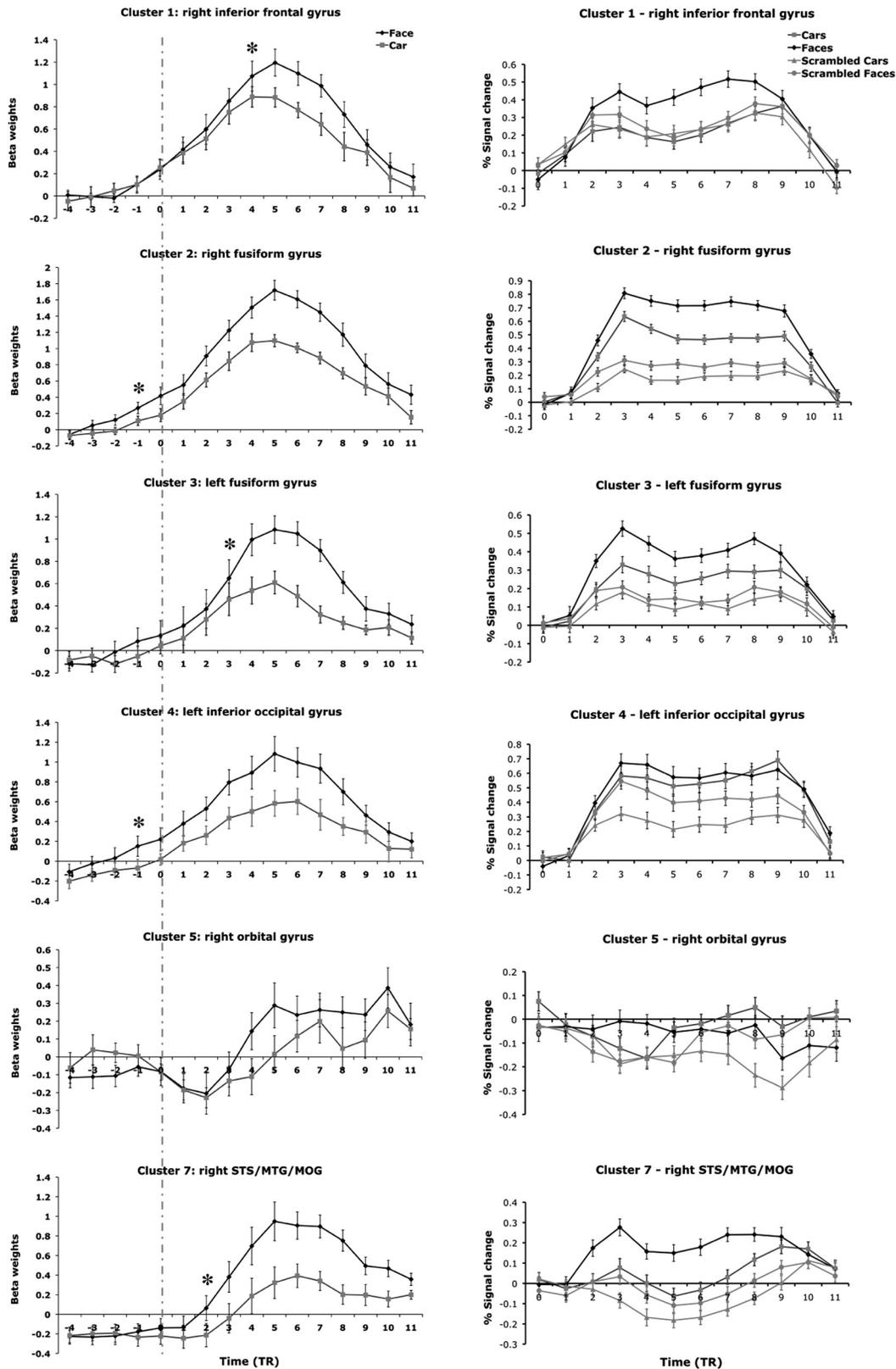


Fig. 7. Activation profile of clusters revealed by whole brain random-effect analysis. The beta weights estimated from the experiment scans are shown at *left*, and the time courses estimated from the localizer scans are shown at *right*. For each cluster, condition-related beta weights were estimated through a deconvolution operation with predictors aligned to participants' behavioral responses (indicated by the gray line). Asterisks indicate the starting time points at which a face sensitivity (i.e., significantly higher beta weight in face condition than in car condition) emerged. Note that the profile of the left amygdala (*cluster 6*) was very similar to that of the right amygdala (*cluster 5*) and thus is not shown. MTG, middle temporal gyrus; MOG, middle occipital gyrus.

analysis (left inferior occipital gyrus, *cluster 4*) led to a further examination of the time course in individually defined object-sensitive areas. Two subregions of the LOC (Grill-Spector et al. 1999), namely, the lateral occipital area (LO) and posterior fusiform area (PF), were localized for each participant using contrast [cars - scrambled cars]. We further excluded the overlap between face-sensitive and object-sensitive regions by eliminating voxels in LOC that showed higher responses for faces than for cars. We also localized a region in the collateral sulcus (near parahippocampal place area, PPA/BA 20; Epstein and Kanwisher 1998) using the contrast [cars - faces] in conjunction with [cars - scrambled cars].

As shown in Fig. 8, consistent with the results from whole brain analysis, the left LO (localized in 11/12 participants; mean Talairach coordinates:  $-40, -75, -10$ ) illustrated relatively early face sensitivity, showing significantly higher activation for face sequences than for car sequences at 15 s after sequence onset [ $t(10) = 2.65, P < 0.013$ ]. The onset of face sensitivity found in the left LO closely followed the onset of face sensitivity in the right FFA, indicating the possible role of LO in subsequent analysis of the scene images once the category of scene has been determined.

Interestingly, a reverse pattern was found in the right PPA/BA20 (localized in 9/12 participants; mean Talairach coordinates:  $28, -34, -16$ ), with significantly higher activation for car sequences than for face sequences at 17.5 s after the sequence onset [ $t(8) = 1.9, P < 0.05$ ]. Although the left PPA/BA20 (localized only in one-half of the participants; mean Talairach coordinates:  $-31, -35, -14$ ) showed a similar trend, the difference between car and face activation did not reach significance at any given time point. Parallel to the finding in the right FFA, the finding of an inverse pattern in the right PPA/BA20 suggests the possible emergence of car sensitivity in high-level regions that respond more to objects than faces.

## DISCUSSION

*Face sensitivity may begin in the higher order right FFA.* Our results support the possibility that the categorization of a visual stimulus as a face may be initiated in a higher order visual area, specifically, the right FFA, rather than lower order visual areas, most notably the most posteriorly located area of the cortical face network, the OFA. More generally, our results

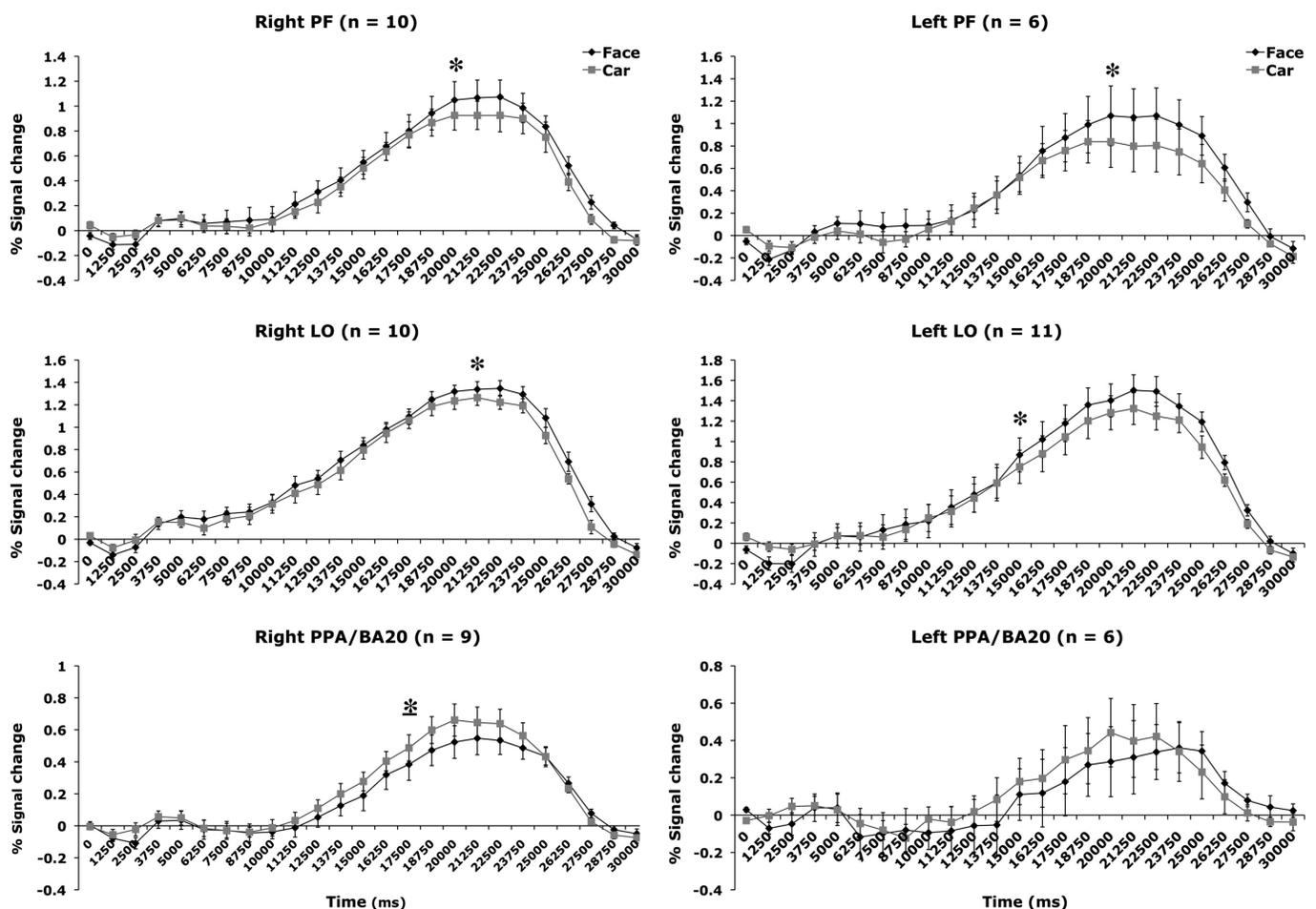


Fig. 8. Averaged time-course data in individually localized non-face-sensitive ROIs.  $n$  indicates the number of participants for which the ROI was defined in the localizer runs. Asterisks indicate the starting time points at which a face sensitivity (i.e., significantly higher activation for visual scenes containing a face condition than for scenes containing a car) emerged in the 2 subregions of the lateral occipital complex, namely, posterior fusiform area (PF) and lateral occipital area (LO). The underscored asterisk indicates the starting time point of a reverse pattern (i.e., significantly higher activation for visual scenes containing a car than for scenes containing a face) emerged in the car-sensitive region near parahippocampal place area (PPA/BA20).

are also consistent with the established dominance of the right hemisphere in face perception, as demonstrated previously by a greater prevalence of prosopagnosia following right than left hemisphere brain damage (e.g., Hécaen and Anguelergues 1962), better performance in lateralized face detection and individualization tasks (e.g., Parkin and Williamson 1987 and Hillger and Koenig 1991, respectively), higher magnitude of brain response to faces (e.g., Rossion et al. 2011; Sergent et al. 1992; Zangenehour and Chaudhuri 2005), and better sensitivity to facial information (e.g., Jiang et al. 2009; Le Grand et al. 2003; Schiltz and Rossion 2006) in the right than the left hemisphere. Beyond this general right hemisphere bias in face perception, our present data indicate that face sensitivity also emerges earlier in the right than in the left hemisphere (i.e., right FFA before left FFA).

Consistent with our claims, in the normal human brain, visual stimuli that are successfully categorized as faces based on prior knowledge or their global configuration elicit activation in the right FFA but not in the OFA (Dolan et al. 1997; Rossion et al. 2011). Also, structural damage to the cortical territory of the OFA does not prevent the observation of robust face-preferential responses in the right FFA (Rossion et al. 2003; Steeves et al. 2006). These observations are in agreement with our present findings in that they indicate that the right OFA is not a mandatory stage to observe face-preferential activation in the right FFA. However, the present findings go further than previous observations by showing that when both areas respond preferentially to faces, the earliest effects are found in the right FFA, not in the ipsilateral OFA.

Taken in toto, we argue that our present results, in conjunction with prior studies, constitute converging evidence that the OFA may not be the first point of face-specific processing in the human brain. Contrary to the most commonly held perspective, we suggest a nonhierarchical model of the early stages of the functional neuroanatomy of face processing. That is, the OFA would exhibit face-preferential responses only following temporally earlier categorization of visual input as a face within the FFA. Although the precise mediator that leads to OFA responses is unknown, we posit that such neuroanatomically earlier responses may arise through putative reentrant connections between these two areas (Rossion et al. 2003; Rossion 2008).

*Evaluating evidence for a hierarchical view of face processing.* What evidence is there that face-sensitive visual information would be fed from lower order areas to higher order areas? Does such evidence present an alternative to our claim that face sensitivity initially arises in the higher order right FFA?

First, when images of faces (and objects) are broken into an increasing number of parts (blocks) that are positioned randomly on an image, fMRI activation decreases along a posterior-anterior axis in the LOC (Lerner et al. 2001). The anterior region corresponding to the FFA shows the highest sensitivity to this breaking into parts (i.e., needing almost the whole stimulus to respond). Although this observation is usually taken to support the hierarchical view, it is also entirely compatible with initial face sensitivity in the right FFA, providing that the first representation of the face is holistic rather than part-based, an issue that we discuss below.

Second, the only strictly feedforward model of the OFA to the FFA obtained using effective connectivity on fMRI data

(dynamic causal modeling, DCM; Friston et al. 2003) is based on a definition of the OFA [faces – scrambled faces] that does not isolate at all face-preferential responses in the inferior occipital cortex (Fairhall and Ishai 2007; see Wiggett and Downing 2008). In fact, the observation that generic (i.e., non-face sensitive) lateral occipital activation also precedes FFA activation in our study (Fig. 4) is compatible with this model. However, what matters for our purpose is that preferential response to faces emerges earlier in the FFA than in the OFA.

Third, transcranial magnetic stimulation (TMS) applied over an average coordinate of the right OFA at a very early stage in time (60–100 ms following visual stimulation) was found to disrupt individualization of faces differing by facial parts (Pitcher et al. 2007). These findings have been taken as evidence supporting an early role for the OFA in face processing. However, disruption of individual face matching following TMS to the right OFA was not found in a follow-up study of the same group (Pitcher et al. 2008). It is also incompatible with the fact that face individualization occurs relatively later in processing, as indicated by electrophysiological recordings on the human scalp (e.g., Jacques and Rossion 2009; Jacques et al. 2007) and also the onset timing of face-selective cells in the monkey inferotemporal cortex (Rolls and Tovee 1995; see Rossion and Jacques 2011). Interestingly, Pitcher et al. (2010) recently provided evidence that TMS disruption of OFA processing at an early time window impairs generic visual categorization, whereas TMS applied at a later time window impairs face processing specifically. These findings are more in line with our current observations. To test our interpretation of these observations with TMS, one would have to apply TMS to the (right) OFA and search for impairment in face detection, as well as for a reduced or abolished neural responses to faces in the FFA (using TMS combined with fMRI).

Finally, a recent study found a correlation between the level activation in the OFA across participants with an early electrophysiological face-sensitive response recorded on the scalp (P1, ~100 ms), whereas the FFA was rather correlated with a later face-sensitive component (N170) (Sadeh et al. 2010). However, whereas the latter correlation was robust, the P1-OFA correlation appears to be driven mainly by a single data point. Moreover, there is evidence that such early face-sensitive responses (P1/M1) recorded on the scalp are based on low-level visual cues (e.g., spatial frequency differences between faces and objects) rather than on face perception per se (Halgren et al. 2000; Rossion and Caharel 2011; Rossion and Jacques, 2008). In summary, we argue that there is currently no strong counterevidence to the view that face-sensitive activation is initiated in the higher order visual FFA.

*The dynamic sequence paradigm: potential limitations and strengths.* To utilize fMRI to investigate (reverse) hierarchical processing in face perception, we acknowledge that we had to rely on an unconventional visual stimulation paradigm, which has both its limitations and strengths.

First, we stimulated with complex visual scenes rather than background-segmented face and non-face stimuli (as is typically done in psychophysical and neuroimaging studies studying object and face perception). We suggest that this kind of visual scene stimulation, in which the faces (and cars) are displayed at different sizes, from different views, at various positions in varying visual scenes, and with variable morphol-

ogy (download images online at [http://www.nefy.ucl.ac.be/facecatlab/PDF/jiang\\_figs\\_stim.zip](http://www.nefy.ucl.ac.be/facecatlab/PDF/jiang_figs_stim.zip)), is somewhat more ecological and better approximates the actual context in which humans typically perform visual object and face recognition (see Crouzet et al. 2010; Peelen et al. 2009; Rousselet et al. 2003). Given this mode of presentation, many (about half) of the face images also contained non-face body parts. Previous studies have found that the FFA, especially when defined at the resolution used in the present study, responds to body parts (e.g., Peelen and Downing 2005; Pinsk et al. 2009; Schwarzlose et al. 2005, 2008; Weiner and Grill-Spector 2010). Moreover, studies that have performed whole brain random effects contrasts of bodies vs. objects consistently find a body-selective peak in the fusiform gyrus (e.g., Hodzic et al. 2009). In contrast, activation in response to body parts has been less consistently found in the OFA, with one study in particular showing activation in response to headless bodies in the FFA but not in the OFA (Peelen and Downing 2005). Therefore, even though more recent surface-based analysis studies have disclosed larger responses to body parts than to objects also in the OFA (Pinsk et al. 2009; Weiner and Grill-Spector 2010, with the latter reporting relatively more activation in response to body parts than to cars in the OFA than in the FFA), we should acknowledge the possibility that the presence of body parts could have contributed to FFA more than to OFA activity and as such constitutes a confound in the study. Moreover, we note that our whole brain analysis disclosed preferential activation to visual scenes containing faces in a left occipital area, as early as in the right FFA. Since this area did not show face sensitivity in the localizer and corresponds to the coordinate of the EBA, it seems that detection of body parts may have occurred as early as face detection in our paradigm, also in different areas and hemispheres.

Second, successful recognition took place on visual stimuli whose structure (phase) was still partially disrupted (i.e., recognition occurred before the stimulation sequence was completed). This phase-scrambling is an artificial manipulation that adds noise to the visual scene, masking diagnostic information that may be useful for categorization. Thus one may argue that our results are due to the right FFA being simply more resistant to a high noise level than the right OFA (although both areas can be activated for faces with high levels of noise; e.g., Righart et al. 2010) and that if the different frames of the sequence were presented one by one, in a random order, the FFA would be activated at frames containing a higher level of noise than the OFA. Since the stimuli are presented in an increasing order of visibility in our paradigm, an earlier response to faces emerges in the right FFA than the right OFA. It is indeed a plausible account of our observations. However, we believe this account to be in line with our claims and reflective of the phenomenon that we attempted to measure: if we objectively disrupt structured visual information (phase-scrambling) and present the stimulus according to an increasing order of visibility, face sensitivity emerges at a higher level of scrambling, that is, earlier, in the right FFA than in the right OFA. In this context, it is worth noting that in real life, rather than revealing objects that are unambiguous and unique, scenes are typically noisy and cluttered, due to occlusions and lighting variation (shadows, luminance edges, and gradients). Thus, apart from the particularly slow increase of diagnostic information in the paradigm, the kind of stimulation used in this

study is not necessarily too far from the real-world conditions of object categorization in complex visual scenes. In fact, it could be argued that this kind of stimulation resembles situations in which faces and objects have to be detected in visual scenes under conditions of low visual acuity and/or contrast sensitivity, occlusion, reduced visibility, or initial perception of shapes in the periphery. Finally, and importantly, we note that phase-scrambling is an objective manipulation of information: it disrupts the structure across the whole stimulus, affecting all frequency scales rather than specific sources of information on the faces (e.g., global organization of the face by moving the parts around, or the parts themselves by blurring; e.g., Goffaux et al. 2011; Lobmaier et al. 2008).

Third, rather than being briefly presented, as in most neuroimaging studies of face and object categorization, visual stimulus information was slowly revealed to the observers at a fixed rate, in a dynamic sequence. This slow presentation rate, which has been used successfully in previous fMRI studies of object/face categorization (e.g., Carlson et al. 2006; Eger et al. 2007; Estermann and Yantis 2010; James et al. 2000; Reinders et al. 2005, 2006), allows us to titrate the contribution of different brain areas over time and establish a temporal sequence for neural events. Critically, given the poor temporal resolution fMRI, our method allowed us to separate events that normally occur within tens of milliseconds at most (Formisano and Goebel 2003). Note that under day-to-day conditions, visual observers survey scenes continuously rather than having pictures appear and disappear before us, so this kind of dynamic stimulation is again perhaps more ecological than one might first think.

Fourth and finally, given the slow mode of presentation and the complex visual scenes that were presented, another important issue to consider is the potential role of top-down factors, or perceptual expectations, to our observations. Several recent fMRI studies have showed a contribution of perceptual expectations in face-sensitive activation in the fusiform gyrus in particular (Egner et al. 2010; Puri et al. 2009; Summerfield et al. 2006; Zhang et al. 2008; also in the inferior occipital cortex: Esterman and Yantis 2010; Righart et al. 2010). Therefore, one should consider the possibility that the earlier face-preferential response disclosed in the FFA than in the OFA in the present study is due, even partly, to such top-down factors. Importantly, we should state that our study differed in several ways from the studies cited above in which top-down factors were manipulated explicitly. First, we did not bias participants in our study toward face or car trials by cueing or presenting higher probabilities of appearance for one category compared with the other. Second, behavioral responses were in agreement with bottom-up information in the very large proportion of trials (i.e., no misperception), and if anything, participants had more “car” than “face” responses. They were also slightly faster to detect the cars than the faces in the scenes, which is not a pattern of result that one would expect to find if anticipation of a face would be responsible for the “early” FFA differential response (with respect to OFA). Third, contrary to these recent fMRI studies, we used complex visual scenes rather than well-segmented face images so that the overall structure of the face (and car) stimuli was not predictable, reducing the likelihood that a predictive code was used to categorize the stimuli (see Summerfield et al. 2006). Nevertheless, we should acknowledge that detecting whether there is a person or a car in

a slowly revealed scene does not simply rely on the bottom-up detection of visual cues of a face or a car stimulus. It relies on many more cues, including contextual cues (a road, a body part) and nonspecific form information (e.g., large vs. small blob), as well as more global scene information (see e.g., Wolfe et al. 2011). These multiple cues might be related more to one of the object category to detect, such as the road and body part examples, and thus could be used by the observer to come up with an initial guess (i.e., a template) about the category that is present in the scene. Given the slow mode of presentation, this template (e.g., “this oval blob there is probably a face”) can then be tested against the slowly revealed image, with observers “trying to see” a face in the blob to confirm their inclination. Once enough evidence is accumulated and a face is clearly perceived, the observer can provide a behavioral response. Admittedly, the initial differential activity observed earlier in the FFA than the OFA might relate to a greater contribution of this top-down search template in the FFA than in the OFA, not only to the face stimulus itself activating the FFA before the OFA. This account is compatible with FFA activity starting before the behavioral response, which was also observed in the data (Figs. 4 and 5).

Considering these particularities of our paradigm and the limitations of fMRI in terms of temporal resolution in general, we acknowledge that our study does not provide any information about the absolute time course of face categorization in the human brain. Moreover, we cannot exclude the possibility that if an isolated face is flashed up without any prior expectation, the right OFA may respond before the right FFA. Therefore, the current paradigm provides one piece of evidence about the respective time course of face perception in the two areas of interest, but we cannot exclude that in typical experimental situations, where responses are time-locked to the onset of a face image (vs. an object image), the right OFA responds before the right FFA. This question cannot be answered directly in such typical paradigms in fMRI and would require measuring the response latencies of neurons in these areas of interest as predefined in neuroimaging, as well as testing the impact of selective inactivation of one of these areas on the other’s face-related activation at the neuronal level. Such experiments would currently be possible only by combining fMRI and neurophysiology in the nonhuman primate brain (Tsao et al. 2006, 2008).

Moreover, a direct comparison between areas in fMRI studies alone assumes that the hemodynamic responses of different cortical areas are nearly identical, an assumption that is almost certainly invalid given the complexities of the cerebral vasculature (Carlson et al. 2006). This is the reason why our approach focused on the relative emergence of face-preferential responses: the onset of a significant and lasting difference in hemodynamic response between two experimental conditions. Nevertheless, despite the idiosyncratic nature of our paradigm as well as its acknowledged limitations, we believe that the present observations are quite useful for informing us regarding the relative time course of face-related activation in different areas of interest. As such, we hold that our results place new constraints on models of the functional neuroanatomy of face perception in the human brain.

*Primacy and temporal precedence of holistic face perception in the human brain.* Why would a neuroanatomically higher level visual area of the cortical face network, the right

FFA, show the earliest sensitivity to faces, and what are the implications of this observation for our understanding of the spatiotemporal course of face perception in the human brain? As indicated in the Introduction, both the FFA and OFA are located outside of visual areas whose borders can be defined thanks to their precise retinotopic organization (Halgren et al. 1999; Weiner and Grill-Spector 2010). However, neurons in higher level visual areas of both the dorsal and ventral stream still present with some degree of retinotopy (Levy et al. 2001; Wandell et al. 2007), which certainly applies to the OFA and even the FFA when optimal stimuli (i.e., faces) are used (Hemond et al. 2007; Sayres et al. 2009; Yue et al. 2011). Since face-selective neurons recorded in the inferotemporal cortex of the monkey brain have receptive fields of 20–50° (area TE; Boussaoud et al. 1991; Tsunoda et al. 2001), it is reasonable to assume that neurons in the human FFA, a rather anterior area in the ventral visual pathway, have a quite large receptive field, certainly encompassing whole faces of various sizes and locations. fMRI-adaptation studies support this view, showing a generalization (i.e., lack of release from adaptation) to substantial changes of face position and size in the FFA (also referred to as pFs; see Grill-Spector et al. 1999; Grill-Spector and Malach, 2001; but see Yue et al. 2011). Also, as mentioned above, the FFA responds maximally when a sufficiently large portion of the visual stimulus is presented so that it can be categorized unambiguously as a face (Lerner et al. 2001). In contrast, the OFA is located about 2 cm posteriorly, much closer to retinotopic visual areas than the FFA, suggesting that populations of neurons in the OFA have a smaller receptive field than those in the FFA, being less sensitive to image scrambling or fragmentation. Indeed, the bias for central vs. peripheral stimulation of the visual field that is found in face-sensitive cortex is much stronger in the OFA than in the FFA (Levy et al. 2001).

Thus it is reasonable to consider that neurons in the FFA should be able to code for a more global representation of a face than in the OFA, in which different aspects of the visual scene (and face) must be coded by different populations of neurons. As such, it may be that a generic “face template” is only effective within the FFA and higher order face-sensitive areas (Nestor et al. 2010). Consistent with this view, a visual stimulus that is perceived as a face by means of its global configuration rather than particular local features (e.g., a Mooney or Arcimboldo face) may also activate the right FFA without evidence of face sensitivity in the right OFA (Dolan et al. 1997; Rossion et al. 2011). This latter finding suggests that the right FFA plays an important role in face categorization/detection at a global/holistic level. In addition, in the monkey brain, face-selective cells identified in the monkey inferotemporal cortex are sensitive to the whole facial organization: removal of a part of the face (Tsunoda et al. 2001) or changing the first-order configuration of the face (Desimone et al. 1984) both produce a marked reduction in neuronal response strength, also suggesting holistic face representations.

Given this body of results, our present finding of initial face-preferential responses in the FFA supports a view in which the first face-specific perceptual representation is that of the whole face. In fact, the RISE sequences that we used might have emphasized this precedence of the holistic face representation: the face appears to emerge from the noise as a global configuration first, with specific features becoming more clear

later (Fig. 1). However, contrary to other kinds of manipulations (e.g., blurring, masking of features, breaking of the face configuration by displacement of features, etc.; e.g., Lewis and Edmonds 2002; Lobmaier et al. 2008; Schwanager et al. 2002), the respective contribution of the whole face structure and of specific features is not explicitly manipulated in this paradigm. Rather, phase information across the entire visual scene is fully randomized and gets reorganized progressively over time. Hence, the initial activation of the FFA that seems to be associated with the precedence of holistic representations emerges naturally in this paradigm.

This view of a (temporal) precedence of a face representation that takes into account all facial features interdependently (a holistic representation; see Sergent 1984; Tanaka and Farah 1993) is compatible with a Gestalt view of the microgenesis of face perception, according to which an initial coarse and holistic initial representation is refined over time (Flavell and Draguns 1957; Sergent 1986; Watt 1987). This view is supported by a number of behavioral, electrophysiological, and neuroimaging observations. For instance, faces can be perceived in very low-resolution pictures, in which distinct features cannot even be extracted (Harmon 1973; Sergent 1986). Moreover, as mentioned above, a visual stimulus can be interpreted as a face based solely on its global configuration, rather than on particular local features (e.g., Mooney or Arcimboldo face stimuli; Moore and Cavanagh 1998).

Electrophysiological studies also support the temporal precedence of holistic representations. For instance, early face-sensitive responses (N170) recorded on the human scalp are delayed if facial features are isolated, if the face is scrambled in visible parts, or if it is cut in two horizontal halves, effects that are most prominent in the right hemisphere (e.g., Bentin et al. 1996; Letourneau and Mitchell 2008). Single-cell recordings in the monkey inferotemporal cortex also indicate that global coarse representations are available before fine-grained details (Sripati and Olson 2009; Sugase et al. 1999).

Finally, a recent fMRI study also supports this view, showing that the right FFA and, to a lesser extent, the right pSTS respond preferentially to low spatial frequency face information in early stages of face processing (i.e., until 75 ms of exposure duration) compared with higher spatial frequency information (Goffaux et al. 2011). Moreover, in that study, the response to finer grained face information, i.e., high spatial frequency, became more significant over time in the bilateral FFAs and in the right OFA, providing further support for the view advocated presently.

In summary, the onset of preferential responses to faces in the right FFA as found in the present study is in agreement with a rather well-supported view of the microgenesis of face perception according to which the initial detection of a face as a face is achieved by considering the whole facial configuration, rather than by treating the features as spatially independent entities. In agreement with this “reverse” hierarchical view of visual perception (e.g., Bullier et al. 2001; Galuske et al. 2002; Hochstein and Ahissar 2002; Hupé et al. 1998; Lamme and Roelfsema 2000; Mumford 1992; see also Bar 2003), lower order visual areas exhibiting later face sensitivity, such as the OFA, may be involved in refining the initial coarse representation that arises in higher order visual areas, for the purpose of finer grained categorization such as face individualization (Rossion 2008; Schiltz et al. 2006).

**Conclusions.** The results of our study obtained during the gradual revealing of visual scenes suggest a nonhierarchical (with respect to neuroanatomy) emergence of face sensitivity among cortical regions. More specifically, the face-preferential response in the right occipital cortex (right OFA) may follow the face-preferential response in the fusiform gyrus (right FFA) (Rossion 2008). Hence, the early face sensitivity in the right FFA may arise independently of any face-sensitive inputs from the inferior occipital cortex (OFA). This initial face categorization within the right FFA may reflect the early detection of faces qua faces, and the later face sensitivity observed in the OFA may emerge as a result of reentrant connections between the FFA and the OFA. Functionally, we speculate that such reentrant connections may facilitate further processing of facial stimuli with the goal of refining exemplar-specific features to better support facial individuation.

#### GRANTS

This research was supported by Communauté Française de Belgique-Actions de Recherche Concertées Grant 07/12-007 and the Fonds National de la Recherche Scientifique (FNRS) Mandat d'Impulsion Scientifique to B. Rossion (2008–2011). F. Jiang is supported by a Human Frontiers Science Program postdoctoral award. B. Rossion is supported by the FNRS. G. Righi and M. Tarr were supported by grants from the James S. McDonnell Foundation to the Perceptual Expertise Network and by National Science Foundation Science of Learning Center Grant SBE-0542013 to the Temporal Dynamics of Learning Center.

#### REFERENCES

- Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM. Cue-invariant networks for figure and background processing in human visual cortex. *J Neurosci* 26: 11695–11708, 2006.
- Bar M. A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci* 15: 600–609, 2003.
- Bentin S, Allison T, Puce A, Perez E, McCarthy G. Electrophysiological studies of face perception in humans. *J Cogn Neurosci* 8: 551–565, 1996.
- Biederman I. Recognition-by-components: a theory of human image understanding. *Psychol Rev* 94: 115–147, 1987.
- Boussaoud DR, Desimone R, Ungerleider LG. Visual topography of area TEO in the macaque. *J Comp Neurol* 306: 554–575, 1991.
- Bullier J, Hupe JM, James AC, Girard P. The role of feedback connections in shaping the responses of visual cortical neurons. *Prog Brain Res* 134: 193–204, 2001.
- Burton AM. Learning new faces in an interactive activation and competition model. *Vis Cogn* 1: 313–348, 1994.
- Carlson T, Grol MJ, Verstraten FA. Dynamics of visual recognition revealed by fMRI. *Neuroimage* 32: 892–905, 2006.
- Crouzet SM, Kirchner H, Thorpe SJ. Fast saccades toward faces: face detection in just 100 ms. *J Vis* 10: 1–17, 2010.
- Deichmann R, Gottfried JA, Hutton C, Turner R. Optimized EPI for fMRI studies of the orbitofrontal cortex. *Neuroimage* 19: 430–441, 2003.
- Desimone R, Albright TD, Gross CG, Bruce C. Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci* 4: 2051–2062, 1984.
- Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RS, Friston KJ. How the brain learns to see objects and faces in an impoverished context. *Nature* 389: 596–599, 1997.
- Downing PE, Jiang Y, Shuman M, Kanwisher N. A cortical area selective for visual processing of the human body. *Science* 293: 2470–2473, 2001.
- Eger E, Henson RN, Driver J, Dolan RJ. Mechanisms of top-down facilitation in perception of visual objects studied by fMRI. *Cereb Cortex* 17: 2123–2133, 2007.
- Egner T, Monti JM, Summerfield C. Expectation and surprise determine neural population responses in the ventral visual stream. *J Neurosci* 30: 16601–16608, 2010.
- Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 392: 598–601, 1998.
- Esterman M, Yantis S. perceptual expectations evokes category-selective cortical activity. *Cereb Cortex* 20: 1245–1253, 2010.

- Fairhall SL, Ishai A. Effective connectivity within the distributed cortical network for face perception. *Cereb Cortex* 17: 2400–2406, 2007.
- Fei-Fei L, Iyer A, Koch C, Perona P. What do we perceive in a glance of a real-world scene? *J Vis* 7: 10, 2007.
- Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1: 1–47, 1991.
- Flavell JH, Draguns J. A microgenetic approach to perception and thought. *Psychol Bull* 54: 197–217, 1957.
- Formisano E, Goebel R. Tracking cognitive processes with functional MRI mental chronometry. *Curr Opin Neurobiol* 13: 174–181, 2003.
- Fox CJ, Iaria G, Barton J. Defining the face-processing network: optimization of the functional localizer in fMRI. *Hum Brain Mapp* 30: 1637–1651, 2009.
- Friston KJ, Harrison L, Penny W. Dynamic causal modelling. *Neuroimage* 19: 1273–1302, 2003.
- Galuske RA, Schmidt KE, Goebel R, Lomber SG, Payne BR. The role of feedback in shaping neural representations in cat visual cortex. *Proc Natl Acad Sci USA* 99: 17083–17088, 2002.
- Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. The fusiform “face area” is part of a network that processes faces at the individual level. *J Cogn Neurosci* 12: 495–504, 2000.
- Gobbini I, Haxby JV. Neural systems for recognition of familiar faces. *Neuropsychologia* 45: 32–41, 2007.
- Goffaux V, Peters J, Haubrechts J, Schiltz C, Jansma B, Goebel R. From coarse to fine? Spatial and Temporal dynamics of cortical face processing. *Cereb Cortex* 21: 467–476, 2011.
- Grill-Spector K, Knouf N, Kanwisher N. The fusiform face area subserves face perception, not generic within-category identification. *Nat Neurosci* 7: 555–562, 2004.
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24: 187–203, 1999.
- Grill-Spector K, Malach R. fMR-adaptation: a tool for studying the functional properties of human cortical neurons. *Acta Psychol (Amst)* 107: 293–321, 2001.
- Halgren E, Dale AM, Sereno MI, Tootell RB, Marinkovic K, Rosen BR. Location of human face-selective cortex with respect to retinotopic areas. *Hum Brain Mapp* 7: 29–37, 1999.
- Halgren E, Raij T, Marinkovic K, Jousmäki V, Hari R. Cognitive response profile of the human fusiform face area as determined by MEG. *Cereb Cortex* 10: 69–81, 2000.
- Harmon LD. The recognition of faces. *Sci Am* 229: 70–83, 1973.
- Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. *Trends Cogn Sci* 4: 223–233, 2000.
- Hécaen H, Angèlergues R. Agnosia for faces (prosopagnosia). *Arch Neurol* 7: 92–100, 1962.
- Hemond CC, Kanwisher NG, Op de Beeck HP. A preference for contralateral stimuli in human object- and face-selective cortex. *PLoS One* 2: e574, 2007.
- Hillger LA, Koenig O. Separable mechanisms in face processing: evidence from hemispheric specialization. *J Cogn Neurosci* 3: 42–58, 1991.
- Hochstein S, Ahissar M. View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* 36: 791–804, 2002.
- Hodzic A, Kaas A, Muckli L, Stirn A, Singer W. Distinct cortical networks for the detection and identification of human body. *Neuroimage* 45: 1264–1271, 2009.
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J Physiol* 160: 106–154, 1962.
- Hupé JM, James AC, Payne BR, Lomber SG, Girard P, Bullier J. Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394: 784–787, 1998.
- Ishai A. Let’s face it: it’s a cortical network. *Neuroimage* 40: 415–419, 2008.
- Jacques C, d’Arripe O, Rossion B. The time course of the face inversion effect during individual face discrimination. *J Vis* 7: 1–9, 2007.
- Jacques C, Rossion B. The initial representation of individual faces in the right occipito-temporal cortex is holistic: electrophysiological evidence from the composite face illusion. *J Vis* 9: 1–16, 2009.
- James TW, Humphrey GK, Gati JS, Menon RS, Goodale MA. The effect of visual object priming on brain activation before and after recognition. *Curr Biol* 10: 1017–1024, 2000.
- Jiang F, Dricot L, Blanz V, Goebel R, Rossion B. Neural correlates of shape and surface reflectance information in individual faces. *Neuroscience* 163: 1078–1091, 2009.
- Jiang X, Rosen E, Zeffiro T, Vanmeter J, Blanz V, Riesenhuber M. Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron* 50: 159–172, 2006.
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17: 4302–4311, 1997.
- Kleinschmidt A, Buchel C, Hutton C, Friston KJ, Frackowiak RS. The neural structures expressing perceptual hysteresis in visual letter recognition. *Neuron* 34: 659–666, 2002.
- Kriegeskorte N, Formisano E, Sorger B, Goebel R. Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc Natl Acad Sci USA* 104: 20600–20605, 2007.
- Lamme VAF, Roelfsema PR. The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci* 23: 571–579, 2000.
- Le Grand R, Mondloch CJ, Maurer D, Brent HP. Expert face processing requires visual input to the right hemisphere during infancy. *Nat Neurosci* 6: 1108–1112, 2003.
- Lerner Y, Hendler T, Ben-Nashat D, Harel M, Malach R. A hierarchical axis of object processing stages in the human visual cortex. *Cereb Cortex* 11: 287–297, 2001.
- Letourneau SM, Mitchell TV. Behavioral and ERP measures of holistic face processing in a composite task. *Brain Cogn* 67: 234–245, 2008.
- Leventhal AG, Wang YC, Schmolesky MT, Zhou Y. Neural correlates of boundary perception. *Vis Neurosci* 15: 1107–1118, 1998.
- Levy I, Hasson U, Avidan G, Hendler T, Malach R. Center-periphery organization of human object areas. *Nat Neurosci* 4: 533–539, 2001.
- Lewis MB, Edmonds AJ. Face detection: mapping human performance. *Perception* 32: 903–920, 2003.
- Lobmaier JS, Klaver P, Loenneker T, Martin E, Mast FW. Featural and configural face processing strategies: evidence from a functional magnetic resonance imaging study. *Neuroreport* 19: 287–291, 2008.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92: 8135–8139, 1995.
- Marr D. *Vision*. San Francisco, CA: Freeman, 1982.
- McKeeff TJ, Tong F. The timing of perceptual decisions for ambiguous face stimuli in the human ventral visual cortex. *Cereb Cortex* 17: 669–678, 2007.
- Menon RS, Kim SG. Spatial and temporal limits in cognitive neuroimaging with fMRI. *Trends Cogn Sci* 3: 207–216, 1999.
- Moeller S, Freiwald WA, Tsao DY. Patches with links: a unified system for processing faces in the macaque temporal lobe. *Science* 320: 1355–1359, 2008.
- Mooney CM. Age in the development of closure ability in children. *Can J Psychol* 11: 219–226, 1957.
- Moore C, Cavanagh P. Recovery of 3D volume from 2-tone images of novel objects. *Cognition* 67: 45–71, 1998.
- Mumford D. On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern* 66: 241–251, 1992.
- Nestor A, Leeds DD, Vettel JM, Tarr MJ. Neurally-derived representations for face detection. Poster presented at the *Fifth International Workshop on Statistical Analysis of Neuronal Data*, Pittsburgh, PA, 2010.
- Nestor A, Vettel JM, Tarr MJ. Task-specific codes for face recognition: how they shape the neural representation of features for detection and individuation. *PLoS One* 3: e3978, 2008.
- Parkin AJ, Williamson P. Cerebral lateralization at different stages of facial processing. *Cortex* 23: 99–110, 1987.
- Peelen MV, Downing PE. Selectivity for the human body in the fusiform gyrus. *J Neurophysiol* 93: 603–608, 2005.
- Peelen MV, Fei-Fei L, Kastner S. Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460: 94–97, 2009.
- Peelen MV, Wiggert AJ, Downing PE. Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron* 49: 815–822, 2006.
- Philiastides MG, Sajda P. EEG-informed fMRI reveals spatiotemporal characteristics of perceptual decision making. *J Neurosci* 27: 13082–13091, 2007.
- Pinsk MA, Arcaro M, Weiner KS, Kalkus JF, Inati SJ, Gross CG, Kastner S. Neural representations of faces and body parts in macaque and human cortex: a comparative fMRI study. *J Neurophysiol* 101: 2581–2600, 2009.
- Pitcher D, Duchaine B, Walsh V, Kanwisher K. TMS evidence for feedforward and feedback mechanisms of face and body perception. Poster pre-

- sented at the *10th Annual Meeting of Vision Science Society*, Naples, FL, 2010.
- Pitcher D, Garrido L, Walsh V, Duchaine BC.** Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. *J Neurosci* 28: 8929–8933, 2008.
- Pitcher D, Walsh V, Yovel G, Duchaine B.** TMS evidence for the involvement of the right occipital face area in early face processing. *Curr Biol* 17: 1568–1573, 2007.
- Puce A, Allison T, Bentin S, Gore JC, McCarthy G.** Temporal cortex activation in humans viewing eye and mouth movements. *J Neurosci* 18: 2188–2199, 1998.
- Puce A, Allison T, Gore JC, McCarthy G.** Face-sensitive regions in human extrastriate cortex by functional MRI. *J Neurophysiol* 74: 1192–1199, 1995.
- Puri AM, Wojciulik E, Ranganath C.** Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Res* 1301: 89–99, 2009.
- Rajimehr R, Young JC, Tootell RBH.** An anterior temporal face patch in human cortex, predicted by macaque maps. *Proc Natl Acad Sci USA* 106: 1995–2000, 2009.
- Reinders AA, Cläscher J, de Jong JR, Willemsen AT, den Boer JA, Büchel C.** Detecting fearful and neutral faces: BOLD latency differences in amygdala-hippocampal junction. *Neuroimage* 33: 805–814, 2006.
- Reinders AA, den Boer JA, Büchel C.** The robustness of perception. *Eur J Neurosci* 22: 524–530, 2005.
- Righart R, Andersson F, Schwartz S, Mayer E, Vuilleumier P.** Top-down activation of fusiform cortex without seeing faces in prosopagnosia. *Cereb Cortex* 20: 1878–1890, 2010.
- Riesenhuber M, Poggio T.** Hierarchical models of object recognition in cortex. *Nat Neurosci* 2: 1019–1025, 1999.
- Rolls ET, Tovee MJ.** Sparseness of the neuronal representation of stimuli in the primate temporal visual-cortex. *J Neurophysiol* 73: 713–726, 1995.
- Rossion B.** Constraining the cortical face network by neuroimaging studies of acquired prosopagnosia. *Neuroimage* 40: 423–426, 2008.
- Rossion B, Caharel S.** ERP evidence for the speed of face categorization in the human brain: disentangling the contribution of low-level visual cues from face perception. *Vision Res* 51: 1297–1311, 2011.
- Rossion B, Caldara R, Seghier M, Schuller AM, Lazeyras F, Mayer E.** A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126: 2381–2395, 2003.
- Rossion B, Dricot L, Goebel R, Busigny T.** Holistic face categorization in higher order visual areas of the normal and prosopagnosic brain: toward a non-hierarchical view of face perception. *Front Hum Neurosci* 4: 225, 2011.
- Rossion B, Jacques C.** Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage* 39: 1959–1979, 2008.
- Rossion B, Jacques C.** The N170: understanding the time-course of face perception in the human brain. In: *The Oxford Handbook of ERP Components*, edited by Kappenman ES, Luck SJ. New York: Oxford University Press, 2011.
- Rousselet GA, Mace MJ, Fabre-Thorpe M.** Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J Vis* 3: 440–455, 2003.
- Sadeh B, Podlipsky I, Zhdanov A, Yovel G.** Event-related potential and functional MRI measures of face-selectivity are highly correlated: a simultaneous ERP-fMRI investigation. *Hum Brain Mapp*, 2010.
- Sadr J, Sinha P.** Object recognition and random image structure evolution. *Cogn Sci* 28: 259–287, 2004.
- Sayres R, Wiener K, Dumoulin S, Wandell B, Grill-Spector K.** Population receptive field measurements in human ventral category-selective cortex. *J Vis* 9: 734, 2009.
- Schiltz C, Rossion B.** Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage* 32: 1385–1394, 2006.
- Schiltz C, Sorger B, Caldara R, Ahmed F, Mayer E, Goebel R, Rossion B.** Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cereb Cortex* 16: 574–586, 2006.
- Schwanninger A, Lobmaier J, Collishaw SM.** Component and configural information in face recognition. *Comput Sci* 2525: 643–650, 2002.
- Schwarzlose RF, Baker CI, Kanwisher N.** Separate face and body selectivity on the fusiform gyrus. *J Neurosci* 25: 11055–11059, 2005.
- Schwarzlose RF, Swisher JD, Dang S, Kanwisher N.** The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci USA* 105: 4447–4452, 2008.
- Sergent J.** An investigation into component and configural processes underlying face perception. *Br J Psychol* 17: 43–63, 1984.
- Sergent J.** Microgenesis of face perception. In: *Aspects of Face Processing*, edited by Ellis H, Jeeves MA, Newcombe F, Young AW. Dordrecht, The Netherlands: Nijhoff, 1986.
- Sergent J.** Face perception and the right hemisphere. In: *Thought Without Language*, edited by Weiskrantz L. Oxford: Oxford University Press, 1988, p. 108–131.
- Sergent J, Ohta S, MacDonald B.** Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* 115: 15–36, 1992.
- Sripati AP, Olson CR.** Representing the forest before the trees: a global advantage effect in monkey inferotemporal cortex. *J Neurosci* 29: 7788–7796, 2009.
- Steeves JK, Culham JC, Duchaine BC, Pratesi CC, Valyear KF, Schindler I, Humphrey GK, Milner AD, Goodale MA.** 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, 2006.
- Sugase Y, Yamane S, Ueno S, Kawano K.** Global and fine information coded by single neurons in the temporal visual cortex. *Nature* 400: 869–873, 1999.
- Summerfield C, Eger T, Mangels J, Hirsch J.** Mistaking a house for a face: neural correlates of misperception in healthy humans. *Cereb Cortex* 16: 500–508, 2006.
- Talairach G, Tournoux P.** *Co-Planar Stereotaxic Atlas of the Human Brain*. New York: Thieme, 1988.
- Tanaka JW, Farah MJ.** Parts and wholes in face recognition. *Q J Exp Psychol A* 46: 225–245, 1993.
- Tsao DY, Freiwald WA, Tootell RB, Livingstone MS.** A cortical region consisting entirely of face-selective cells. *Science* 311: 670–674, 2006.
- Tsao DY, Moeller S, Freiwald WA.** Comparing face patch systems in macaques and humans. *Proc Natl Acad Sci USA* 105: 19514–19519, 2008.
- Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M.** Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nat Neurosci* 4: 832–838, 2001.
- Ullman S.** Object recognition and segmentation by a fragment-based hierarchy. *Trends Cogn Sci* 11: 58–64, 2007.
- Van Essen DC, Maunsell JH.** Hierarchical organization and functional streams in the visual cortex. *Trends Neurosci* 6: 370–375, 1983.
- VanRullen R.** On second glance: still no high-level pop-out effect for faces. *Vision Res* 46: 3017–3027, 2006.
- VanRullen R, Thorpe SJ.** Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception* 30: 655–668, 2001.
- Wandell BA, Dumoulin SO, Brewer AA.** Visual field maps in human cortex. *Neuron* 56: 366–383, 2007.
- Watt RJ.** Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *J Opt Soc Am A* 4: 2006–2021, 1987.
- Weiner KS, Grill-Spector K.** Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage* 52: 1559–1573, 2010.
- Wiggett AJ, Downing PE.** The face network: overextended? *Neuroimage* 40: 420–422, 2008.
- Wolfe JM, Võ ML, Evans KK, Greene MR.** Visual search in scenes involves selective and nonselective pathways. *Trends Cogn Sci* 15: 77–84, 2011.
- Yin RK.** Looking at upside-down faces. *J Exp Psychol* 81: 141–145, 1969.
- Yue X, Cassidy BS, Devaney KJ, Holt DJ, Tootell RB.** Lower-level stimulus features strongly influence responses in the fusiform face area. *Cereb Cortex* 21: 35–47, 2011.
- Zangenehpour S, Chaudhuri A.** Patchy organization and asymmetric distribution of the neural correlates of face processing in monkey inferotemporal cortex. *Curr Biol* 15: 993–1005, 2005.
- Zhan CA, Baker CL Jr.** Boundary cue invariance in cortical orientation maps. *Cereb Cortex* 16: 896–906, 2006.
- Zhang H, Liu J, Huber DE, Rieth CA, Tian J, Lee K.** Detecting faces in pure noise images: a functional MRI study on top-down perception. *Neuroreport* 19: 229–233, 2008.