

Being BOLD: The Neural Dynamics of Face Perception

Francesco Gentile,^{1,2,3*} Justin Ales,⁴ and Bruno Rossion^{2,3,5}

¹Netherlands Institute for Neuroscience (NIN), Amsterdam,
The Netherlands

²Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience,
Maastricht University, Maastricht, The Netherlands

³Maastricht Brain Imaging Center (M-BIC), Maastricht University, Maastricht,
The Netherlands

⁴School of Psychology and Neuroscience, University of St. Andrews, St. Andrews,
United Kingdom

⁵Psychological Sciences Research Institute and Institute of Neuroscience (IPSY),
University of Louvain, Louvain, Belgium

Abstract: According to a non-hierarchical view of human cortical face processing, selective responses to faces may emerge in a higher-order area of the hierarchy, in the lateral part of the middle fusiform gyrus (fusiform face area [FFA]) independently from face-selective responses in the lateral inferior occipital gyrus (occipital face area [OFA]), a lower order area. Here we provide a stringent test of this hypothesis by gradually revealing segmented face stimuli throughout strict linear descrambling of phase information [Ales et al., 2012]. Using a short sampling rate (500 ms) of fMRI acquisition and single subject statistical analysis, we show a face-selective responses emerging earlier, that is, at a lower level of structural (i.e., phase) information, in the FFA compared with the OFA. In both regions, a face detection response emerging at a lower level of structural information for upright than inverted faces, both in the FFA and OFA, in line with behavioral responses and with previous findings of delayed responses to inverted faces with direct recordings of neural activity were also reported. Overall, these results support the non-hierarchical view of human cortical face processing and open new perspectives for time-resolved analysis at the single subject level of fMRI data obtained during continuously evolving visual stimulation. *Hum Brain Mapp* 00:000–000, 2016. © 2016 Wiley Periodicals, Inc.

Key words: face perception; neural dynamics; non-hierarchical model; fMRI; OFA; FFA; face inversion

Additional Supporting Information may be found in the online version of this article.

The authors declare no competing financial interests.

Contract grant sponsor: ERC grant (faceSSVEP 284025) and by the Belgian National Fund for Scientific Research (FNRS).

*Correspondence to: F. Gentile; Netherlands Institute for Neuroscience (NIN), Meibergdreef 47, 1105BA, Amsterdam, The Netherlands.
E-mail: f.gentile@nin.knaw.nl/f.gentile@maastrichtuniversity.nl

Received for publication 29 March 2016; Revised 19 July 2016;
Accepted 3 August 2016.

DOI: 10.1002/hbm.23348

Published online 00 Month 2016 in Wiley Online Library
(wileyonlinelibrary.com).

INTRODUCTION

Face perception is one of the most important functions of the human brain. Lesion studies, intracerebral recordings and stimulation as well as neuroimaging in humans provides converging evidence that this function depends on a widely distributed network of areas and their interconnections in the ventral occipito-temporal cortex, with a right hemispheric dominance [Allison et al., 1999; Fox et al., 2008; Haxby et al., 2000; Jonas et al., 2016; Kanwisher and Yovel, 2006; Puce et al., 1998; Rossion et al., 2012a; Sergent et al., 1992; Thomas et al., 2009; Weiner and Grill-Spector, 2013; Zhen et al., 2015]. Neuroimaging studies of the healthy brain point to the lateral sections of the inferior occipital gyrus (IOG) and of the posterior/middle fusiform gyrus (MFG) as two regions of the ventral occipito-temporal cortex where larger responses to faces than non-face objects are consistently observed. These “face-selective” regions are usually labeled as the “occipital face area” (OFA) [Gauthier et al., 2000] and the “fusiform face area” (FFA) [Kanwisher et al., 1997], respectively. They are thought to play an important role in distinguishing faces from non-face objects but also in extracting information about invariant aspects of faces such as facial identity [e.g., Davies-Thompson et al., 2009; Ewbank et al., 2013; Gauthier et al., 2000; Gentile and Rossion, 2014; Jonas et al., 2014]. Structural [Gomez et al., 2013; Gschwind et al., 2012; Pyles et al., 2013; Tavor et al., 2014] and functional [Fairhall and Ishai, 2007; Moeller et al., 2008; Turk-Browne et al., 2010; Zhu et al., 2011] connectivity findings suggest that the OFA and FFA are interconnected and the strength of functional connectivity between the two regions is correlated with behavioral aspects of face processing [Zhu et al., 2011].

According to the hierarchical view of face perception, FFA activation depends on face-selective inputs coming from the posteriorly located OFA, which is generally considered as the lower order face-selective region of the cortical face network [Fairhall and Ishai, 2007; Haxby et al., 2000; Pitcher et al., 2011]. This model has been challenged by the presence of face-selective responses in the right FFA in patients with long-term damage in the right or bilateral IOG with no OFA activation [Rossion et al., 2003; Sorger et al., 2007; Steeves et al., 2006]. Such findings suggest that face-selectivity in the higher-order FFA can arise without the contribution of the lower-order OFA, potentially using connections from early visual areas [Rossion, 2008a; Rossion et al., 2003]. In this regard the hierarchical model of cortical face processing has been revised to incorporate a direct connection between early visual areas and the FFA which would run in parallel to connections from these early visual areas to the OFA [Atkinson and Adolphs, 2011; Duchaine and Yovel, 2015; Rossion, 2008a].

However, whether, during the course of face perception, high-level (i.e., structural) specific information about faces can reach the FFA even before the OFA remains debated. The relevance of answering to this question arises from the different response properties of the two areas. For instance,

the FFA contains face representations that have larger receptive fields [Hemond et al., 2007] and are more holistically integrated [Axelrod and Yovel, 2010; Harris and Aguirre, 2008, 2010; Rossion and Boremanse, 2011; Schiltz and Rossion, 2006] than representations in the OFA. Thus, identifying which area between the FFA and the OFA initially process a face as a face is important to understand the type of initial process of face perception: holistically or rather part-by-part [Rossion, 2014]. In principle the temporal resolution of functional magnetic resonance imaging (fMRI) is too low to provide information about the relative onset of face-selective activation in the OFA and FFA. However, if information processing takes several seconds, fMRI can track the relative—rather than absolute—latency and time-course of cortical activations across brain regions during perceptual or cognitive tasks [Formisano and Goebel, 2003]. Given that visual processing, and face perception in particular, is extremely fast [Crouzet et al., 2010], one strategy is to slow down, that is, gradually reveal, the visual stimulation in order to potentially disclose relative timing differences between cortical areas of the visual system. This approach has been used by Jiang et al., [2011] with fMRI in the context of face processing. Strikingly, this study showed that slowly (i.e., over several seconds) revealed pictures of faces were discriminated from pictures of cars significantly earlier in the stimulation sequence in the right FFA than in the right OFA [see also Jiang et al., 2015].

However, it is fair to say that these studies have been limited by several factors: (1) the relatively low sampling rate (i.e., TR of 1,250 ms) which makes it more difficult to identify timing differences between functional areas in fMRI; (2) the presence of body parts in a fraction of the stimuli, which might have contributed to FFA more than to OFA activity [e.g., Peelen and Downing, 2005; Pinsk et al., 2009; Schwarzlose et al., 2005, 2008; Weiner and Grill-Spector, 2010]; (3) the comparison of faces to a non-face object category—cars—which was not matched for low-level features; (4) a nonlinear descrambling of phase (i.e., structure) of the stimuli over time [Sadr and Sinha, 2004], which might have created large transient in the stimulation [Ales et al., 2012]; (5) a group analysis that is statistically suboptimal given the large amount of variability not only in terms of localization of the face-selective regions [Rossion et al., 2012a; Zhen et al., 2015], but also in terms of the magnitude of face-selectivity [Frost and Goebel, 2012; Zhen et al., 2015], and the shape of the BOLD response [Aguirre et al., 1998; Handwerker et al., 2004].

The present fMRI study aimed to overcome these limitations and to provide a more stringent test of the non-hierarchical view of the emergence of face-selectivity in the human brain. To do so, we first reduced the sampling rate of the fMRI signal acquisition to half a second (i.e., TR = 500 ms) and in order to achieve a reasonable spatial resolution ($3 \times 3 \times 3 \text{ mm}^3$) we sampled only few slices, focusing on the predefined FFA and OFA. Second, in order to avoid any bias in favor of the FFA activation there were no body parts

or external features in the stimuli used. Third, the non-face objects were only employed for the definition of the region of interest (i.e., FFA and OFA), but not in the actual testing of the onset differences between the functionally defined regions. Instead, we contrasted the BOLD responses across face-selective regions for faces and phase-scrambled faces, which were matched for low-level visual properties (i.e., power spectrum). This is particularly important, considering that low-level statistical properties of images contained in the amplitude spectrum contribute to fast face detection responses [e.g., Crouzet et al., 2010; Honey et al., 2008], and that category-selectivity of ventral visual stream region as the FFA appears to be sensitive to low-level statistical image properties [Coggan et al., 2016; Rice et al., 2014]. Moreover rather than using picture of non-face objects, we also introduced a simple and well-known stimulus transformation—picture-plane inversion. Inverted faces activate both the FFA and OFA robustly, albeit less than upright faces [e.g., Gilaei-Dotan et al., 2010; Haxby et al., 1999; Mazard et al., 2006; Yovel and Kanwisher, 2005]. Most importantly, this manipulation has the advantage of preserving low-level global visual information, while disrupting and slowing down the perception of a face as a face [e.g., Purcell and Stewart, 1988a; Rousselet et al., 2003], in particular when such a stimulus is presented during gradual phase descrambling as employed here [Liu-Shuang et al., 2015]. Thus, including inverted faces in the design provides a direct assessment of relative onset and time-course differences between face-selective cortical areas during gradual revealing of information. A further improvement we made in this version of the experiment was related to the stimuli creation. We used a recently developed algorithm in which the phase angle is linearly interpolated with a direction of interpolation that corresponds to the minimum distance between phases, irrespective of modulus boundaries [Ales et al., 2012]. The use of the minimum distance between phases preserves the uniformity of the phase distribution around the unit circle and provides equal sized steps, so that the face structure (i.e., phase) can truly increase parametrically over time. Finally, we analyzed and displayed the fMRI data at the single subject level. This was possible since our slowly revealing stimulation mode isolates the face detection response from the transient visual response(s), and thus it provides high signal-to-noise responses that can be tested statistically at the single subject level.

In summary, we tested six participants in two separate fMRI sessions. First, face-selective regions of the ventral occipito-temporal cortex were predefined for each subject with a classical face localizer. Within those regions the data from the second session, the main experiment, was analyzed. During the experiment, subjects were required to detect faces progressively emerging by phase-descrambling. We predicted that the progressive emergence of the face would lead to an increase of the fMRI signal significantly earlier in the FFA than the OFA, and that in both regions upright faces would be associated with a relatively earlier response than inverted faces.

METHODS

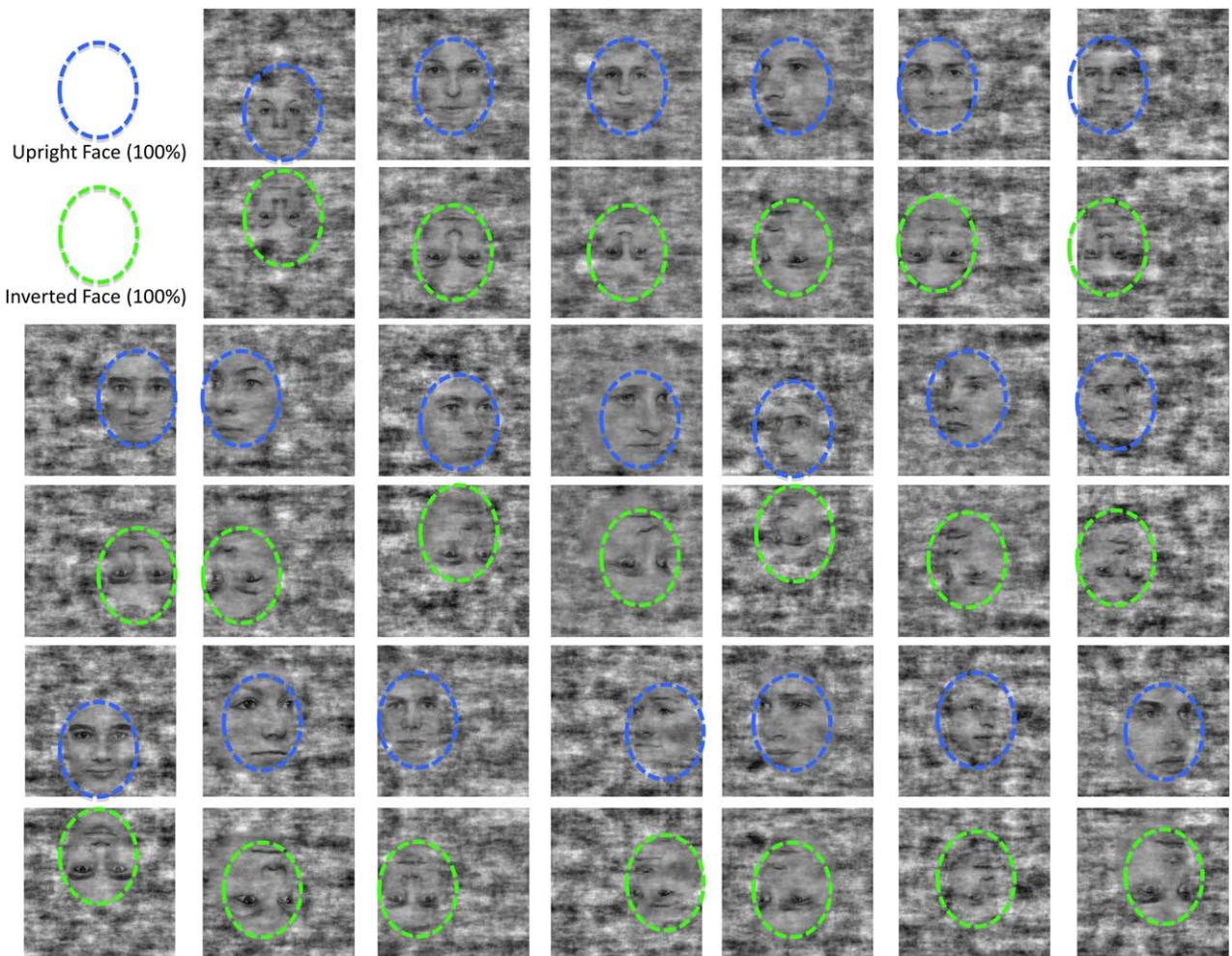
Subjects

Six healthy volunteers (5 females and 1 male, right handed) with normal or corrected visual acuity took part in the study. Four subjects (S1, S2, S3, and S6) performed the two experiments in two different sessions, and the remaining two (S4, S5) in a single session. All participants were undergraduate or PhD students (mean age: 25.33 and SD = 3.72) recruited at Maastricht University. Their participation was compensated with cash money. After explanation of the procedures, participants signed an informed consent form. The ethical committee of the faculty of Psychology in Maastricht approved the study.

Stimuli

The original set of images consisted of 20 high-quality photographic face images, namely 15 from the study of Ales et al., [2012] and 5 new faces, taken from a larger dataset as reported in Laguesse et al. [2012] (see Fig. 1 in that article). In order to minimize any top-down expectations about the location and shape of the stimulus to be detected as a face, after removing external features such as hair, the faces were placed at various spatial locations on a uniform rectangular white background and they were varied in size (three levels) and viewpoint (eight full-front, six left profile, and six right profile). The making of the stimuli was fully described in Ales et al. [2012]. Below we describe the key steps of making the stimuli, with additional information and illustration in Supporting Information.

Face visibility was varied parametrically by creating a graded sequence of 20 images from each stimulus with decreasing degrees of phase-scrambling. Phase-scrambling maintains the same distribution of low-level image statistics (i.e., equal power spectra and mean luminance). The image background remained fully scrambled throughout the entire sequence. There were two distinct processes involved in the creation of the stimuli. First, a set of face exemplars on noise backgrounds with identical power spectra from a set of unscrambled isolated face images (Supporting Information Fig. S1) was created. The average power spectrum was calculated over the set of 20 isolated face exemplars and combined with the phase spectrum of each exemplar to create intermediate images with identical power spectra. The face regions of Figure 1 still contain noise because they are still 100% phase coherent with the face exemplars; the noise in the face regions is a result of balancing the power spectrum across the set of exemplars (the amount of noise added to the face regions as a result of changing the amplitude spectrum is shown in Supporting Information Fig. S1). Thus, the 100% coherent face stimulus is fully phase coherent in the face region, but is not 100% amplitude coherent. In order to limit the inclusion of a local contrast cue that would occur if isolated faces were scrambled, each face was embedded in a random noise background of the same power spectrum as

**Figure 1.**

The face stimuli ($n = 20$) used in the study. They are shown here at 100% phase-coherence in their upright and inverted version. The blue and green circles identify the upright and inverted faces within the “noise” background in which the face stimuli

are embedded in (note that the circles were not present in the stimulus, they are only used here for didactical reasons). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

the faces. A set of background images was, then, created from the average power spectrum image so that each had a uniform random phase distribution, and the isolated faces were blended with the background images. To eliminate this discontinuity between the face region and the background region of the final images, complementary spatial blending masks that smoothly transitioned between regions were created. The blending masks were made such that they started within the face and ended by the face outline. Complementary masks for faces and the backgrounds were used to avoid an increase in contrast in the transition region. The complementary face and background images were then added to create the final equalized power spectrum faces. The second step in the creation of the stimuli was to generate a series of images that had progressively greater

amounts of scrambling of the phase structure of the face image. Phase angle was linearly interpolated, choosing the direction of interpolation that corresponded to the minimum distance between phases, irrespective of modulus boundaries. Using the minimum distance between phases preserves the uniformity of the phase distribution around the unit circle and provides equal sized steps. The 20 steps that were swept for one face exemplar (one trial) are shown in Figure S1 of Supporting Information. For each face we interpolated between a starting image that had 100% randomized phases and the final unscrambled face exemplar. There were 20 equal steps in the interpolation (Fig. 2b). In order to remove temporal correlations in luminance between successive scrambled images, the starting, fully random, image for the interpolation for each step in the sequence was chosen

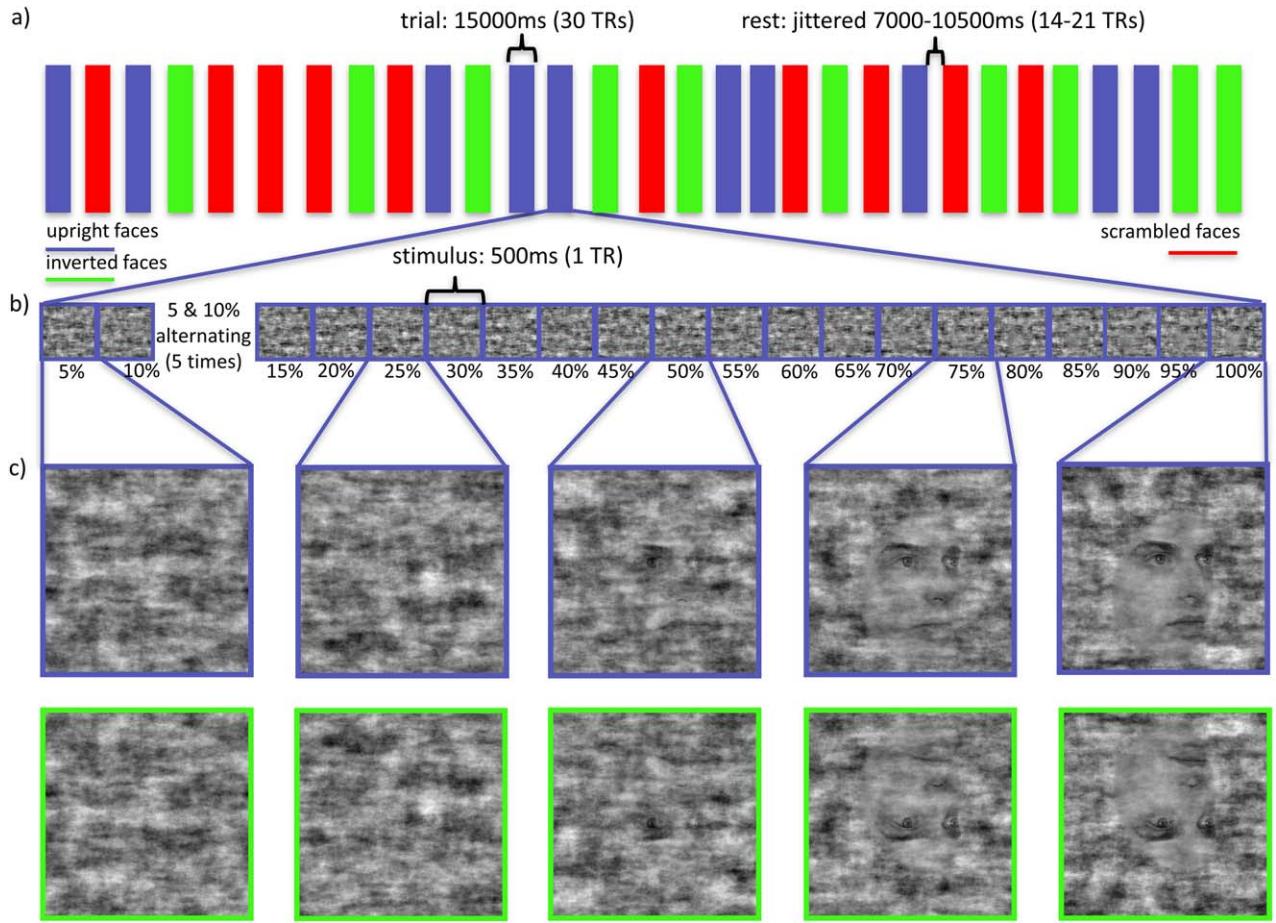


Figure 2.

(a) Experimental design. Each bar represent one condition-block, color-coded according to the specific condition. (b) Stimuli ($n = 30$) composing one condition-block. In the example here the stimuli presented during the upright condition are shown. (c) Five examples

out of the 30 stimuli in a condition-block at five different phase-coherence level, namely at 5%, 25%, 50%, 75%, and 100%. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

independently. The effects of the independent noise images can be seen by noting that on each step the noise background has been updated, and thus the noise masking of the face is different both because a new noise has been used and because the phase-coherence is different. A total of 20 graded face image sequences were created for this study. The least scrambled image of each face exemplar is shown in Figure 1.

Each sequence included 20 steps, ranging from 5% to 100% interpolation of the original and random phase spectra, with 5% change in coherence per step, as opposed to 5.26% used in Ales et al., [2012]. A coherence level of 0 corresponded to a fully randomized phase spectrum of the original image and a coherence level of 100% corresponded to an unaltered phase spectrum. The 20 inverted face stimuli were generated here by flipping the upright face stimuli with respect to the horizontal midline. Twenty fully phase-scrambled sequences of

stimuli, with a random phase values at every “step,” were also created. In total there were 1,200 images used in the study (3 conditions \times 20 faces \times 20 steps).

Procedure

A single run of the main experiment consisted of the presentation of 30 trials, each lasting for 15 seconds (30 volumes of 500 ms), where three different conditions were presented (Fig. 2a). Participants performed a total of 3 runs, leading to 30 trials per condition per subject. In the upright (UPR) condition an upright face gradually emerged from a 5% coherence image (Movie 1 in Supporting Information); in the inverted (INV) trials a face also slowly appeared from a noisy background but it was presented upside-down (Movie 2 in Supporting Information); no faces were shown in the scrambled (SCR) condition (Movie 3 in

Supporting Information). Specifically, in the upright (and inverted) faces condition the first 12 images consisted of alternating two images presented for 500 ms each containing upright (or inverted) faces with 5% and 10% of phase-coherence. After these initial 6 seconds, each of the remaining 18 stimuli was presented in succession, with increasing phase-coherence (i.e., increasing face visibility) progressively from 15% to 100% phase-coherence (Fig. 2b, Videos in Supporting Information). An automatic algorithm ensured that the order of the trials for the 3 conditions was randomized for every run. Within a run, a specific series of images (30) was presented only once for each of the three conditions and trials. However, as the total amount of images was 20 (see Stimuli) and the total amount of trials per condition was 30, one run (out of 3) was repeated in term of images presented (but the order of conditions was still randomized). The stimuli were delivered in E-Prime (Psychology Software Tool, Inc.).

A run started with a blank screen lasting for 3,000 ms after which the first trial was presented. Each 15 second trial was followed by a resting period (baseline, blank screen) jittered between 7,000 and 10,500 ms (14–21 volumes). During a single trial (sequence of 30 images) the participants were instructed to attend to the center of the screen and press a response key as soon as they perceive that a face (upright or inverted) was emerging from the “noise.” Their responses were considered as correct if the button presses were made before the end of the sequence (below 12,500 ms or 75% phase-coherence). The stimuli, including the rectangular background, subtended approximately 19° (height) × 19° (width) of visual angle and were presented on the center of the screen. The participants were not aware of any of the manipulations described above. The fMRI session lasted approximately 45 minutes.

Localization of Individuals FFA and OFA

The localization of the right FFA and OFA in each individual brain was performed in an independent functional localizer [Rossion et al., 2012a]. Four different categories of stimuli photographs were presented: faces, cars, scrambled faces, and scrambled cars. External features of faces (e.g., hair) were eliminated by editing the original photographs of those stimuli. Faces and cars were presented in color, in frontal view and they were embedded in a gray rectangle. Both categories consisted of 43 different stimuli (22 were female). In order to create the scrambled version of face and car stimuli a Fourier phase randomization procedure was performed. This was done to preserve the global low-level properties of the original image (luminance, contrast, spectral energy, etc.) while completely degrading category-related information. More specifically the algorithm performs a Fourier transformation of the face and car stimuli and it replaces the phase spectrum with random values while keeping the amplitude spectrum of the image unaltered [Nasanen, 1999]. The pictures of the four categories

of stimuli subtended equal shape, size, and contrast against background.

The four different categories of stimuli were presented in blocks. Participants performed three functional runs. In a single run a total of 24 blocks were presented, 6 for each category. One block consisted of 24 stimuli, each lasting for 750 ms for a total of 18 seconds (no resting condition was interleaved between two consecutive stimuli). Within a block, the same stimulus could be consecutively repeated two or three times.

A single stimulus was presented centrally and its size was 6.2° (height) × 5.5° (width) of visual angle. The stimulus location randomly varied from the central location in horizontal (6%) and vertical (8%) direction at each presentation. This was done in order to control that specific element belonging to the face/car stimuli (e.g., the eyes or headlights) were not shown at the same location in two consecutive trials. A rest condition lasting for 9 seconds and consisting of a cross on a black background centrally located was presented between two different blocks. Subjects were asked to perform a one-back task over the duration of the entire run (2 or 3 targets per block; 30 targets for each condition in total). Each run lasted for about 11 minutes, for a total duration of the face localizer experiment of 33 minutes.

Data Acquisition

Main experiment

Images for four subjects (S1, S2, S3, S6) were acquired on a Whole-body 3T Siemens Magnetom Trio scanner (Siemens Medical System, Erlangen, Germany) using a 32-channel head coil. The remaining two subjects (S4, S5) were measured using a Whole-body 3T Siemens Magnetom Prisma scanner (Siemens Medical System, Erlangen, Germany) using a 64-channel head-neck coil.

Eight oblique axial slices (in-plane resolution: 3.5 mm × 3.5 mm, slice thickness: 3.5 mm, interslice distance 0 mm) mainly covering the mid-inferior part of the occipital and temporal lobe were acquired using an echo planar imaging sequence (TR = 500 ms, TE = 30 ms (Trio) or TE = 28 ms (Prisma), matrix: 64 × 64, flip angle = 45°). For each run of the main experiment 1,421 volumes were acquired. The first four volumes were discarded from the analysis due to the T1 saturation effect.

Face localizer

Images for two (S3 and S6) subjects were acquired on a 3T Siemens Magnetom Allegra scanner (Siemens Medical System, Erlangen, Germany) using a birdcage volume coil. Two subjects (S1 and S2) were measured using a Whole body 3T Siemens Magnetom Trio scanner (Siemens Medical System, Erlangen, Germany) with a 32-channel head coil. And the images related to the remaining 2 participants (S4, S5) were acquired at a Whole body 3T Siemens

Magnetom Prisma scanner (Siemens Medical System, Erlangen, Germany) using a 64-channel head-neck coil.

Thirty-six oblique axial slices (in-plane resolution: $3.5 \text{ mm} \times 3.5 \text{ mm}$, slice thickness: 3.5 mm , interslice distance 0 mm) covering the entire cortical volume were acquired using an echo planar imaging sequence (same protocol for all subjects: repetition time (TR) = $2,250 \text{ ms}$, echo time (TE) = 30 ms , matrix: 64×64 , flip angle = 90°). For each run 293 volumes were acquired. The first four volumes were discarded from the analysis due to the T1 saturation effect.

Functional slices (of the main experiment and the face localizer) were aligned to a high resolution 3D anatomical dataset acquired in the middle of the entire session and consisting of 192 slices (ADNI sequence: TR = $2,250 \text{ ms}$; TE = 2.6 ms (Allegra) or 2.17 ms (Trio) or 2.21 ms (Prisma); flip angle = 9° , voxel dimension = $1 \times 1 \times 1 \text{ mm}^3$).

The participants were placed comfortably in the scanner and their head was fixated with foam pads. They saw the stimuli projected on a screen through a mirror mounted on the head coil. The visual field was perceived at a distance of 57 (Allegra) or 53.5 (Trio and Prisma).

Analysis

Both the functional and the anatomical data were analyzed using the BrainVoyager QX package 2.2.1 (Brain Innovation B.V., Maastricht, The Netherlands). The anatomical scans were used to project the statistical results from the functional data onto high-resolution anatomical images.

Functional data were pre-processed and aligned to the anatomical images. The pre-processing procedure started with correcting the data for motion artifacts in three dimensions and for slice scan-time differences. Subsequently, linear drifts were removed from the signal and data were high-pass filtered to remove slow frequency drifts up to 2 cycles per time course. Although the functional runs were analyzed at the individual level, after the pre-processing, they were aligned to the high-resolution anatomical images and normalized to the standard 3D Talairach space. This normalization procedure was performed only in order to compare the 3D coordinates of the two regions of interest namely, the FFA and the OFA with previous studies. The final version of the functional data consisted of a 4-dimensional (x, y, z, t) dataset in Talairach space for each run and participant.

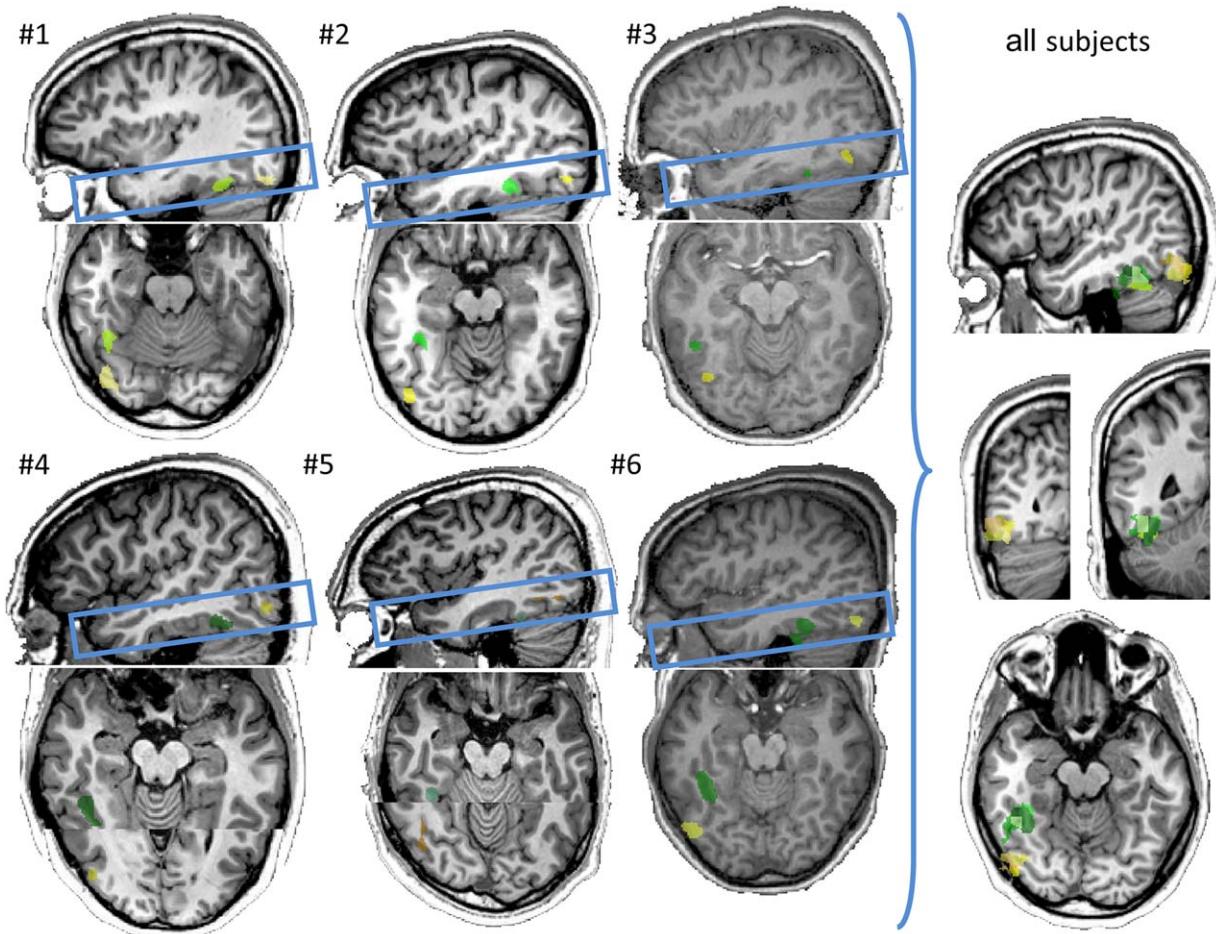
The FFA and the OFA were localized for each individual brain by performing a standard General Linear Model (GLM) analysis with the four stimulus types (*faces, cars, scrambled faces, and scrambled cars*) as predictors on the face-localizer functional runs. More specifically this localization consisted of several steps. First, a conjunction analysis between (*faces vs. cars*) and (*faces vs. scrambled faces*) was performed [Rossion et al., 2012a]. Clusters of voxels that were statistically significant after correction for multiple comparison via the false discovery rate approach (FDR

$- q$ (FDR) < 0.05) [Genovese et al., 2002] were selected. In order to define the final set of face-selective areas anatomical landmarks were also used [Weiner and Grill-Spector, 2012]. In particular, we focused on the MFG and the lateral part of the IOG, where the FFA and the OFA are typically located [e.g., Haxby et al., 2000; Ishai, 2008; Rossion et al., 2012a; Weiner and Grill-Spector, 2010].

The analysis of the main experiment was restricted to the OFA and FFA (localized via the procedure described above). We aimed to identify the point in time (in terms of volume of acquisition) at which upright and inverted face condition significantly differed from the scrambled face condition in each of the face-selective areas. The detection of this time-point was constrained such that the fMRI signal from the two conditions compared (upright vs. scrambled faces or inverted vs. scrambled faces) diverged significantly for at least three consecutive volumes of acquisition (i.e., $1,500 \text{ ms}$). In that case, the first of those three volumes was considered as the volume at which upright (inverted) and scrambled faces are functionally distinguished. The same analysis was also performed to identify the absolute timing of activation, that is, the time at which the signal differs from baseline.

The statistical comparison of the fMRI signal from two conditions volume by volume was performed as follows. For each condition a set 40 shifted stick functions were defined (for a total of 120 predictors). They comprised the number of volumes related to the 30 images presented in one event (30 volumes) and 10 additional volumes covering part of the following resting condition. The reason we included the latter predictors (i.e., representing the resting condition) is that we expected, based on the outcome of the EEG study of Ales et al. [2012] that the onset of the hemodynamic response related to the processing of a face would start at around the 16–17th volume (when the face in the background reaches 30%–35% of phase-coherence). Therefore, additional predictors were needed in order to “cover” most of the temporal extend of a typical hemodynamic response. With the reasonable assumption that the hemodynamic response profile is the same across conditions (upright, inverted and scrambled) and across the FFA and OFA, we used the 40 stick functions to estimate the % BOLD signal change in terms of beta values at each time point for each condition for the two areas via a GLM deconvolution analysis. Finally, the beta values related to upright and inverted conditions were statistically compared for a specific time point in each region of interest with zero (baseline) and with the betas from the scrambled conditions.

We also computed two parameters that allowed us to estimate eventual differences in term of hemodynamic response between the FFA and OFA: the time-to-peak (ttP) and the “full width at half maximum” (FWHM). The ttP consisted on the amount of volumes needed to reach the maximum beta from the first beta significantly different than zero. The FWHM represented the bandwidth of the response calculated as the interval (in term of volumes) between the

**Figure 3.**

Localization of the right FFA and OFA for the six subjects in Talairach space. The right FFA is color coded in green and the right OFA in yellow (a specific color tone is used for each subject). The consistency of localization across subjects is represented under the “all subjects” title. In this regard, both areas, localized for each of the six subjects were projected in the brain of a single subject (in the figure on the right). Note that, for

matter of visualization, the transversal view of subject #4 and #5 is composed by two parts: the top part shows the localization of the right FFA and the bottom part was used to show the right OFA. In fact, for those two subjects it was not possible to visualize on the same brain slice (viewed transversally) both the right FFA and right OFA. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

two points when the response reaches half the maximum (both when increases to and decreases from the peak).

the upright faces corresponded to the 18th volume, or 40% coherence, which is similar to what was found by Ales et al. [2012].

RESULTS

Behavior

The proportion of responses below 75% phase coherence was comparable for upright and inverted conditions (83.33% and 82.22%, respectively, $P = 0.38$). However, participants detected the face significantly faster for upright than inverted stimuli (9,442.53 ms and 9,739.75 ms, respectively, $P = 0.002$, with RT differences ranging between 139.09 ms and 528.66 ms). The average behavioral RT for

fMRI

The OFA and FFA were identified for each of the six participants in the right hemisphere (Fig. 3; Table I for coordinates) as well as in the left hemisphere. As the results from the left and right hemisphere were identical, for the sake of clarity and due to the higher relevance of the right hemisphere in face perception, we report in the main body of the text the findings from the right OFA and

TABLE I. X,Y and Z coordinates (in Talairach space) and amount of voxels (Vx) for the RH OFA and the RH FFA for each of the six subjects.

Subject #	RH OFA				RH FFA			
	X	Y	Z	Vx	X	Y	Z	Vx
#1	38	-75	-18	684	38	-50	-22	1082
#2	43	-77	-13	335	36	-41	-17	990
#3	39	-68	-11	276	47	-51	-18	1718
#4	43	-69	-7	168	40	-46	-15	702
#5	41	-60	-9	122	36	-44	-18	603
#6	45	-71	-17	585	38	-42	-22	2479
average	42 + -3	-70 + -6	-12 + -4		39 + -4	-46 + -4	-19 + -3	

FFA only (see Supporting Information for the fMRI results from the left OFA and FFA).

FFA versus OFA (upright faces)

The absolute activation level, that is, the percent signal change versus baseline (blank screen in between the individual trials) did not differ between the FFA and OFA for four out six subjects (only two subjects, #2 and #5 showed a significantly larger response in the FFA than the OFA, $P = 0.005$ and $P = 0.046$, respectively, all other P -values >0.05). In line with the more posterior localization of the OFA with respect to the FFA and the general hierarchical organization of the visual system, the absolute onset of activation (the time at which the percent signal change in the upright condition starts to differ significantly from the baseline, see Analysis) was earlier in the OFA than the FFA for five out of six subjects (Figs. 4, 5 and Table II). However, strikingly, for all individual participants, the BOLD signal significantly differed, during the stimulation sequence, for upright faces as compared with scrambled faces earlier, between 2 and 4 TRs of 500 ms, in the FFA than in the OFA (Figs. 4, 5 and Table II).

Upright versus inverted faces: Right FFA

Only one subject (#1) showed a significant difference in the magnitude of the BOLD response between upright and inverted faces ($P = 0.043$; all other $P > 0.05$). However, for six out of six subjects, we observed a difference in the time onset at which the fMRI signal in the upright and inverted condition diverged from the phase-scrambled condition. More specifically, the time at which upright and scrambled faces differed systematically preceded the time onset of the difference between inverted and scrambled faces (between 1 and 3 TRs, and 8 TRs for subject #5 who had a low response for inverted faces, in the FFA, see Fig. 6 and Table II).

Upright versus inverted faces: Right OFA

In the right OFA, there was a difference between the two conditions in the level of activation only for subject #1

and #3 ($P = 0.028$ and $P = 0.038$, respectively; all other $P > 0.05$).

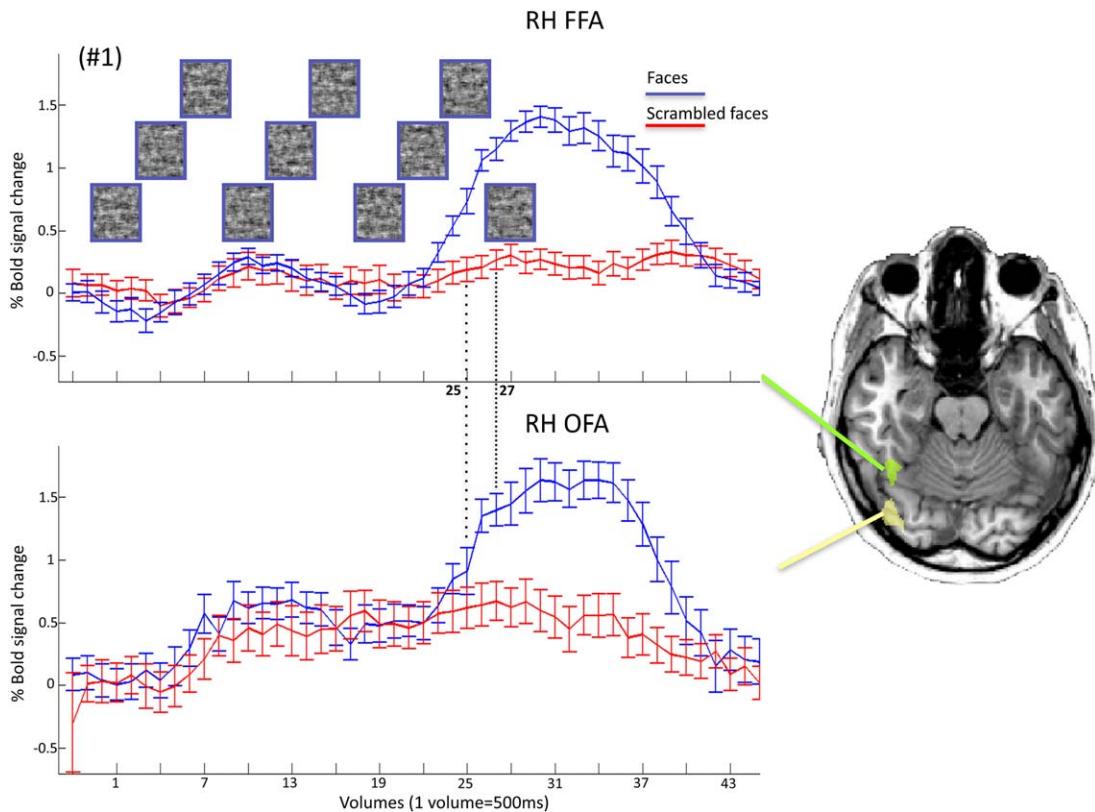
Yet, for five out of six subjects, upright faces diverged from the scrambled faces earlier during the stimulation sequence than the inverted faces (Fig. 7 and Table II). Moreover, for the subject who showed the opposite trend (upright faces diverged from scrambled faces later than inverted faces), we observed a significant difference between upright and scrambled faces already at the 27th volume ($P = 0.013$, i.e., before the 28th volume of inverted vs. scrambled faces). However, we did not consider the 27th volume as the "representative" volume, since the difference between upright and scrambled faces of the following volume (the 28th) was only marginally significant ($P = 0.061$) (according to our criteria the volume representing the first difference between two condition needed to be the first of three consecutive significant values, see Analysis). Therefore, with a slightly more liberal threshold, all six subjects showed the same effect.

TTP and FWHM for FFA and OFA

For five out of six subjects the time-to-peak was much longer for the OFA than the FFA both for the upright and inverted conditions. The same observation was made for the bandwidth of the upright and inverted response in the two regions, namely we observed for the upright response that for four out of six subjects the FWHM was larger in correspondence of OFA than FFA and for one participant that parameter was identical. In the inverted condition the bandwidth was larger for the OFA than the FFA for five subjects (out of six) (Table III).

DISCUSSION

In this study, we presented 15 seconds stimulation sequence of upright, inverted and scrambled faces changing in local contrast occurring every TR (500 ms) while preserving low-level visual information (i.e., power spectrum) constant throughout the sequence. The appearance of a face shape in the sequence triggered highly significant fMRI activation in the predefined face-selective regions FFA and OFA of all six individual brains tested, both for upright and

**Figure 4.**

Time courses of BOLD activation in the right hemisphere FFA and OFA for the upright and scrambled face conditions for a representative subject (#1). Each time point corresponds to the BOLD signal for a specific TR (500 ms) and stimulus. Pictures in the top graph represent 10 (out of 30) examples of the face

stimuli at 5%, 10%, 5%, 10%, 25%, 40%, 55%, 70%, 85%, and 100% phase-coherence presented in the upright condition. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

inverted faces. Three main effects were observed. First, the BOLD signal rose above baseline in the OFA before any differentiation between upright and scrambled faces, but not in the FFA for the majority of subjects. Second and in contrast with this result, the difference between upright and scrambled faces emerged later in the OFA than the FFA, showing face detection based on structural information first in the higher-order region of the cortical face network. Third and in line with behavioral responses, the face versus scrambled face difference appeared earlier in the sequence for upright than inverted stimuli, both in the FFA and OFA, thus revealing a face inversion effect in the time-domain for the first time in fMRI. We discuss these findings in details in the following sections.

An Early Onset of Activation in the OFA Due to Local Low-Level Changes

The statistical analysis at the individual level showed that the BOLD signal in response to upright faces rose

above baseline relatively early in the OFA, significantly earlier than in the FFA. Importantly, this initial increase of activation in the OFA is observed both for faces and phase-scrambled faces, with no difference between the two conditions. This early increase of the OFA in the stimulation sequence could be due to specific statistical properties of the images contained in the amplitude spectrum [Crouzet et al., 2010; VanRullen, 2006] which are preserved by phase-scrambling. However, for four participants (out of six) the increase of the fMRI signal over baseline in the OFA occurs before the 13th TR, that is, when the phase-coherence merely alternate between 5% and 10%. Since there is no evidence of a face-related response at all in EEG signals below 30% phase-coherence in such stimulation sequences [Ales et al., 2012; Liu-Shuang et al., 2015], this early OFA activity increase cannot be due to the contribution of additional face-related low-level information, such as differences in local cues between faces and non-faces. Rather, it indicates that low-level visual information, here represented by local changes in contrast depending

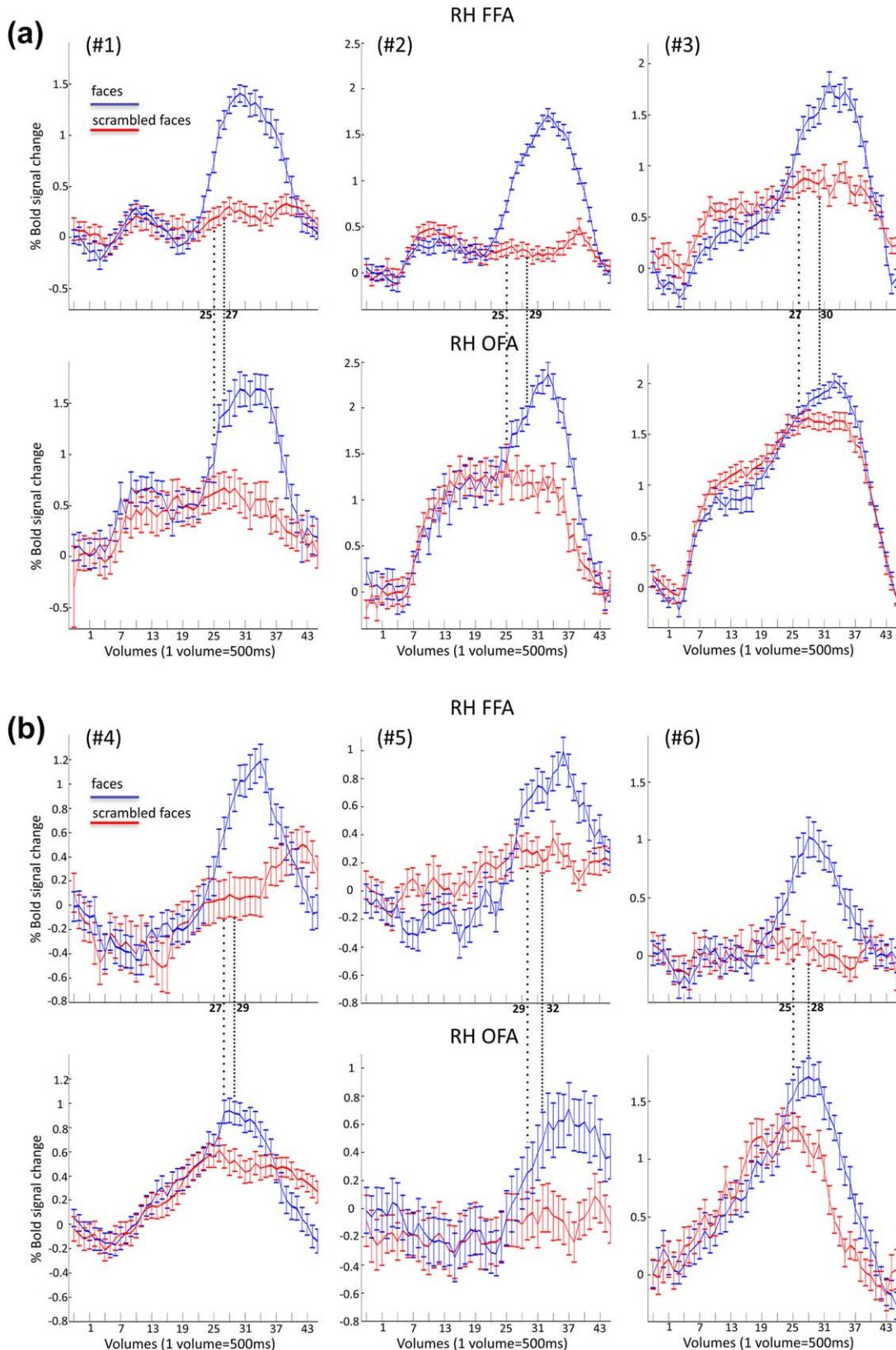


Figure 5.

Time courses related to the right FFA (top) and the right OFA (bottom) for upright and scrambled condition for each of the six subjects (#1, #2, #3 in **a**; #4, #5, #6 in **b**). The dashed line represents the time point at which the BOLD signal in the upright face condition is significantly larger than the scrambled condition in

the right FFA (with the constraint that a significant difference has to be found on the next consecutive two points). The dotted line provides the same type of information but within the right OFA. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

TABLE II. The time point (Onset) in TRs and corresponding P-value at which: the UPR condition is significantly larger than baseline (leftward column), the UPR condition is significantly larger than the SCR condition (central column) and the INV condition is significantly larger than the SCR condition (rightward column). All those values were computed for the RH FFA (a) and the RH OFA (b) for each of the six subjects.

(a) RH FFA

	UPR vs. baseline		UPR vs. SCR		INV vs. SCR	
	Onset TR	P-value	Onset TR	P-value	Onset TR	P-value
Subject #						
#1	24	0.047	25	0.003	26	0.025
#2	24	0.013	25	0.010	26	0.004
#3	22	0.004	27	0.008	29	0.007
#4	27	0.017	27	0.027	30	0.025
#5	28	0.042	29	0.043	37	0.009
#6	23	0.015	25	0.009	28	0.018

(b) RH OFA

	UPR vs. baseline		UPR vs. SCR		INV vs. SCR	
	Onset TR	P-value	Onset TR	P-value	Onset TR	P-value
Subject #						
#1	10	0.004	27	0.000	28	0.018
#2	10	0.001	29	0.001	28	0.024
#3	8	0.000	30	0.045	31	0.003
#4	18	0.016	29	0.003		
#5	34	0.012	32	0.016		
#6	8	0.040	28	0.009	32	0.028

on the randomization of phase at every step, triggers and contributes to the response of the OFA, a region that is closer to retinotopic visual areas and in which populations of neurons have smaller receptive fields than the FFA [Henriksson et al., 2015; Sayres and Grill-Spector, 2008].

Strikingly, with the exception of the data of two subjects (#2 and #3), the FFA does not rise at all to these local changes of contrast as showed by the lack of activation to the fully phase-scrambled faces (Fig. 5). This finding indicates that the OFA is more sensitive than the FFA to local low-level visual information changes and perhaps to power spectrum information characterizing face stimuli. At first sight this finding contradicts previous observations suggesting that the FFA responds to large differences of low-level visual information (e.g., size, position, global contrast, color, global power-spectrum) [Andrews et al., 2010; Rossion et al., 2012a; Yue et al., 2011]. However, it is important to point out that here all these low-level changes were strictly controlled, and only local, but not global, contrast changed from one phase-scrambled step to the next.

As a consequence of the higher sensitivity of the OFA compared with FFA to low-level information changes, the time-to-peak (i.e., from onset of activation to peak) differs between the two regions, being much longer in the OFA than FFA.

Earlier Face Detection in the FFA than the OFA During Gradual Revealing of Structural Face Information

Despite being less or no sensitive to low-level visual information, the FFA showed earlier response to the emergence of structural (i.e., high level) face information than the OFA, that is, the difference between upright and scrambled faces emerged earlier in the FFA than the OFA. This finding replicates and extends previous observations of Jiang and colleagues with visual scenes containing people and vehicles [Jiang et al., 2011, 2015]. Here, faces were not compared with other object categories. However, the functional ROIs where the signal from the main experiment was extracted from, the OFA and the FFA, were defined based on their selective response to faces (as compared with objects). Therefore, the onset difference between the FFA and OFA in terms of face versus scrambled face discrimination is likely to be due to face-selective processes. This claim is strengthened by the large hemodynamic response (ranging between 1% and 1.7% signal change across participants) observed in the upright face condition within these regions.

Most importantly, since the absolute level of the BOLD response was comparable across the two regions, the onset latency difference between regions cannot be attributed to

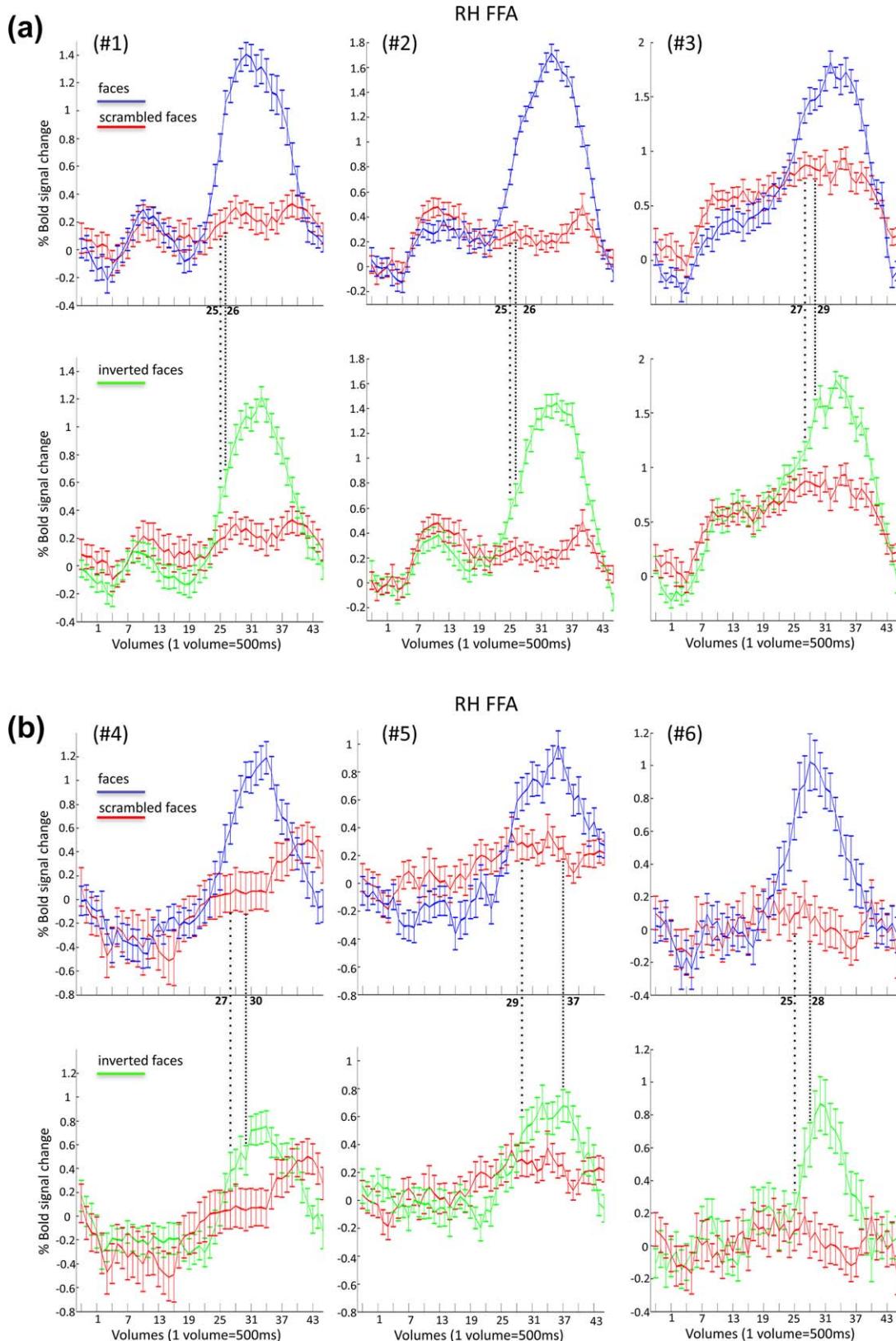


Figure 6.

Time courses related to the right FFA for upright versus scrambled (top) and inverted versus scrambled condition (bottom) for each of the six subjects (#1, #2, #3 in a; #4, #5, #6 in b). The dashed line represents the time point at which the BOLD signal in the upright condition is significantly larger than the scrambled condition in the right FFA (constraint to the fact

that the successive two time points have to show the same significant difference as well). The dotted line provides the same type of information but comparing inverted and scrambled condition. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

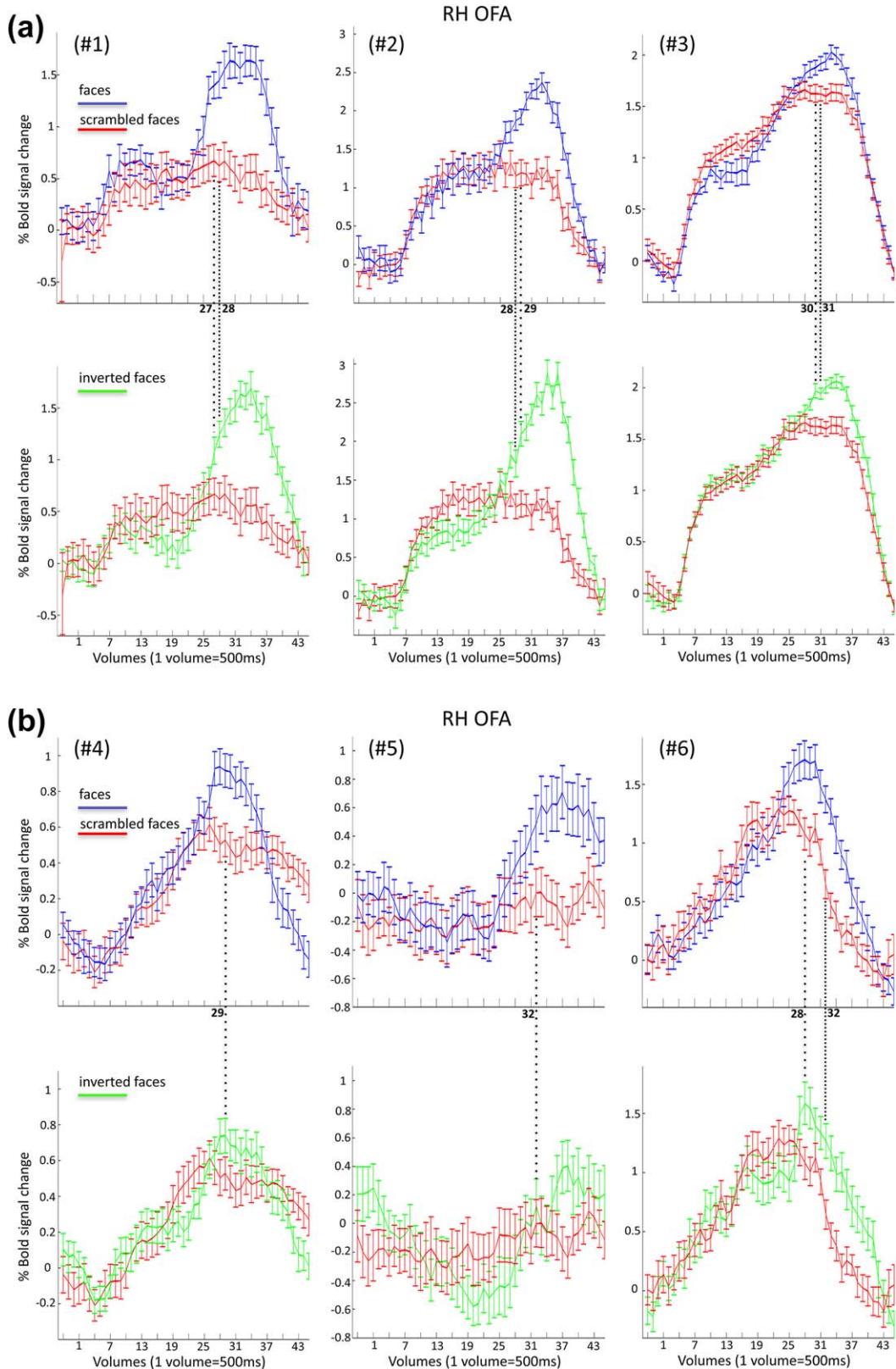


Figure 7.

Time courses related to the right OFA for upright versus scrambled (top) and inverted versus scrambled condition (bottom) for each of the six subjects (#1, #2, #3 in a; #4, #5, #6 in b). The dashed line represents the time point at which the BOLD signal in the upright condition is significantly larger than the scrambled face condition in the right OFA (with the

constraint that a significant difference has to be found on the next consecutive two time points). The dotted line provides the same type of information but comparing inverted and scrambled face conditions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

TABLE III. Leftward column: the time point in TRs at which the fMRI response peaks (Peak); central column: interval of time in TRs needed to reach the maximum beta from the first beta significantly different than zero (Ttp); rightward column: interval of time in TRs between the two points when the response reaches half the maximum (both when increases to and decreases from the peak) (FHW_M). All those values were computed for the UPR and INV response in the RH FFA (a) and the RH OFA (b) for each of the six subjects.

(a) RH FFA

Subject #	Peak (TR)		Ttp (TR interval)		FHW _M (TR interval)	
	UPR	INV	UPR	INV	UPR	INV
#1	31	34	7	8	13	12
#2	34	35	10	10	13	12
#3	33	34	11	24	14	13
#4	35	35	8	5	11	8
#5	37	34	9	4	11	10
#6	29	31	6	4	10	9

(b) RH OFA

Subject #	Peak (TR)		Ttp (TR interval)		FHW _M (TR interval)	
	UPR	INV	UPR	INV	UPR	INV
#1	32	35	22	11	13	13
#2	34	35	24	25	16	14
#3	34	35	26	28	19	25
#4	29	30	11	6	13	14
#5	38		4		7	
#6	31	29	23	22	19	26

a larger response in the FFA than OFA [Thompson et al., 2014]. Moreover, the significance of the temporal difference between the FFA and OFA across subjects at the *individual level* made this effect stronger and more consistent than the difference observed previously [Jiang et al., 2011]. This is certainly partly due to the fast sampling rate used here (500 ms), providing a 2.5 increase in time resolution compared with previous studies.

Here the earlier activation to gradually revealed faces observed in FFA compared with OFA cannot be accounted for by other cues than the face, since there were no body parts or specific backgrounds presented with the face stimuli. This issue is particularly relevant as it has been shown that the FFA responds to body parts, for example when presenting headless bodies, more consistently than the OFA [e.g., Peelen and Downing, 2005; Pinski et al., 2009; Schwarzkopf et al., 2005, 2008; Weiner and Grill-Spector, 2010].

Another element that distinguishes the present study from previous studies using a similar approach in fMRI

[Esterman and Yantis, 2010; Reinders et al., 2005; Reinders et al., 2006], is the strictly linear increase of phase information throughout a stimulation sequence [Ales et al., 2012; Liu-Shuang et al., 2015]. This procedure is particularly advantageous because it avoids any abrupt increase or decrease in phase coherence that might blur onset differences in face-selective response between brain regions.

Moreover, the same type of sequences has already been used at a faster rate (i.e., 6 Hz or 166 ms/image) while recording high temporal resolution electrophysiological signals to investigate the threshold at which faces are detected [Ales et al., 2012; Liu-Shuang et al., 2015]. In these studies, face-related EEG signal typically emerges at 35%–40% phase coherence, which would correspond to the 17th–18th TR here (Fig. 2a). Interestingly, here the face-selective response emerges significantly at the 25th TR in most participants (Fig. 5). Hence, in this paradigm, the delay between neuronal populations' early face-selective response and its translation into a detectable hemodynamic response appears to be of about 3 seconds.

The Non-Hierarchical View of Cortical Face Detection

Overall, our findings support a non-hierarchical view of cortical face processing [Jiang et al., 2011, 2015; Rossion, 2008a; Rossion et al., 2003, 2011; see also Atkinson and Adolphs, 2011; Duchaine and Yovel, 2015], inspired from reverse hierarchical cortical processing in the visual system in general [Ahissar and Hochstein, 2004; Bullier, 2001; Mumford, 1992]. More specifically, we showed that gradually revealing structural face information activates the independently-defined FFA before, and thus independently, of activation related to structural information in the OFA. This claim is consistent with findings that brain damage to the cortical territory of the OFA does not prevent the observation of robust face-selective responses in the right FFA [Rossion et al., 2003; Sorger et al., 2007; Steeves et al., 2006]. It also agrees with activation in the right FFA but not in the OFA for ambiguous visual stimuli successfully categorized as faces based on prior knowledge [Dolan et al., 1997], their global configuration rather than their local properties [Rossion et al., 2011], or limited low spatial frequency information [Goffaux et al., 2011].

Importantly, the latter observation suggests that the timing difference observed here between the OFA and FFA with regard to the latency of face-related activation may be due to a higher resistance to degraded or noisy stimuli in the FFA compared with the OFA. That is, if the different frames of the sequence were presented one by one in random order rather than in an ordered sequence, the FFA could respond to a larger extent than the OFA to faces presented at an intermediate level of phase-descrambling. Therefore, presenting the stimuli with an increasing order of visibility leads to an earlier activation in the FFA than the OFA in response to faces. As previously discussed

elsewhere [Jiang et al., 2011, 2015] such a higher sensitivity of the FFA compared with the OFA to a face embedded in a noisy background is a plausible account of our observations, and it actually reflects exactly the phenomenon that we attempted to investigate. In fact, in real life conditions, when faces appear far away or in the periphery, or have to be detected in visual scenes under conditions of low visual acuity and/or contrast sensitivity, occlusion, or reduced visibility, there is always an increasing order of visibility and accumulation of information (e.g., lower to higher spatial frequencies) [Goffaux et al., 2011; Hegde, 2008; Sargent, 1986]. The present display merely slows down artificially this gradual accumulation of evidence in the visual system in order to disclose latency differences between functional brain regions of interest [see also Ramon et al., 2015]. Obviously, these latency differences are relative rather than being absolute: the 2–3 TR (i.e., 1–1.5 seconds) onset difference in absolute latency of the face response between the FFA and the OFA could be increased/decreased if smaller/larger steps than 5% per TR were presented in the stimulation sequence.

The earlier activation for faces in the FFA than the OFA in the ordered sequence in our study suggests the presence of anatomico-functional connections from low-level visual areas which would bypass the OFA. Findings from both functional [Fairhall and Ishai, 2007; Turk-Browne et al., 2010; Zhu et al., 2011] and structural [Gomez et al., 2013; Gschwind et al., 2012; Pyles et al., 2013; Tavor et al., 2014] connectivity studies suggest that face-selective regions are inter-connected and at least two studies reported white matter tracts connecting early visual areas to face-selective regions in correspondence to the FG and STS [Gschwind et al., 2012; Kim et al., 2006]. Therefore, these tracts could provide additional routes for information to reach the FFA independently of OFA, and may explain the earlier onset of face-selective responses in the FFA.

Interestingly, despite the earlier onset of activation in the FFA than in the OFA for structural face information, there was either no difference in the latency of the peak of activation between these regions or this peak difference was smaller than the onset latency difference. Consequently, the time onset-to-peak is shorter in the OFA than the FFA, suggesting that the OFA follows the FFA only in terms of onset of face-related activation: once the OFA is activated by structural face information, the two areas may act in concert to process faces, together with more anterior face-selective areas of the ventral occipito-temporal cortex [e.g., Jonas et al., 2015; Rajimehr et al., 2009].

In summary, we argue that the present results, in combination with previous studies specifically targeting the dynamic and relationship between the FFA and OFA in the context of face perception, constitute converging evidence that the OFA does not necessarily constitutes the first relay station of structural face processing in the human brain. In this context and contrary to the most commonly held perspective, a non-hierarchical model is a

more likely framework of the early stages of the functional neuroanatomy of face processing. That is, the OFA could exhibit high-level (structural) face-selective responses only following previous categorization of face stimuli within the FFA. Although the exact mediator that leads to the OFA responses is still unknown, we posit that face-selective responses in the IOG (i.e., the OFA) arise through putative re-entrant connections with areas anteriorly located as FFA [Rossion, 2008a; Rossion et al., 2011].

A Temporal Face Inversion Effect in Face-Selective Areas

Picture-plane inversion is the most documented manipulation that disproportionately impairs perception for faces compared with objects [Yin, 1969 for inversion; see Farah et al., 1998; Rossion, 2008b for reviews]. However, this behavioral effect of inversion have been most often reported in the context of individual face discrimination (i.e., the differentiation of facial identities) [e.g., Freire et al., 2000] whereas the effect of inversion on face detection (i.e. the perception of a face as a face) is less well described. Behavioral studies that have addressed this issue generally show that inversion reduces accuracy but also slows down face detection relative to upright faces [Garrido et al., 2008; Lewis and Edmonds, 2003, 2005; Parkin and Williamson, 1987; Purcell and Stewart, 1986a,b; Rousselet et al., 2003; VanRullen, 2006]. In line with the recent EEG study using the same stimulation sequences as used here but at a much faster changing rate (6Hz) [Liu-Shuang et al., 2015], in the present study participants were also behaviorally slower at detecting the appearance of inverted than upright faces.

In terms of temporal delay at the neural level, inversion increases the latency of the face-selective neurons in the monkey infero-temporal cortex [Perrett et al., 1982, 1998], the face-sensitive N170 component over the human scalp [Bentin et al., 1996; Jeffreys, 1993; Rossion et al., 1999] and the steady-state visual evoked potential (SSVEP) elicited by faces [Rossion et al., 2012b]. This increase is in the order of about 10 ms for face stimuli presented in full view. However, when gradually revealing stimuli in a noisy background, the response to inverted faces can be substantially delayed. For instance, in the recent EEG study of Liu-Shuang et al., [2015], the face detection threshold increased by about 15% coherence. Here, we also observed, in the individually-defined FFAs (for all subjects) and OFAs (for five out of six subjects), a significant delay in the onset of face-related response to upside-down as compared with upright faces. In both the FFA and OFA, this temporal shift cannot be due to an overall difference in amplitude between the conditions, as this was not significant at the individual level (except for subject #1 in the FFA and subjects #1 and #3 in the OFA).

CONCLUSIONS AND PERSPECTIVES

By combining a fast sampling rate with a slowly increasing visual stimulation display, we showed that the most anterior face-selective region targeted here, the FFA, was characterized by an earlier response to structural face information as compared with the OFA, despite an earlier activation of the OFA above baseline. Moreover, both the FFA and the OFA showed an inversion effect in terms of increased latency, that is, a delayed significant difference between faces and scrambled faces for inverted than upright stimuli.

One outstanding issue is whether other face areas of the face-selective network, such as the posterior superior temporal sulcus (pSTS) [Puce et al., 1998], the amygdala [e.g., Vuilleumier et al., 2001], the anterior fusiform gyrus (AFG) [Jonas et al., 2015], the anterior collateral sulcus ("AT") [Nasr and Tootell, 2012], the anterior inferior temporal lobe [Leveroni et al., 2000] or the inferior frontal gyrus (IFG) [e.g., Nakamura et al., 1999], dissociate from the OFA and FFA regarding their onset time of activation. In previous studies with people in visual scenes, the FFA was systematically associated with the earliest onset time of activation compared with all other activated areas of the face-selective network mentioned above [Jiang et al., 2011, 2015]. Here, in order to maximize the chance to observe timing differences by increasing the TR, we were not able to address this issue since we only acquired few slices centered on the FFA and OFA. This issue may be addressed in future studies with higher magnetic fields providing better temporal and spatial resolution.

In the present experiment, the subject was actively searching for a face stimulus in the gradually revealing stimulation display. An open question is whether this procedure speeds up the onset time of face detection as opposed to a passive viewing of the stimulation, or the search for another kind of stimulus [Jiang et al., 2015]. Recent studies provide evidence for both scenarios. On one hand, data from the study of Jiang et al. [2015] indicate that category-search enhance and speed up category-selective responses in the human brain. For instance, searching actively for inverted faces only may reduce the inversion effect obtained in amplitude and latency of the BOLD response. On the other hand, we also observed that the coherence level at which face-related activity emerges in EEG did not differ when subjects performed an active face detection task [Ales et al., 2012] versus passive viewing [Liu-Shuang et al., 2015], although in the latter case, only one face stimulus was presented in all trials, making the stimulation highly predictable.

Finally and more generally, the present work highlights the relevance for fMRI studies of visual categorization of a continuously changing mode of visual stimulation. In typical (fMRI) studies, various categories of stimuli are presented transiently, that is, flashed, to the visual system. This sudden onset of complex visual items evokes a large visual response that primarily includes responses un-related to the

specific difference between the stimuli. This common activation, which is typically removed by post-hoc subtraction procedures rather than being neutralized in the design, may blur the difference between the conditions of interest. In contrast, by gradually revealing category information—while keeping low-level visual information constant—the specific difference (here between upright, inverted and scrambled faces) occurs at a distance from the transient stimulation and one is able to capture it fully in the BOLD response. This procedure maximizes contrast and has significant impact on the statistical power of the entire experiment as it provides a high signal-to-noise ratio and consequently significant responses in individual brains within few trials (here 30). Future fMRI studies could take full advantage of this approach to disclose differences at the population level for between and within category changes of information occurring in a continuous stimulation flow, as used, for instance, recently in a familiar/unfamiliar face categorization task [Ramon et al., 2015].

REFERENCES

- Aguirre GK, Zarahn E, D'Esposito M (1998): The variability of human, BOLD hemodynamic responses. *Neuroimage* 8: 360–369.
- Ahissar M, Hochstein S (2004): The reverse hierarchy theory of visual perceptual learning. *Trends Cogn Sci* 8:457–464.
- Ales JM, Farzin F, Rossion B, Norcia AM (2012): An objective method for measuring face detection thresholds using the sweep steady-state visual evoked response. *J Vis* 12:18.
- Allison T, Puce A, Spencer DD, McCarthy G (1999): Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cereb Cortex* 9:415–430.
- Andrews TJ, Clarke A, Pell P, Hartley T (2010): Selectivity for low-level features of objects in the human ventral stream. *NeuroImage* 49:703–711.
- Atkinson AP, Adolphs R (2011): The neuropsychology of face perception: Beyond simple dissociations and functional selectivity. *Philos Trans R Soc Lond Ser B, Biol Sci* 366:1726–1738.
- Axelrod V, Yovel G (2010): External facial features modify the representation of internal facial features in the fusiform face area. *NeuroImage* 52:720–725.
- Bentin S, Allison T, Puce A, Perez E, McCarthy G (1996): Electrophysiological studies of face perception in humans. *J Cogn Neurosci* 8:551–565.
- Bullier J (2001): Integrated model of visual processing. *Brain Res Brain Res Rev* 36:96–107.
- Coggan DD, Liu W, Baker DH, Andrews TJ (2016): Category-selective patterns of neural response in the ventral visual pathway in the absence of categorical information. *NeuroImage* 135:107–114.
- Crouzet SM, Kirchner H, Thorpe SJ (2010): Fast saccades toward faces: Face detection in just 100 ms. *J Vis* 10:16 1–17.
- Davies-Thompson J, Gouws A, Andrews TJ (2009): An image-dependent representation of familiar and unfamiliar faces in the human ventral stream. *Neuropsychologia* 47:1627–1635.
- Dolan RJ, Fink GR, Rolls E, Booth M, Holmes A, Frackowiak RS, Friston KJ (1997): How the brain learns to see objects and faces in an impoverished context. *Nature* 389:596–599.

- Duchaine B, Yovel G (2015): A revised neural framework for face processing. *Annu Rev Vis Sci* 1:24.
- Esterman M, Yantis S (2010): Perceptual expectation evokes category-selective cortical activity. *Cereb Cortex* 20:1245–1253.
- Ewbank MP, Henson RN, Rowe JB, Stoyanova RS, Calder AJ (2013): Different neural mechanisms within occipitotemporal cortex underlie repetition suppression across same and different-size faces. *Cereb Cortex* 23:1073–1084.
- Fairhall SL, Ishai A (2007): Effective connectivity within the distributed cortical network for face perception. *Cereb Cortex* 17: 2400–2406.
- Farah MJ, Wilson KD, Drain M, Tanaka JN (1998): What is “special” about face perception? *Psychol Rev* 105:482–498.
- Formisano E, Goebel R (2003): Tracking cognitive processes with functional MRI mental chronometry. *Curr Opin Neurobiol* 13: 174–181.
- Fox CJ, Iaria G, Barton JJ (2008): Disconnection in prosopagnosia and face processing. *Cortex* 44:996–1009.
- Freire A, Lee K, Symons LA (2000): The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception* 29:159–170.
- Frost MA, Goebel R (2012): Measuring structural-functional correspondence: Spatial variability of specialised brain regions after macro-anatomical alignment. *Neuroimage* 59:1369–1381.
- Garrido L, Duchaine B, Nakayama K (2008): Face detection in normal and prosopagnosic individuals. *J Neuropsychol* 2:119–140.
- Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW (2000): The fusiform “face area” is part of a network that processes faces at the individual level. *J Cogn Neurosci* 12: 495–504.
- Genovese CR, Lazar NA, Nichols T (2002): Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15:870–878.
- Gentile F, Rossion B (2014): Temporal frequency tuning of cortical face-sensitive areas for individual face perception. *NeuroImage* 90:256–265.
- Gilaie-Dotan S, Gelbard-Sagiv H, Malach R (2010): Perceptual shape sensitivity to upright and inverted faces is reflected in neuronal adaptation. *NeuroImage* 50:383–395.
- Goffaux V, Peters J, Haubrechts J, Schiltz C, Jansma B, Goebel R (2011): From coarse to fine? Spatial and temporal dynamics of cortical face processing. *Cereb Cortex* 21:467–476.
- Gomez J, Yoon J, Golarai G, Grill-Spector K (2013): Development of right inferior longitudinal fasciculus correlates specifically with face perception. *J Vis* 13:178.
- Gschwind M, Pourtois G, Schwartz S, Van De Ville D, Vuilleumier P (2012): White-matter connectivity between face-responsive regions in the human brain. *Cereb Cortex* 22:1564–1576.
- Handwerker DA, Ollinger JM, D’Esposito M (2004): Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage* 21:1639–1651.
- Harris A, Aguirre GK (2008): The representation of parts and wholes in face-selective cortex. *J Cogn Neurosci* 20:863–878.
- Harris A, Aguirre GK (2010): Neural tuning for face wholes and parts in human fusiform gyrus revealed by fMRI adaptation. *J Neurophysiol* 104:336–345.
- Haxby JV, Ungerleider LG, Clark VP, Schouten JL, Hoffman EA, Martin A (1999): The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* 22:189–199.
- Haxby JV, Hoffman EA, Gobbini MI (2000): The distributed human neural system for face perception. *Trends Cogn Sci* 4: 223–233.
- Hegde J (2008): Time course of visual perception: Coarse-to-fine processing and beyond. *Prog Neurobiol* 84:405–439.
- Heimond CC, Kanwisher NG, Op de Beeck HP (2007): A preference for contralateral stimuli in human object- and face-selective cortex. *PloS One* 2:e574.
- Henriksson L, Mur M, Kriegeskorte N (2015): Faciotopy-A face-feature map with face-like topology in the human occipital face area. *Cortex* 72:156–167.
- Honey C, Kirchner H, VanRullen R (2008): Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *J Vis* 8:9 1–13.
- Ishai A (2008): Let’s face it: It’s a cortical network. *NeuroImage* 40:415–419.
- Jeffreys DA (1993): The influence of stimulus orientation on the vertex positive scalp potential evoked by faces. *Exp Brain Res* 96:163–172.
- Jiang F, Badler JB, Righi G, Rossion B (2015): Category search speeds up face-selective fMRI responses in a non-hierarchical cortical face network. *Cortex* 66:69–80.
- Jiang F, Dricot L, Weber J, Righi G, Tarr MJ, Goebel R, Rossion B (2011): 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.
- Jonas J, Rossion B, Brissart H, Frismand S, Jacques C, Hossu G, Colnat-Coulbois S, Vespiagnani H, Vignal JP, Maillard L (2015): Beyond the core face-processing network: Intracerebral stimulation of a face-selective area in the right anterior fusiform gyrus elicits transient prosopagnosia. *Cortex* 72:140–155.
- Jonas J, Rossion B, Krieg J, Koessler L, Colnat-Coulbois S, Vespiagnani H, Jacques C, Vignal JP, Brissart H, Maillard L (2014): Intracerebral electrical stimulation of a face-selective area in the right inferior occipital cortex impairs individual face discrimination. *NeuroImage* 99:487–497.
- Jonas J, Jacques C, Liu-Shuang J, Brissart H, Colnat-Coulbois S, Maillard L, Rossion B (2016): A face-selective ventral occipitotemporal map of the human brain with intracerebral potentials. *Proc Natl Acad Sci U S A* 113:E4088.
- Kanwisher N, Yovel G (2006): The fusiform face area: A cortical region specialized for the perception of faces. *Philos Trans R Soc Lond Ser B, Biol Sci* 361:2109–2128.
- Kanwisher N, McDermott J, Chun MM (1997): The fusiform face area: A module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302–4311.
- Kim M, Ducros M, Carlson T, Ronen I, He S, Ugurbil K, Kim DS (2006): Anatomical correlates of the functional organization in the human occipitotemporal cortex. *Magn Reson Imaging* 24: 583–590.
- Laguesse R, Dormal G, Biervoye A, Kuefner D, Rossion B (2012): Extensive visual training in adulthood significantly reduces the face inversion effect. *J Vis* 12:14.
- Leveroni CL, Seidenberg M, Mayer AR, Mead LA, Binder JR, Rao SM (2000): Neural systems underlying the recognition of familiar and newly learned faces. *J Neurosci* 20:878–886.
- Lewis MB, Edmonds AJ (2003): Face detection: Mapping human performance. *Perception* 32:903–920.
- Lewis MB, Edmonds AJ (2005): Searching for faces in scrambled scenes. *Vis Cogn* 12:1309–1336.

- Liu-Shuang J, Ales J, Rossion B, Norcia AM (2015): Separable effects of inversion and contrast-reversal on face detection thresholds and response functions: A sweep VEP study. *J Vis* 15:11.
- Mazard A, Schiltz C, Rossion B (2006): Recovery from adaptation to facial identity is larger for upright than inverted faces in the human occipito-temporal cortex. *Neuropsychologia* 44:912–922.
- Moeller S, Freiwald WA, Tsao DY (2008): Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science* 320:1355–1359.
- Mumford D (1992): On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol Cybern* 66: 241–251.
- Nakamura K, Kawashima R, Ito K, Sugiura M, Kato T, Nakamura A, Hatano K, Nagumo S, Kubota K, Fukuda H, Kojima S (1999): Activation of the right inferior frontal cortex during assessment of facial emotion. *J Neurophysiol* 82:1610–1614.
- Nasanen R (1999): Spatial frequency bandwidth used in the recognition of facial images. *Vis Res* 39:3824–3833.
- Nasr S, Tootell RB (2012): Role of fusiform and anterior temporal cortical areas in facial recognition. *NeuroImage* 63:1743–1753.
- Parkin AJ, Williamson P (1987): Cerebral lateralization at different stages of facial processing. *Cortex* 23:99–110.
- Peelen MV, Downing PE (2005): Selectivity for the human body in the fusiform gyrus. *J Neurophysiol* 93:603–608.
- Perrett DI, Rolls ET, Caan W (1982): Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47: 329–342.
- Perrett DI, Oram MW, Ashbridge E (1998): Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition* 67:111–145.
- Pinsk MA, Arcaro M, Weiner KS, Kalkus JF, Inati SJ, Gross CG, Kastner S (2009): Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *J Neurophysiol* 101:2581–2600.
- Pitcher D, Dilks DD, Saxe RR, Triantafyllou C, Kanwisher N (2011): Differential selectivity for dynamic versus static information in face-selective cortical regions. *NeuroImage* 56: 2356–2363.
- Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998): Temporal cortex activation in humans viewing eye and mouth movements. *J Neurosci* 18:2188–2199.
- Purcell DG, Stewart AL (1986a): The face-detection effect. *B Psychonomic Soc* 24:118–120.
- Purcell DG, Stewart AL (1986b): The face-detection effect - splitting is as good as inverting. *B Psychonomic Soc* 24: 349–349.
- Purcell DG, Stewart AL (1988a): The face-detection effect - configuration enhances detection. *Percept Psychophys* 43:355–366.
- Purcell DG, Stewart AL (1988b): The face detection effect - effects of positional uncertainty and contrast. *B Psychonomic Soc* 26: 517–517.
- Pyles JA, Verstynen TD, Schneider W, Tarr MJ (2013): Explicating the face perception network with white matter connectivity. *PLoS One* 8:e61611.
- Rajimehr R, Young JC, Tootell RB (2009): An anterior temporal face patch in human cortex, predicted by macaque maps. *Proc Natl Acad Sci U S A* 106:1995–2000.
- Ramon M, Vizioli L, Liu-Shuang J, Rossion B (2015): Neural microgenesis of personally familiar face recognition. *Proc Natl Acad Sci U S A* 112:E4835–E4844.
- Reinders AA, den Boer JA, Buchel C (2005): The robustness of perception. *Eur J Neurosci* 22:524–530.
- Reinders AA, Glascher J, de Jong JR, Willemse AT, den Boer JA, Buchel C (2006): Detecting fearful and neutral faces: BOLD latency differences in amygdala-hippocampal junction. *NeuroImage* 33:805–814.
- Rice GE, Watson DM, Hartley T, Andrews TJ (2014): Low-level image properties of visual objects predict patterns of neural response across category-selective regions of the ventral visual pathway. *J Neurosci* 34:8837–8844.
- Rossion B (2008a): Constraining the cortical face network by neuroimaging studies of acquired prosopagnosia. *NeuroImage* 40: 423–426.
- Rossion B (2008b): Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychol* 128:274–289.
- Rossion B (2014): Understanding face perception by means of human electrophysiology. *Trends Cogn Sci* 18:310–318.
- Rossion B, Boremanse A (2011): Robust sensitivity to facial identity in the right human occipito-temporal cortex as revealed by steady-state visual-evoked potentials. *J Vis* 11:16.
- Rossion B, Delvenne JF, Debatisse D, Goffaux V, Bruyer R, Crommelinck M, Guerit JM (1999): Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biol Psychol* 50:173–189.
- Rossion B, Caldara R, Seghier M, Schuller AM, Lazeyras F, Mayer E (2003): A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain* 126:2381–2395.
- Rossion B, Dricot L, Goebel R, Busigny T (2011): Holistic face categorization in higher order visual areas of the normal and prosopagnosic brain: Toward a non-hierarchical view of face perception. *Front Human Neurosci* 4:225.
- Rossion B, Hanseeuw B, Dricot L (2012a): Defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain Cogn* 79:138–157.
- Rossion B, Prieto EA, Boremanse A, Kuefner D, Van Belle G (2012b): A steady-state visual evoked potential approach to individual face perception: Effect of inversion, contrast-reversal and temporal dynamics. *NeuroImage* 63:1585–1600.
- Rousselet GA, Mace MJM, Fabre-Thorpe M (2003): Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J Vis* 3:440–455.
- Sadr J, Sinha P (2004): Object recognition and random image structure Evolution. *Cogn Sci* 28:259–287.
- Sayres R, Grill-Spector K (2008): Relating retinotopic and object-selective responses in human lateral occipital cortex. *J Neurophysiol* 100:249–267.
- Schiltz C, Rossion B (2006): Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage* 32: 1385–1394.
- Schwarzlose RF, Baker CI, Kanwisher N (2005): Separate face and body selectivity on the fusiform gyrus. *J Neurosci* 25: 11055–11059.
- Schwarzlose RF, Swisher JD, Dang S, Kanwisher N (2008): The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci U S A* 105:4447–4452.
- Sergent J (1986): Microgenesis of face perception. In: Ellis HD, Jeeves MA, Newcombe F, Young AM, editors. *Aspects of Face Processing*. Dordrecht: Nijhoff. pp 17–33.

- Sergent J, Ohta S, MacDonald B (1992): Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain* 115:15–36.
- Sorger B, Goebel R, Schiltz C, Rossion B (2007): Understanding the functional neuroanatomy of acquired prosopagnosia. *NeuroImage* 35:836–852.
- Steeves JK, Culham JC, Duchaine BC, Pratesi CC, Valyear KF, Schindler I, Humphrey GK, Milner AD, Goodale MA (2006): The fusiform face area is not sufficient for face recognition: Evidence from a patient with dense prosopagnosia and no occipital face area. *Neuropsychologia* 44:594–609.
- Tavor I, Yablonski M, Mezer A, Rom S, Assaf Y, Yovel G (2014): Separate parts of occipito-temporal white matter fibers are associated with recognition of faces and places. *NeuroImage* 86:123–130.
- Thomas C, Avidan G, Humphreys K, Jung KJ, Gao F, Behrmann M (2009): Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nat Neurosci* 12:29–31.
- Thompson SK, Engel SA, Olman CA (2014): Larger neural responses produce BOLD signals that begin earlier in time. *Front Neurosci* 8:159.
- Turk-Browne NB, Norman-Haignere SV, McCarthy G (2010): Face-specific resting functional connectivity between the fusiform gyrus and posterior superior temporal sulcus. *Front Human Neurosci* 4:176.
- VanRullen R (2006): On second glance: Still no high-level pop-out effect for faces. *Vis Res* 46:3017–3027.
- Vuilleumier P, Armony JL, Driver J, Dolan RJ (2001): Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron* 30:829–841.
- Weiner KS, Grill-Spector K (2010): Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage* 52:1559–1573.
- Weiner KS, Grill-Spector K (2012): The improbable simplicity of the fusiform face area. *Trends Cogn Sci* 16:251–254.
- Weiner KS, Grill-Spector K (2013): Neural representations of faces and limbs neighbor in human high-level visual cortex: Evidence for a new organization principle. *Psychol Res* 77: 74–97.
- Yin RK (1969): Looking at Upside-down Faces. *J Exp Psychol* 81: 141.
- Yovel G, Kanwisher N (2005): The neural basis of the behavioral face-inversion effect. *Curr Biol* 15:2256–2262.
- Yue X, Cassidy BS, Devaney KJ, Holt DJ, Tootell RB (2011): Lower-level stimulus features strongly influence responses in the fusiform face area. *Cereb Cortex* 21:35–47.
- Zhen Z, Yang Z, Huang L, Kong XZ, Wang X, Dang X, Huang Y, Song Y, Liu J (2015): Quantifying interindividual variability and asymmetry of face-selective regions: A probabilistic functional atlas. *NeuroImage* 113:13–25.
- Zhu Q, Zhang J, Luo YL, Dilks DD, Liu J (2011): Resting-state neural activity across face-selective cortical regions is behaviorally relevant. *J Neurosci* 31:10323–10330.