cortex 66 (2015) 69–80

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Category search speeds up face-selective fMRI responses in a non-hierarchical cortical face network

Fang Jiang [a,b], *Jeremy B. Badler* [a,c], *Giulia Righi* [d] *and Bruno Rossion* [a,*]

[a] *Psychological Sciences Research Institute and Institute of Neuroscience, University of Louvain, Belgium*
[b] *Vision and Cognition Group, University of Washington, USA*
[c] *Playable Innovative Technologies Lab, Northeastern University, USA*
[d] *Department of Psychology, University of Massachusetts at Ahmerst, USA*

ARTICLE INFO

Article history:
Received 5 April 2014
Reviewed 25 July 2014
Revised 25 October 2014
Accepted 30 January 2015
Action editor Stefan Schweinberger
Published online 21 February 2015

Keywords:
Face perception
FFA
OFA
Attention
RISE method

ABSTRACT

The human brain is extremely efficient at detecting faces in complex visual scenes, but the spatio-temporal dynamics of this remarkable ability, and how it is influenced by category-search, remain largely unknown. In the present study, human subjects were shown gradually-emerging images of faces or cars in visual scenes, while neural activity was recorded using functional magnetic resonance imaging (fMRI). Category search was manipulated by the instruction to indicate the presence of either a face or a car, in different blocks, as soon as an exemplar of the target category was detected in the visual scene. The category selectivity of most face-selective areas was enhanced when participants were instructed to report the presence of faces in gradually decreasing noise stimuli. Conversely, the same regions showed much less selectivity when participants were instructed instead to detect cars. When "face" was the target category, the fusiform face area (FFA) showed consistently earlier differentiation of face versus car stimuli than did the "occipital face area" (OFA). When "car" was the target category, only the FFA showed differentiation of face versus car stimuli. These observations provide further challenges for hierarchical models of cortical face processing and show that during gradual revealing of information, selective category-search may decrease the required amount of information, enhancing and speeding up category-selective responses in the human brain.

1. Introduction

The human brain is able to categorize a wide range of visual stimuli as faces and discriminate those from other categories in visual scenes accurately and rapidly (e.g., Crouzet, Kirchner, & Thorpe, 2010; Rousselet, Macé, & Fabre-Thorpe, 2003). To date, the spatio-temporal dynamics of this remarkable ability remain largely unknown. Multiple cortical areas that respond more to pictures of faces than other object categories have been identified ("face-selective areas", e.g., Puce, Allison, Gore, & McCarthy, 1995; Sergent, Ohta, & MacDonald,

---

* *Corresponding author*. Psychological Sciences Research Institute (IPSY), Université de Louvain, 10, Place Cardinal Mercier , 1348 Louvain-la-Neuve, Belgium.

E-mail address: bruno.rossion@uclouvain.be (B. Rossion).

http://dx.doi.org/10.1016/j.cortex.2015.01.025

1992; Rossion, Hanseeuw, & Dricot, 2012; Weiner & Grill-Spector, 2010). These areas are thought to be organized largely in a feedforward fashion, with the face-selective area in the lateral part of the inferior occipital gyrus ("occipital face area", OFA, Puce, Allison, Asgari, Gore, & McCarthy, 1996; Gauthier et al., 2000) being the first stage in the hierarchical face perception network (Fairhall & Ishai, 2007; Haxby, Hoffman, & Gobbini, 2000; Pitcher, Walsh, & Duchaine, 2011). However, neuroimaging studies on brain-damaged patients with prosopagnosia have challenged such a strict hierarchical view, showing that the middle fusiform gyrus (the "fusiform face-selective area", FFA, Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997) responds to faces despite no potential inputs from a damaged OFA (Rossion, 2014; Rossion et al., 2003; Steeves et al., 2006 for a review). In line with these observations, manipulations of stimulus and presentation times have revealed activation of the FFA in the typical human brain in the absence of, or before, face-selective responses emerging in the OFA (Goffaux et al., 2011; Jiang et al., 2011; Rossion, Dricot, Goebel, & Busigny, 2011; Rossion et al., 2012) suggesting the presence of processing pathways that bypass the OFA (Atkinson & Adolphs, 2011; Rossion, 2008; see also Fig. 5 in Pitcher et al., 2011).

Besides the FFA and OFA, the face-perception network is generally thought to include an area in the posterior superior temporal sulcus (pSTS; Puce, Allison, Bentin, Gore, & McCarthy, 1998), and there is recent evidence that this area could also be reached independently of the OFA (Pitcher, Duchaine, & Walsh, 2014). All these areas show a right hemisphere advantage (for other areas identified in the individual brain, see Rossion et al., 2012; Weiner & Grill-Spector, 2010). Additional face-selective areas include the amygdala, responding to emotional content (e.g., Vuilleumier, Armony, Driver, & Dolan, 2001); the collateral sulcus ("AT", Nasr & Tootell, 2012); the anterior inferior temporal lobe, involved in the retrieval of semantic information (Leveroni et al., 2000; Rajimehr, Young, & Tootell, 2009); the orbitofrontal cortex, responding to facial attractiveness (Ishai, 2007); the inferior frontal gyri (IFG), activated during the assessment of emotion (Nakamura et al., 1999; Phillips et al., 1997) and facial imagery tasks (Ishai, Haxby, & Ungerleider, 2002).

An important feature of the face-processing cortical network is that it is not solely influenced by visual stimuli, as top-down neural processes also play a critical role. For instance, many "visual" face-selective areas show robust activation if they are cued to expect a face (Egner, Monti, & Summerfield, 2010; Esterman & Yantis, 2010; Puri, Wojciulik, & Ranganath, 2009), as well as if a task requires the detection of faces within a mixed stimulus series (Summerfield, Egner, & Greene, et al., 2006). It is even possible to observe the neural signature of a detected face in response to noise (Righart, Andersson, Schwartz, Mayer, & Vuilleumier, 2010; Zhang et al., 2008) or nonface (Summerfield, Egner, Mangels, & Hirsch, 2006) stimuli. Furthermore, feature-based attention (Maunsell & Treue, 2006) activates a broad cortical network (Corbetta & Shulman, 2002) that has considerable overlap with face-processing areas (Palermo & Rhodes, 2007). In face-specific tasks, attention has been shown to modulate activity in the FFA (e.g., Pessoa, McKenna, Gutierrez, & Ungerleider, 2002; Summerfield, Egner, & Greene, et al., 2006; Vuilleumier et al., 2001; Wojciulik, Kanwisher, & Driver, 1998) and other face-selective areas (Ishai et al., 2002; Pessoa et al., 2002; Summerfield, Egner, & Greene, et al., 2006).

The aforementioned studies were primarily concerned with response magnitude; to our knowledge, no study has specifically examined the effect of expectation, or category-search, on the timing of activation of the different components of the face network. This presents a significant information gap, given the widespread influence of top-down processes, and the importance of inter-area response latency for determining how specific face areas fit into the overall processing stream. Thus our goal for the present study was twofold. First, we wished to assess top-down influences (searching for certain stimulus category) on areas of the face-processing network. Second, we wished to see whether the relative timing relationships between individual areas remained consistent when influenced by searching for a specific category.

A small clarification is in order. Usually, expectation is related to the degree of certainty about an upcoming stimulus; if a certain stimulus has a high probability of appearance, preparatory processes can facilitate its detection and the associated responses (Cisek & Kalaska, 2010). For instance, certainty can be manipulated by preceding the stimulus with a cue (e.g., Egner et al., 2010; Esterman & Yantis, 2010; Puri et al., 2009) or manipulating the content of a stimulus series to enhance the predictability of individual stimuli (e.g., Maljkovic & Nakayama, 1994). However, one can manipulate category-search by instructing observers to consider only one of the potential categories as the relevant detection stimulus, even if on any given trial one target category is just as likely to emerge as another. This is what we manipulated in our study, by asking participants to detect only faces or only objects (cars) in separate blocks in which exemplars of both kinds of categories were presented randomly. Note that in every trial, there is only one object category presented at a time, unlike object-based attention paradigms in which two or more objects are presented and only one is attended and/or task-relevant (e.g., Vuilleumier et al., 2001). For the second goal of assessing relative timing, we used a paradigm in which a visual stimulus gradually emerges from phase-scrambled background noise (Sadr & Sinha, 2004).

Thus, in the current paradigm (modeled after Jiang et al., 2011) either a face or a car picture emerged gradually, and human observers had to indicate when they first detected an object that matched a pre-instructed category while undergoing fMRI. We hypothesized that 1) FFA and OFA would show a strong influence of the searched category, and that, in line with our previous study (Jiang et al., 2011), 2) face-selective responses would emerge in the middle fusiform gyrus (i.e., "FFA") before the inferior occipital gyrus (i.e., "OFA").

2. Materials and methods

2.1. *Participants*

Twelve paid participants (8 females and 4 males; mean age 23, all right-handed) were included in the study. All had normal or corrected vision. Written informed consent was obtained from all participants prior to the experiment, following procedures approved by University of Maastricht where all imaging took place.

2.2. Stimuli

The stimuli used in this experiment were identical to those of our previous study (Jiang et al., 2011), and more details may be found there. Briefly, the image sequences were generated with Random Image Structure Evolution (RISE) methods (Sadr & Sinha, 2004), which 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. Based on an original image, a RISE sequence was generated by combining progressively degraded/randomized phase spectra and the original intact amplitude spectrum. The result was a sequence of images in which a recognizable object gradually evolved from randomness. Note that all the images belonging to a sequence had identical amplitude spectra and identical overall luminance/contrast (see Fig. 3 in Sadr & Sinha, 2004).

A total of 48 face and 48 car grayscale images were selected as the originals to create the stimuli for this study. Some examples are given in Fig. 1A–B. The images were visual scenes containing faces and cars that were highly variable in their appearance, size, and spatial location. Before implementing RISE, we equalized the luminance across all 96 original images. Each RISE sequence included 15 frames, ranging from 10% to 50% interpolation of the random and original phase spectra, in 14 steps of 2.86% (Fig. 1C). Note that an interpolation level of 0% corresponds to the unaltered phase spectrum of the original image, and an interpolation level of 100% corresponds to a completely random phase spectrum. In order to reduce the length of the sequences, we selected eight frames to be used in the actual experiment, indicated by the purple borders in Fig. 1C.

2.3. Design and procedure

The experiment consisted of presenting a continuous series of 8-frame RISE sequences, in which a recognizable object (face or car) gradually evolved from degraded images. Participants were asked to press a button as soon as they were sure that they detected the target object during the presentation of the sequences. Critically, for each experimental run they were only asked to respond for only one object category (i.e., either a face or a car). No button press was required for trials containing image sequences of the non-target object. Participants were informed that faces and cars could vary in size, appearance, and their spatial location within the image. Note that after the participant indicated that they detected the target object, the presentation of the sequence continued and participants were instructed to maintain attention until the end of the sequence. Participants were requested to maintain a constant level of confidence in their judgment across trials.

Each participant performed four experimental runs, including two runs with faces as the target category and two runs with cars as the target. The order of the target categories was counterbalanced between participants, with half of the participants starting with two face detection runs and half of the participants starting with two car detection runs. To minimize any item-specific effect, the assignment of faces or cars as the target category to each experimental run was also counterbalanced between participants.

Each run lasted approximately 12 min and consisted of a total of 24 trials (12 face stimulus trials and 12 car stimulus

Fig. 1 – Stimuli. Sample images of faces (A) and cars (B) used for the experiment. (C) RISE sequence. A total of 15 frames were generated, ranging from 10% to 50% interpolation between noise and target image (from top left to bottom right). Eight frames were selected for the actual experiment (purple borders).

trials). Trials were presented in a randomized order for each participant. Each trial contained a unique 20 s sequence of 8 images, each displayed for 2.5 sec. The sequence was followed by 2.5 sec of blank screen and a long fixation before the start of the next sequence. The fixation duration varied between 5000, 6250, and 7500 msec. The onset times of any two subsequent trials, therefore, were separated by an average of 8750 msec (7500–10,000 msec/6–8 TRs).

Stimuli were back projected onto a screen located above the participant's head. A PC running E-prime 1.1 (PST Inc., Sharpsburg, PA) was used to present stimuli and collect behavioral responses. All stimuli were sized to 575 × 383 pixels. The images were always presented in the center of the screen and subtended approximately a visual angle of 8.54°.

2.4. *Localizer scans*

Independent localizer scans were performed separately from the main experiment, in order 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 sec each) with 9 sec fixation intervals in between. During each block, 18 images were presented for 750 msec followed by a 250 msec blank screen. All images of faces and cars were presented in color with equalized luminance and their scrambled version was created with a Fourier phase randomization (Sadr & Sinha, 2004). Note that unlike the images of faces and cars used in the experimental scans, all those used in the localizer scans were segmented (see Jiang et al., 2011; Rossion et al., 2012 for details).

2.5. *Image acquisition*

All participants were scanned at the Maastricht Brain Imaging Center. Data were collected using a 3T Allegra head scanner (Siemens AG, Munich, Germany). Functional data in the localizer scans were obtained from 36 transverse slices with a spatial resolution of 3.5 mm × 3.5 mm × 3.5 mm (acquisition matrix 64 × 64), using a repeated single-shot echo-planar imaging sequence (TE = 50 msec, TR = 2250 msec, FA = 90°, FOV = 224 mm). Functional data in the experiment scans were obtained from 20 transverse slices, with a spatial resolution 3.5 mm × 3.5 mm × 5 mm. To have a relatively good temporal resolution, TR was set to 1250 msec 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 msec, TR = 2250 msec, FA = 9°, FOV = 256 mm). These T1-weighted images provided detailed anatomical information. A 25° angle perpendicular to the main magnetic field B_0 was used to reduce magnetic artifacts and signal dropout, allowing us to record up to the anterior inferior temporal lobe (Deichmann, Gottfried, Hutton, & Turner, 2003).

2.6. *fMRI data preprocessing*

Data were analyzed using Brain Voyager QX (Version 2.1, Brain Innovation, Maastricht, Netherlands). The first four volumes of each functional dataset were discarded to cope with T2* contrast saturation effects. Prior to statistical analysis, the functional data underwent a series of preprocessing steps, namely slice scan time correction, 3D motion correction (with realignment to the first volume), linear trend removal, and high pass filtering (removing frequencies lower than 3 cycles/session, ≈.004 Hz for experimental runs and .005 Hz for localizer runs). Both anatomical and functional data were transformed into Talairach space (Talairach & Tournoux, 1988).

2.7. *Regions of interest (ROIs)*

Areas responding preferentially to faces were defined independently for each individual participant based on the localizer scans, using the contrast [faces − cars] in conjunction with the contrast [faces − scrambled faces] (Fig. 2). The conjunction analysis ensured that the activation in face-sensitive regions was not related to low-level differences between faces and non-face object categories (Rossion et al., 2012). For each participant, all contiguous voxels in the middle fusiform gyrus and inferior/middle occipital gyrus (corresponding to FFA and OFA, respectively) in both hemispheres with a minimum significance of q(False Discovery Rate, FDR) < .001 were selected. We lowered the q(FDR) to .05 for four participants, due to the relatively smaller size of their regions.

Using the same method, we also selected two additional ROIs that we could identify on more than half of our participants, the right pSTS and the right IFG (Fig. 2). These additional face-preferential ROIs were defined using a minimum statistical threshold of q(FDR) < .05. The mean Talairach coordinates, cluster sizes and standard deviation of all face-preferential ROIs selected in this study are reported in Table 1. A total of six regions (bilateral FFA, bilateral OFA, right pSTS and right IFG) were thus considered for subsequent analysis.

2.8. *Time course analysis*

In individually defined ROIs, we examined the top-down influence of task instruction on the first appearance of a significant difference in BOLD activity between responses to face and car stimulus sequences. Note that only subjects in which a particular ROI was detectable were used for the respective analysis (see Table 1 for n-values). The onset of differential responses was separately defined for the face and car detection runs, in which participants were instructed to detect faces or cars, respectively. The method used was similar to that of our previous study (Jiang et al., 2011). Specifically, for each subject and each ROI, event-related averages were computed for relevant face and car trials. Using two-tailed pairwise *t* tests, BOLD responses to face and car sequences were then compared at group level to determine the time points at which responses became differentiable (i.e., significantly higher for one type of stimulus sequence compared to the other). Only if three successive time points showed a significant difference at the $p < .05$ level, the separation onset time was defined as the first of those points. Note that by adding temporal continuity constraints (instead of multiple comparison correction), our method of determining onset of

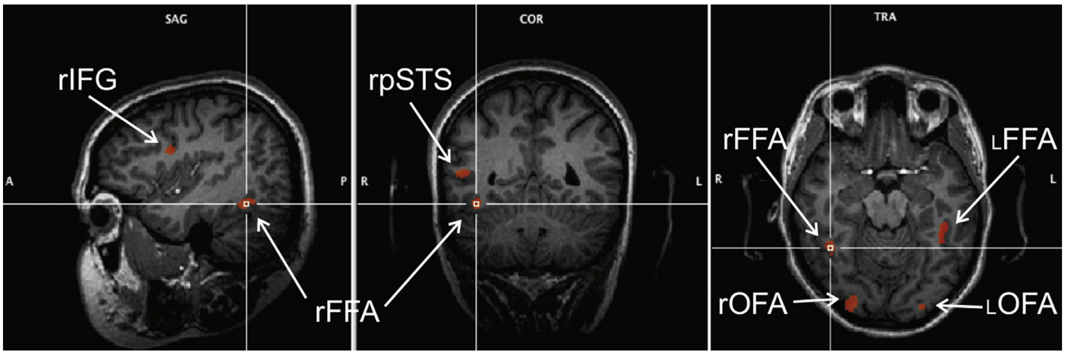

Fig. 2 — An example of individually localized regions of interest that responded preferentially to faces based on localizer scans using contrast [faces − cars] in conjunction with contrast [faces − scrambled faces]. Sagittal view shows the right FFA (cross coordinates: 40, −50, −16) and right IFG. Coronal views show the right FFA and right pSTS. Bilateral FFA and OFA are shown in transverse view. Note for this subject, a minimum significance of q(FDR) < .001 was used to localize all ROIs except the right IFG, for which q(FDR) was lowered to .05.

difference between two time courses ensured both reliability and sensitivity (Jiang et al., 2011).

To further verify the validity of the calculated separation times, a bootstrap analysis was performed. For each ROI and instruction condition, subjects were randomly selected with replacement from the pool until the original n value was reached. The BOLD curve separation time was then calculated as before. After 10000 iterations, the procedure yielded a bootstrap distribution of separation time values. The median and interquartile range of the distribution was used as a measure of the reliability and stability of the original separation calculation.

3. Results

3.1. *Behavioral results*

Participants successfully detected 91% of the car trials and 88.5% of the face trials. The level of accuracy was similar to what we previously reported (93% and 89% in Jiang et al., 2011). The accuracy difference between car trials and face trials was not significant [$F(1,11) = .79$, *ns*]. Subjects made on average slightly faster correct responses in car trials (mean RT 14.56 sec) than in face trials (mean RT 15.32 sec), but the difference did not reach significance [$F(1,11) = 3.87$, $p > .08$].

3.2. *Neuroimaging results*

We analyzed time course data in each individually localized ROI (see Table 1 for mean Talairach coordinates and number of participants in which each ROI was defined). As indicated in the Methods, the time course data were averaged separately for face and car detection runs, where different instructions about the target category were given to participants. Only correctly-identified stimuli were used for the analysis.

Shown in Fig. 3 are the time courses of fMRI activity in the right FFA, for both face detection runs (red curves) and car detection runs (blue curves). The responses to face stimuli are represented by solid curves, while car stimuli responses are dotted. The critical comparison is between the time point when the responses to face and car stimuli in a face detection task run become differentiable (vertical red line), and the time when the different stimuli responses become differentiable in a car detection task run (vertical blue line). In both ROIs, the fMRI activity rose slowly as the face and car in the visual scene gradually emerged from randomness. When participants were instructed to detect faces, fMRI activity to face sequences in the right FFA rose significantly above the activity to car sequences at 7.5 sec post-sequence onset and significant face-selective responses lasted for 18.75 sec (see Table 2 for onset and duration of face-selective responses for FFA and OFA). In contrast, for the exact same stimuli, when participants were

Table 1 — Average Talarich coordinates, cluster size, and their standard deviation of face-preferential ROIs.

ROI	#Subj	Talairach coordinates							
		Mean x	Mean y	Mean z	Mean size	stdev x	stdev y	stdev z	stdev size
Right hemisphere									
FFA	12	38	−51	−16	506	3	5.1	3.4	259
OFA	11	31	−83	−11	455	6.8	8.1	7.1	293
pSTS	8	49	−44	6	596	3.7	6.6	3.7	205
IFG	7	44	13	25	310	5.8	9.2	7.2	172
Left hemisphere									
FFA	11	−39	−49	−16	644	4.1	7	3	296
OFA	11	−32	−80	−14	411	5.5	9.7	6.2	251

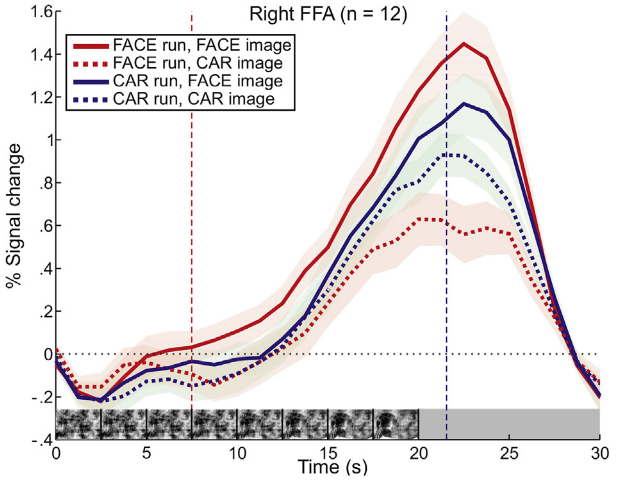

Fig. 3 – Subject-mean BOLD activity for right FFA. Color indicates the instruction (red, face detection run; blue, car detection run) and line style indicates the stimulus (solid, face image; dotted, car image). Shaded regions denote standard error of the mean. Dashed vertical lines indicate the time point where the difference between face and car stimuli became statistically significant (see Section 2.8, Time course analysis) for the indicated run. A representative stimulus sequence viewed by the subject is shown in the lower inset. Note the strong influence of category search; stimuli evoked differential activity much earlier in face detection runs than in car detection runs.

instructed to detect cars the difference emerged much later, at 22.5 sec, and lasted much shorter, i.e., for 3.75 s.

BOLD activity traces for the remaining areas are shown in Fig. 4. In the right OFA (Fig. 4C), a significant differential response occurred *only* in the face detection runs, at 13.75 sec post-sequence onset and lasted for 12.5 sec. When subjects were instructed to identify cars, whether the stimulus itself was a face or a car had no effect on the BOLD activity (the overlapping blue curves).

Thus, only in the face detection runs, both areas showed face-selective responses during the stimulation sequence. In these runs, the difference between the right FFA and OFA in the onset of face-selectivity, as evidenced previously in Jiang et al. (2011) was further supported by a two-way repeated measures analysis, in which we tested the effects of *area* (FFA versus OFA) and *time points* (in the unit of TR), and their interaction on the difference in BOLD activity between responses to face and car stimulus sequences (in 11 subjects in which the two areas could be identified). When "face" was the target category, there was a main effect of *area* [F(1,10) = 33.09, $p < .001$] and *time points* [F(24,240) = 7.37, $p < .001$], and most importantly a significant interaction between *area* and *time points* [F(24,240) = 12.13, $p < .001$]. As seen in Fig. 4, the interaction was due to a larger faces-cars differential response in the right FFA than the OFA ($p < .05$, one-tailed) at the consecutive time points 6 and 7 (7.5 sec & 8.75 sec), at which only the FFA showed a face-selective response, and then from time points 11 to 21 (13.75 sec–27.5 sec), during which face-selectivity was larger in the FFA then in the OFA ($p < .05$, one-tailed).

Table 2 – Onset and duration (in seconds) of face-selective responses for FFA and OFA.

Right hemisphere			Left hemisphere	
ROI	Onset	Duration	Onset	Duration
Face detection run				
FFA	7.5	18.75	8.75	17.5
OFA	13.75	12.5	20	7.5
Car detection run				
ROI	Onset	Duration	Onset	Duration
FFA	22.5	3.75	20	5
OFA	N/A	N/A	N/A	N/A

The left hemisphere homologues of FFA (Fig. 4B) and OFA (Fig. 4D) showed similar results to their right side counterparts (Fig. 4A and C, respectively). In left FFA, face preferential responses occurred at 8.75 sec after the onset of image sequences in the face detection runs and remained significant for 17.5 sec. In car detection runs the response difference did not reach significance until 20.0 sec. The left OFA showed face preferential responses only in the face detection runs, and even later than the right side, at 20.0 sec.

The last two ROIs examined were both on the right side, the pSTS (Fig. 4E) and IFG (Fig. 4F). The right pSTS (Fig. 4E) showed a pattern of activity somewhat different to that of the other areas. Specifically, there did not seem to be a large effect of task instruction: a differential face/car response was significant at 17.5 sec during face detection runs and at 20.0 sec during car detection runs. Although there was a hint of enhanced activity for face detection runs in the interval 11–15 sec, it was not statistically reliable. Rather, the dominant feature of pSTS was a segregation of BOLD activity according to stimulus type, with the face stimulus curves (solid) forming one cluster and the car stimulus curves (dashed) another. The face cluster also peaked at a higher level and approximately 5 sec earlier than the car cluster, although this was not quantified and will not be explored further.

Finally, the right IFG (Fig. 4F) combined features of both FFA and OFA. Like FFA, it showed early face-preferential activation during face detection runs (7.5 sec post-sequence onset). On the other hand, the preferential activation was completely absent during car detection runs, similar to what was observed in OFA. Note that some caution is needed in interpreting results in both right pSTS and IFG due to the fact that we were only able to localize them in a limited number of participants (Table 1).

Fig. 4 provides a fairly detailed area-by-area breakdown of how the latency of face-related activity was affected by task instruction. The big picture may be visualized in Fig. 5. Here the onset times of face/car differential activity are plotted, for both face and car detection runs in all areas (color-coded). The data are sorted by their bootstrap estimates (see Methods), showing the distribution medians (colored circles) as well as their respective interquartile ranges (error bars). The directly-calculated differentiation times are also shown (black squares). Note that the differentiation time was not calculable for all regions and conditions (notably, right IFG and bilateral OFA during car detection runs), nor was its 75th percentile estimate. The distribution of behavioral responses is also indicated.

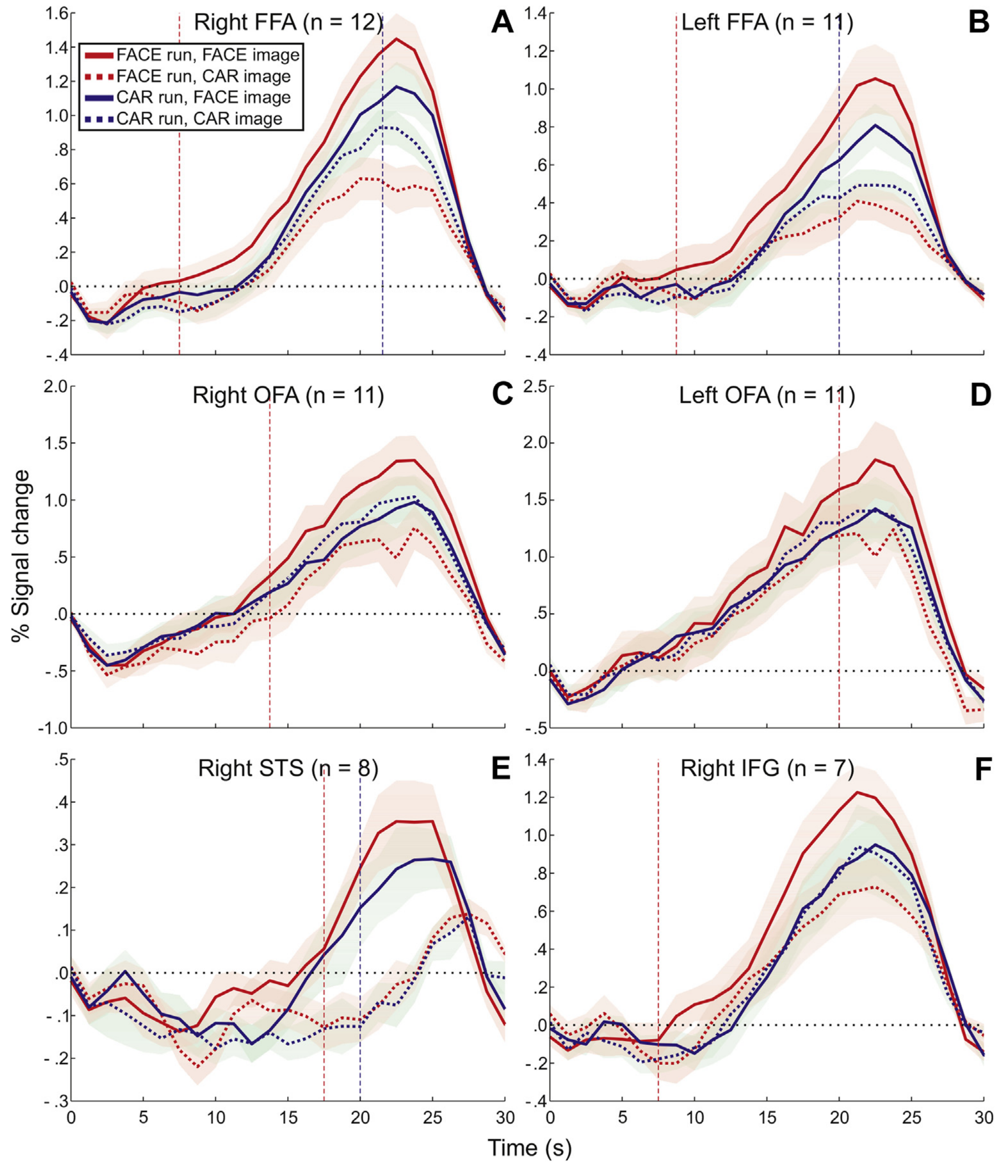

Fig. 4 – Subject-mean BOLD activity for all areas. Details as in Fig. 3. Note different y-axis values. Right FFA (A; repeated for comparison) and left FFA (B) showed a similar pattern of activity. Right (C) and left (D) OFA also showed category-based modulation, but compared to FFA it occurred consistently later. Note that there was no discernible activity difference between stimuli during car detection runs. Unlike the other areas, right pSTS (E) showed little influence of task. Finally, right IFG (F) showed a strong category influence, with early differentiation during face detection runs (like FFA) and no differentiation during car detection runs (like OFA).

There are several salient features of Fig. 5. First, note that the rank order of the calculated face/car differentiation times (black squares) agreed with that of their bootstrap estimates (circles), and always fell within the interquartile range. This supports the reliability of the calculated differentiation time for comparing between different areas and conditions. Second, every region differentiated faces from cars earlier during face detection runs (red) than during car detection runs (blue). Finally, if FFA and OFA are collapsed across hemispheres, the resulting four regions each showed a different activation profile. FFA was consistently activated the earliest and OFA consistently the latest; pSTS was intermediate but least affected by the task, and IFG was as early as FFA but critically dependent on the task. Each area had a unique temporal signature; the Discussion will examine those signatures in the context of functions of these areas.

4. Discussion

The principle result of the current study was that the face selectivity of bilateral FFA, bilateral OFA and right IFG was enhanced when participants were instructed to report the presence of faces in gradually decreasing noise stimuli. Conversely, the same regions showed much less selectivity when participants were instructed instead to detect cars. Only

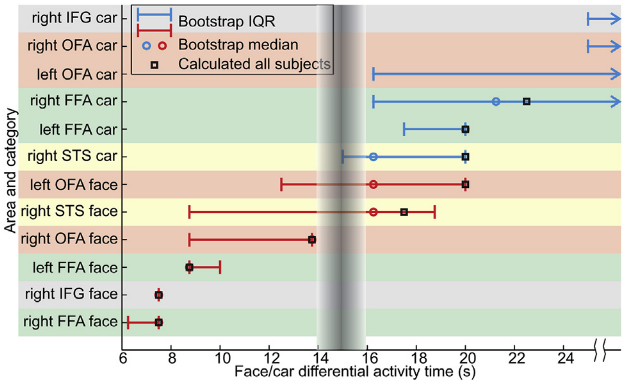

Fig. 5 – Summary of timing characteristics for all tested regions during both face (red) and car detection runs (blue). Circles with error bars indicate median and interquartile range (IQR) of bootstrap estimates of differentiation times (see text); note that these can overlap if the distribution is particularly narrow. Differentiation times calculated using standard data are plotted as black squares, and correspond to the vertical dashed lines in Figs. 3 and 4. For the car detection runs, not all areas yielded values using either method; in these cases no points are plotted (or indicated by arrowheads for the 75th percentile values). The shaded bars towards the middle of the figure indicate the distribution of behavioral response times (mean ± SD). Individual ROIs are color coded: right and left FFA, light green; right and left OFA, light red; right pSTS, yellow; and right IFG, gray.

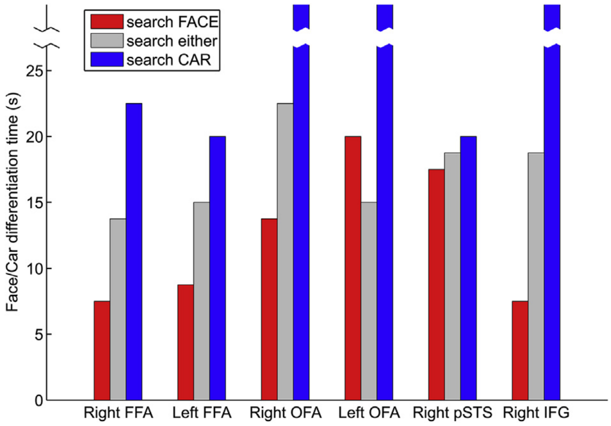

Fig. 6 – Graded effect of category search. The red (search FACE) and blue (search CAR) bars show the calculated differentiation times for face and car detection runs, respectively. The gray bars (search either) plot the differentiation times reported in Jiang et al. (2011). For the three cases where there was no detectable difference in the BOLD signals, the differentiation time was assumed to be infinite. Note that progressively increased attentiveness to the face stimulus category correspondingly reduced the latency of the differential BOLD response.

the right posterior STS showed little influence of task instruction. Regardless of the instruction, FFA showed consistently earlier differentiation of face versus car stimuli than did OFA. When the present results are compared with the earlier study of Jiang et al. (2011), an interesting trend emerges (Fig. 6). There is a graded effect of category search across our sample of the face-network: in almost all cases, the no-preferred category search task of Jiang et al. (2011) yielded differentiation times intermediate to those obtained during the "pure"-category-search conditions of the current study.

4.1. *Category attention, timing and function*

In the following paragraphs we walk through each ROI examined in the present study, in order to integrate our new results with previous findings. We begin with the FFA, which showed the earliest differentiation time. The right and left FFA will be lumped together here, since their timing and BOLD activation patterns were very similar (compare Fig. 4A and B). Although in general the right hemisphere is dominant for face processing in neuroimaging (e.g., Kanwisher et al., 1997; Puce et al., 1995; Rossion et al., 2012; Sergent et al., 1992), bilateral activation is almost always observed (e.g., Weiner & Grill-Spector, 2010).

We found the BOLD response to face stimuli became different from the response to car stimuli significantly earlier when faces were the searched category (Fig. 5). This observation suggests that category search serves to increase nonspatial attention. Numerous studies have observed attention-enhanced FFA activation in tasks where different faces were compared (e.g., Pessoa et al., 2002; Vuilleumier et al., 2001; Wojciulik et al., 1998). Summerfield, Egner, and Greene, et al. (2006) used a mixed stimulus series (faces, cars and houses) and manipulated attention by asking subjects to classify each one as a face or not ('face set'), or as a house or not ('house set'). They found elevated FFA BOLD activity for face stimuli in face sets relative to face stimuli in house sets, but did not report differential activation times. Analogous results were obtained in another study where attention was manipulated indirectly, via the expectation of an upcoming face or house stimulus (Egner et al., 2010). Two additional expectation studies did investigate activation times (Esterman & Yantis, 2010; Puri et al., 2009), and both found enhanced FFA activation prior to any discriminable visual stimulus when a face image was guaranteed to appear.

The modulation of face/car differentiation times that we observed here is consistent with the theory that category expectation or attention act by pre-emptively enhancing visually-related activity for one class of stimulus at the expense of another, in order to bias competition between competing stimuli (Desimone & Duncan, 1995). In the context of faces, the right FFA may instantiate a "template" by which faces can be rapidly recognized as holistic configurations (Rossion et al., 2011). During face detection runs, the requirement to detect only faces within scenes presumably pre-loads the face template into working memory.

As with the FFA, the OFA was also modulated by category search, similar to what has been observed in previous studies (Pessoa et al., 2002; Summerfield, Egner, & Greene, et al., 2006). However, two important differences distinguish the OFA from the FFA: first, it showed differential activity significantly later

(Fig. 5). The relative latency difference has important implications for the face-processing network as a whole, to which we will return in the next section (see also Jiang et al., 2011 for an in-depth discussion). The other key difference is that the OFA lost its face selectivity entirely during car blocks. To be precise, when the task did not explicitly require the subject to indicate the presence of a face, the BOLD activity profile in the OFA was indistinguishable for visual scenes that contained a face versus visual scenes that contained a car (see Fig. 4C and D).

At first, the apparent lack of face-specific activity in the OFA during the car-detection task seems puzzling, and warrants further discussion. First, our stimuli were natural scenes, where the face images themselves comprised a limited fraction of the total area and could be in one of many orientations or positions (see Fig. 1). By contrast, an earlier study that did find OFA face-selectivity in low attention blocks (Summerfield, Egner, & Greene, et al., 2006) used full frontal images of faces. Other studies (e.g., Crouzet et al., 2010; Jiang et al., 2011) that used visual scenes also used face-related tasks that implicitly focused attention on faces. One possibility is that participants may expect cars at different locations than faces.

Finally, in the present study the gradual descrambling of the visual scenes via the RISE algorithm promoted a slow buildup of BOLD activity. While suitable for distinguishing inter-area timing differences (Carlson, Grol, & Verstraten, 2006), the technique is less suited to detecting differences in maximal response magnitude. Furthermore, the detection task itself did not require the decoding of individual face parts (as an identity discrimination task would). Thus it appears that the OFA visual response to faces was largely minimized by the paradigm leaving only the attentional response, which in turn was suppressed when an irrelevant category (cars) was the focus of the task (Peelen, Fei-Fei, & Kastner, 2009).

Of the three core face areas, pSTS was the only one where substantial category modulation was *not* observed (Fig. 4E). The area showed a reliable BOLD activity difference between face and car stimuli, with a latency intermediate to that of the FFA and OFA. However, the onset of the difference was relatively unaffected by the category-related focus of the task. Other studies have shown attention-related activity in STS, for both face (Pessoa et al., 2002) and nonface (e.g., Hopfinger, Buonocore, & Mangun, 2000) tasks. Expectation-related activity has also been reported (Esterman & Yantis, 2010).

It should be noted that the absence of significant modulation does not imply the total absence of modulation. Upon closer inspection, STS activation has been observed for spatial attention shifts (Corbetta et al., 1998) though not for nonspatial tasks such as color detection (Giesbrecht, Woldorff, Song, & Mangun, 2003). It appears to be more closely linked with the ventral network specialized for salience-driven (bottom-up) attention, rather than the dorsal network controlling goal-directed (top-down) processes (Corbetta & Shulman, 2002). Assuming our paradigm manipulated top-down attention by changing the stimulus category focus, it may not have been optimal for resolving subtle effects particular to the STS. The gradual-onset, visual scene stimuli might also have contributed to a low-magnitude effect, as discussed previously for the OFA.

The final area of interest in the present study was the IFG. Its BOLD activity profile was somewhat of a hybrid between the FFA and OFA (Fig. 4F). Like the FFA, it showed very early differential activity during face detection runs, but like the OFA, the difference completely disappeared during car detection runs. Like much of frontal cortex, the IFG has been implicated in a wide range of executive functions outside of the face-processing domain, including attentional control (Hampshire, Chamberlain, Monti, Duncan, & Owen, 2010). Both top-down (Giesbrecht et al., 2003) and bottom-up (Corbetta, Kincade, & Shulman, 2002) modulation have been observed. For face tasks, attention-related activity in the IFG is often reported in the context of emotion processing (Nakamura et al., 1999; Pessoa et al., 2002; Phillips et al., 1997) or imagery (Ishai et al., 2002). The role of the IFG in emotion is particularly relevant for the present study, since several faces in the stimulus set were smiling or had non-neutral expressions (see Fig. 1). Thus, the early IFG selectivity for faces during face detection runs might represent a positive interaction between attention and emotion, in line with the view that attention is *necessary* for processing the emotional content of faces (Pessoa et al., 2002; but see Palermo & Rhodes, 2007 for additional discussion).

4.2. *Implications for the face-processing network*

The preceding discussion highlighted the different ways in which category-search, as manipulated by task instruction, affected different components of the face-processing network. Let us now reexamine the network as a whole.

The feedforward hierarchical view of face processing (e.g., Fairhall & Ishai, 2007; Haxby et al., 2000; Pitcher et al., 2011) states that information from early visual cortex first enters the network at the level of the OFA, a region that is absolutely critical for face perception as indicated by its common association with prosopagnosia (Bouvier & Engel, 2006; Rossion et al., 2003) and face recognition impairments following electrical intracerebral stimulation (Jonas et al., 2012, 2014). From the three core areas, the information then proceeds to the extended network where emotional information and other higher attributes are processed. The OFA is well placed as an entry node, receiving more direct projections from the visual regions around the calcarine sulcus than either FFA or pSTS (Gshwind, Pourtois, Schwartz, Van De Ville, & Vuilleumier, 2012; Pyles, Verstynen, Schneider, & Tarr, 2013). It also shows early selectivity, with an intra-cortical recording study suggesting that the right OFA plays a major role in generating the N170 face-sensitive response (Jonas et al., 2012; see also Sadeh et al., 2011). TMS studies suggest an even earlier role of this region in face processing, since face discrimination, but not face detection, is disrupted when successive transcranial magnetic pulses are sent at 60 msec and 100 msec (i.e., indicating a role of the region after the second pulse if these pulses add up; Pitcher, Walsh, Yovel, & Duchaine, 2007; Solomon-Harris, Mullin, & Steeves, 2013).

Nonetheless, there are several lines of evidence that suggest alternative pathways exist for feedforward information flow through the cortical face network. First, prosopagnosic patients such as PS and DF have extensive lesions of inferior occipital cortex, including right OFA. Despite this fact, they both show FFA activation (Rossion et al., 2003; Steeves et al., 2006; for PS and DF respectively), an observation that is difficult to explain if the OFA were a critical relay node (Atkinson & Adolphs, 2011; Rossion, 2008, 2014). Supporting this view, a recent study showed that TMS applied to the OFA did not impair a difficult face categorization task (Solomon-Harris et al., 2013).

A second line of evidence derives from fMRI studies of the normal brain, including the present one. In a face localizer task, FFA showed an earlier onset time of BOLD activation than OFA (Rossion et al., 2012). Jiang et al. (2011) found consistently earlier differential FFA activation between faces and cars when there was no prior information or instruction that depended on stimulus category (see Fig. 6). In another study, Mooney or Arcimboldo stimuli, which can be perceived as faces only through their global structure, activated primarily the right FFA and pSTS without face-selective responses in the OFA (Rossion et al., 2011). Goffaux et al. (2011) presented face photographs contained only coarse information (low spatial frequencies, LSF), or more details (middle and high spatial frequencies, MSF and HSF, respectively). These stimuli were flashed for various durations (75, 150 or 300 msec) and then masked. At very short time durations (75 msec), LSF evoked larger responses in the (right) FFA than did HSF and MSF faces, and the pattern reversed for longer durations. However, the rOFA did not show any face-selective response (faces versus scrambled faces) to LSF faces at any exposure duration. This latter observation is in line with the findings described above, showing first that face-selectivity can emerge in the FFA without such effects in the OFA, and second that the OFA shows face-selectivity only when face-likeliness can be extracted from local information.

Note that in the present study, as well as in the previous study of Jiang et al. (2011), the nature of the information driving the earlier face-selective response in the FFA than the OFA remains unknown. There is evidence that face-selectivity in these regions is driven partly by category-related low-level visual information such as power spectrum differences (Andrews, Davies-Thompson, Kingstone, & Young, 2010; Rossion et al., 2012) or other cues such as contrast or size (see Yue, Cassidy, Devaney, Holt, & Tootell, 2011), and that such information play a role in rapid face detection in natural scenes for instance (Honey, Kirchner, & VanRullen, 2008). Here, face-selective areas are defined by a comparison that eliminates the contribution of differences in power spectrum (Rossion et al., 2012; see Methods). Moreover, the grayscale visual scenes were equalized in power spectrum, minimizing the potential contribution of low-level visual information to the emergence of face-selectivity during the dynamic displays. Nevertheless, different categories and categories of environments exhibit different orientations and spatial frequency distributions (Torralba & Oliva, 2003), with pictures of cars containing more horizontal lines as compared to faces in general. This diagnostic information could be potentially used by the cortical face network to discriminate faces from cars in visual scenes, even though the differences between object categories in our stimulus set are certainly not as large as when comparing segmented faces to cars or other objects, as in ours and most face localizer paradigms. Most importantly, this factor is unlikely to account for earlier face-selectivity in the FFA than the OFA.

Another important point concerning the interpretation of the earlier selectivity in the FFA than the OFA during the face task is the following. One may argue that our results are due to the FFA being simply more resistant to image noise than the OFA 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. In other words, the differential timing observed here may reflect better detection of faces at a higher level of scrambling in the FFA relative to the OFA. As discussed previously (Jiang et al., 2011), this is a plausible account of our observations, which we believe to be in line with our claims and to reflect of the phenomenon that we attempt to measure: in the visual world, information is not presented in random order, but accumulates over time, albeit too rapidly to capture timing differences between areas with a low temporal resolution technique such as fMRI. Here, when we objectively disrupt structured visual information (phase scrambling) and present the stimulus according to an artificially slow increasing order of visibility (Carlson et al., 2006), face-selectivity emerges at a higher level of scrambling, that is, earlier, in the FFA than in the OFA. Compared to previous studies that reported FFA activation without OFA (e.g., Rossion et al., 2011), the advantage of the present approach is that both areas are activated by the stimuli, at least here during the face task.

Finally, the present study showed that earlier FFA activation was maintained when search was limited to a single category. Most strikingly, when attention was *not* directed specifically towards faces, there was no difference whatsoever between OFA BOLD activity for face and car stimuli (see the blue curves in Fig. 4C and D). This suggests that the OFA does not passively, automatically distinguish faces (or face parts) from other stimuli if not required to by task demands. Rather than acting as a foundation without which further information integration is impossible, it is likely that the OFA functions as part of a holistic face processing network (Andrews et al., 2010; Harris & Aguirre, 2008; Rossion, 2008, 2013, 2014; Rossion et al., 2011; Schiltz & Rossion, 2006), where multiple components act on visual information in parallel according to their functional specializations.

5. Conclusions

By manipulating the category-search conditions of a task and using a gradual stimulus onset paradigm, we were able to examine the responses and timing for several different loci with the face-processing network. We confirmed and extended the critical role of selective category-search in determining neural activity, quantified the effects of this category-search versus visual stimuli, and developed an approximate sequence of area activations. The results suggest both similarities and differences between individual face-processing areas. Together, the areas form a synergy that allows humans to function efficiently and effectively in a highly social world.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

The authors wish to thank Dan Nemrodov, David Pitcher and two anonymous reviewers for helpful comments on an earlier version of the manuscript. This research was supported by an FRSM 3.4601.12 (Belgian National Fund for Scientific Research) grant and an ARC grant 13/18-053.

REFERENCES

Andrews, T. J., Davies-Thompson, J., Kingstone, A., & Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *The Journal of Neuroscience, 30*(9), 3544–3552.

Atkinson, A. P., & Adolphs, R. (2011). The neuropsychology of face perception: beyond simple dissociations and functional selectivity. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1571), 1726–1738.

Bouvier, S. E., & Engel, S. A. (2006). Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cerebral Cortex, 16*(2), 183–191.

Carlson, T., Grol, M. J., & Verstraten, F. A. (2006). Dynamics of visual recognition revealed by fMRI. *NeuroImage, 32*(2), 892–905.

Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience, 33*, 269–298.

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron, 21*(4), 761–773.

Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience, 14*(3), 508–523.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 215–229.

Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *Journal of Vision, 10*(4), 16.

Deichmann, R., Gottfried, J. A., Hutton, C., & Turner, R. (2003). Optimized EPI for fMRI studies of the orbitofrontal cortex. *NeuroImage, 19*(2), 430–441.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*(1), 193–222.

Egner, T., Monti, J. M., & Summerfield, C. (2010). Expectation and surprise determine neural population responses in the ventral visual stream. *The Journal of Neuroscience, 30*(49), 16601–16608.

Esterman, M., & Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex, 20*, 1245–1253.

Fairhall, S. L., & Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex, 17*(10), 2400–2406.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience, 12*(3), 495–504.

Giesbrecht, B., Woldorff, M. G., Song, A. W., & Mangun, G. R. (2003). Neural mechanisms of top-down control during spatial and feature attention. *NeuroImage, 19*(3), 496–512.

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. *Cerebral Cortex, 21*(2), 467–476.

Gshwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2012). White-matter connectivity between face-responsive regions in the human brain. *Cerebral Cortex, 22*(7), 1564–1576.

Hampshire, A., Chamberlain, S. R., Monti, M. M., Duncan, J., & Owen, A. M. (2010). The role of the right inferior frontal gyrus: inhibition and attentional control. *NeuroImage, 50*(3), 1313–1319.

Harris, A., & Aguirre, G. K. (2008). The representation of parts and wholes in face-selective cortex. *Journal of Cognitive Neuroscience, 20*(5), 863–878.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*(6), 223–233.

Honey, C., Kirchner, H., & VanRullen, R. (2008). Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *Journal of Vision, 8*(12). http://dx.doi.org/10.1167/8.12.9, 9.1–13.

Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience, 3*(3), 284–291.

Ishai, A. (2007). Sex, beauty and the orbitofrontal cortex. *International Journal of Psychophysiology, 63*(2), 181–185.

Ishai, A., Haxby, J. V., & Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *NeuroImage, 17*(4), 1729–1741.

Jiang, F., Dricot, L., Weber, J., Righi, G., Tarr, M. J., Goebel, R., et al. (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. *Journal of Neurophysiology, 106*, 2720–2736.

Jonas, J., Descoins, M., Koessler, L., Colnat-Coulbois, S., Sauvee, M., Guye, M., et al. (2012). Focal electrical intracerebral stimulation of a face-sensitive area causes transient prosopagnosia. *Neuroscience, 222*, 281–288.

Jonas, J., Rossion, B., Krieg, J., Koessler, L., Colnat-Coulbois, S., Vignal, J.-P., et al. (2014). Intracerebral electrical stimulation of a face-selective area in the right occipito-temporal cortex impairs individual face discrimination. *NeuroImage, 99*, 487–497.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17*(11), 4302–4311.

Leveroni, C. L., Seidenberg, M., Mayer, A. R., Mead, L. A., Binder, J. R., & Rao, S. M. (2000). Neural systems underlying the recognition of familiar and newly learned faces. *The Journal of Neuroscience, 20*(2), 878–886.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & cognition, 22*(6), 657–672.

Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends in Neurosciences, 29*(6), 317–322.

McCarthy, G., Puce, A., Gore, J., & Allison, T. (1997). Face-specific processing in the fusiform gyrus. *Journal of Cognitive Neuroscience, 9*, 605–610.

Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., et al. (1999). Activation of the right inferior frontal cortex during assessment of facial emotion. *Journal of Neurophysiology, 82*(3), 1610–1614.

Nasr, S., & Tootell, R. B. H. (2012). Role of fusiform and anterior temporal cortical areas in facial recognition. *NeuroImage, 63*(3), 1743–1753.

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia, 45*, 75–92.

Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature, 460*(7251), 94–97.

Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences, 99*(17), 11458–11463.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature, 389*(6650), 495–498.

Pitcher, D., Duchaine, B., & Walsh, V. (2014). Combined TMS and fMRI reveals dissociable cortical pathways for dynamic and static face perception. *Current Biology, 24*(17), 2066–2070.

Pitcher, D., Walsh, V., & Duchaine, B. C. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research, 209*, 481–493.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. C. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology, 17*(18), 1568–1573.

Puce, A., Allison, T., Asgari, M., Gore, J. C., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and texture: a functional magnetic resonance imaging study. *Journal of Neuroscience, 16*, 5205–5215.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience, 18*(6), 2188–2199.

Puce, A., Allison, T., Gore, J. C., & McCarthy, G. (1995). Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology, 74*, 1192–1199.

Puri, A. M., Wojciulik, E., & Ranganath, C. (2009). Category expectation modulates baseline and stimulus-evoked activity in human inferotemporal cortex. *Brain Research, 1301*, 89–99.

Pyles, J. A., Verstynen, T. D., Schneider, W., & Tarr, M. J. (2013). Explicating the face perception network with white matter connectivity. *PLoS One, 8*(4), e61611.

Rajimehr, R., Young, J. C., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences, 106*, 1995–2000.

Righart, R., Andersson, F., Schwartz, S., Mayer, E., & Vuilleumier, P. (2010). Top-down activation of fusiform cortex without seeing faces in prosopagnosia. *Cerebral Cortex, 20*, 1878–1890.

Rossion, B. (2008). Constraining the cortical face network by neuroimaging studies of acquired prosopagnosia. *NeuroImage, 40*, 423–426.

Rossion, B. (2013). The composite face illusion: a whole window into our understanding of holistic face perception. *Visual Cognition, 21*(2), 139–253.

Rossion, B. (2014). Understanding face perception by means of prosopagnosia and neuroimaging. *Frontiers in Bioscience (Elite Ed.), 6*, 308–317.

Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain, 126*(11), 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. *Frontiers in Human Neuroscience, 4*, 225.

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: a large-scale factorial fMRI face localizer analysis. *Brain Cognition, 79*(2), 138–157.

Rousselet, G. A., Macé, M. J., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *Journal of Vision, 3*(6), 440–455.

Sadeh, B., Pitcher, D., Brandman, T., Eisen, A., Thaler, A., & Yovel, G. (2011). Stimulation of object-category selective areas modulates ERP to their preferred categories. *Current Biology, 21*(22), 1894–1899.

Sadr, J., & Sinha, P. (2004). Object recognition and random image structure evolution. *Cognitive Science, 28*(2), 259–287.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage, 32*(3), 1385–1394.

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing: a positron emission tomography study. *Brain, 115*(1), 15–36.

Solomon-Harris, L. M., Mullin, C. R., & Steeves, J. K. (2013). TMS to the "occipital face area" affects recognition but not categorization of faces. *Brain Cognition, 83*(3), 245–251.

Steeves, J. K. E., Culham, J. C., Duchaine, B. C., Pratesi, C. C., Valyear, K. F., Schindler, I., et al. (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.

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science, 314*(5803), 1311–1314.

Summerfield, C., Egner, T., Mangels, J., & Hirsch, J. (2006). Mistaking a house for a face: neural correlates of misperception in healthy humans. *Cerebral Cortex, 16*, 500–508.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: An approach to cerebral imaging*. New York: Thieme Medical Publishers.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network, 14*, 391–412.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: an event-related fMRI study. *Neuron, 30*(3), 829–841.

Weiner, K. S., & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage, 52*(4), 1559–1573.

Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI study. *Journal of Neurophysiology, 79*, 1574–1578.

Yue, X., Cassidy, B. S., Devaney, K. J., Holt, D. J., & Tootell, R. B. H. (2011). Lower-level stimulus features strongly influence responses in the fusiform face area. *Cerebral Cortex, 21*(1), 35–47.

Zhang, H., Liu, J., Huber, D. E., Rieth, C. A., Tian, J., & Lee, K. (2008). Detecting faces in pure noise images: a functional MRI study on top-down perception. *Neuroreport, 19*, 229–233.