Brain Structure and Function
https://doi.org/10.1007/s00429-018-1630-4

ORIGINAL ARTICLE

Fast periodic stimulation (FPS): a highly effective approach in fMRI brain mapping

Xiaoqing Gao[1] · Francesco Gentile[1,2] · Bruno Rossion[1,3]

Received: 17 November 2017 / Accepted: 14 February 2018

Abstract

Defining the neural basis of perceptual categorization in a rapidly changing natural environment with low-temporal resolution methods such as functional magnetic resonance imaging (fMRI) is challenging. Here, we present a novel fast periodic stimulation (FPS)-fMRI approach to define face-selective brain regions with natural images. Human observers are presented with a dynamic stream of widely variable natural object images alternating at a fast rate (6 images/s). Every 9 s, a short burst of variable face images contrasting with object images in pairs induces an objective face-selective neural response at 0.111 Hz. A model-free Fourier analysis achieves a twofold increase in signal-to-noise ratio compared to a conventional block-design approach with identical stimuli and scanning duration, allowing to derive a comprehensive map of face-selective areas in the ventral occipito-temporal cortex, including the anterior temporal lobe (ATL), in all individual brains. Critically, periodicity of the desired category contrast and random variability among widely diverse images effectively eliminates the contribution of low-level visual cues, and lead to the highest values (80–90%) of test–retest reliability in the spatial activation map yet reported in imaging higher level visual functions. FPS-fMRI opens a new avenue for understanding brain function with low-temporal resolution methods.

Keywords fMRI · Brain mapping · Frequency tagging · Visual categorization · Face

Introduction

A fundamental goal of neuroscience is to build a comprehensive map of the human brain, i.e., to define the structure and function of each of its regions and networks (Brodmann 1909; Amunts and Zilles 2015; Glasser et al. 2016). The advent of the non-invasive, spatially resolved functional magnetic resonance imaging (fMRI) technique in the early 1990s (Ogawa et al. 1990, 1992) has provided an unprecedented opportunity to reach this goal, and this technique has now become a major player in Systems and Cognitive Neuroscience. Starting with visual perception, the dominant modality in primates, occupying a substantial fraction of the cortex, researchers have used fMRI to define retinotopic maps (Sereno et al. 1995; Engel et al. 1997; Wandell and Winawer 2011) as well as areas specifically involved in processing low-level visual attributes such as color or motion (e.g., McKeefry and Zeki 1997; Tootell et al. 1995; Winawer and Withoft 2015). More recently, this approach has been extended to build maps of higher level areas responding differentially to different categories of the visual world, such as faces, places and body parts (as reviewed by Grill-Spector and Weiner 2014), and to decode other categories as distributed patterns of variable neural activity across smaller brain volumes (i.e., voxels; e.g., Haxby et al. 2001; Kriegeskorte et al. 2007; Huth et al. 2012).

With a low-temporal resolution method such as fMRI measuring neural activity indirectly (i.e., the blood oxygenation level-dependent, BOLD, response), identifying category-selective brain regions requires presenting visual

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00429-018-1630-4) contains supplementary material, which is available to authorized users.

✉ Bruno Rossion
bruno.rossion@univ-lorraine.fr

[1] Psychological Sciences Research Institute (IPSY), Institute of Neuroscience (IoNS), University of Louvain, 10, Place Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium

[2] Department of Cognitive Neuroscience, Maastricht Brain Imaging Center (M-BIC), Maastricht University, 6211 LK Maastricht, The Netherlands

[3] Université de Lorraine - CHRU-Nancy, CNRS, CRAN, Service de Neurologie, F-54000 Nancy, France

Published online: 03 March 2018

stimuli belonging to different categories at a relatively slow rate, i.e., separated by several seconds, to isolate neural activity to each category. Most often though, stimuli from the same condition/category are presented consecutively for 10–20 s, i.e., a block design, which typically provides the largest BOLD response and contrast between conditions (Dale 1999). The estimated BOLD response during the whole block with respect to a baseline measure (activity to a uniform visual field, or the average activity across all blocks of stimuli) is then considered as reflecting the brain's response to this category (e.g., Aguirre and D'Esposito 1999; Weiner and Grill-Spector 2010). Then, by subtracting neural responses to different categories from one another, category-selective maps, either of regions or patterns of voxels, can be identified (Grill-Spector and Weiner 2014; Kanwisher 2017).

Although this category-selective localizer approach provides important information regarding human brain cartography, it also relies on a number of unwarranted assumptions, such as pure insertion (i.e., that adding a process to a set of cognitive processes does not affect these latter processes, Friston et al. 1996; D'Esposito 2010) and uniformity of the modeled hemodynamic response function (HRF) for different brain regions (Boynton et al. 1996; Buxton et al. 2004). Most importantly, this approach does not consider three key aspects of perceptual categorization when designing stimulation parameters. First, perceptual categorization processes can occur at a high speed in a continuous or quasi-continuous stimulation mode in the visual world (Potter 2012; Retter and Rossion 2016; Thorpe et al. 1996). This high speed implies that high-level visual areas could be stimulated at faster rates than traditionally used in fMRI designs (Gentile and Rossion 2014), with an optimal stimulation rate taking into consideration both the minimal stimulus duration to elicit a full categorization process and the duration of this process for setting stimulus onset asynchrony (see Retter and Rossion 2016). Second, block designs in fMRI do not take into account category-specific adaptation (e.g., Kovacs 2005) that occurs when exemplars of the same category are presented consecutively. Rather, within a block, direct comparisons (i.e., contrasts) between categories should be measured to increase sensitivity. The third and final aspect concerns the contribution of low-level visual cues such as differences in global contrast or amplitude spectrum to perceptual categorization (VanRullen 2006; Crouzet and Thorpe 2011; Andrews et al. 2015). Studies measuring neural responses of perceptual categorization often ignore this issue entirely (e.g., see the review of Berman et al. 2010 on face localizers), or take two extreme positions: on the one hand, using ecological stimuli that lead to large but partially unspecific differential responses (e.g., the "dynamic face localizer approach", Fox et al. 2009) or, on the other hand, using stimuli from different categories normalized for low-level visual cues at the expense of ecological validity (e.g., grayscale segmented full frontal face and house stimuli equalized for power spectra, e.g., Rousselet et al. 2008). These unwarranted assumptions and suboptimal stimulation parameters should add extra noise to the existing physiological, thermal, and scanner noises (Kruger and Glover 2001), lowering the signal-to-noise ratio (SNR) of fMRI measurement. As a result, extensive trial averaging is often required to achieve effective signal detection (Murphy et al. 2007) and category-selective responses may be biased by low-level image statistics (Andrews et al. 2015). For these reasons, fMRI studies in cognitive neuroscience may sometimes suffer from low sensitivity, specificity and reliability (Bennet and Miller 2010).

Here, we take into account these factors to introduce an effective approach to localize category-selective neural responses in the human brain. We use this approach as a model to measure brain function with fMRI or low-temporal resolution methods in general. In this approach, visual stimuli are presented at a *fast* presentation rate (6 images/s, 6 Hz) allowing one fixation by image—largely sufficient for perceptual categorization (Potter 2012; Retter and Rossion 2016; Thorpe et al. 1996)—throughout the entire recording of neural activity (Fig. 1a; Supplemental Movie S1). This dynamic stimulation sets a high baseline level of activity in low-level visual areas as well as in non-category-selective high-level visual areas. Then, we introduce transient switches from non-target object categories to a target category, here faces (Fig. 1a, red bins, and Fig. 1b for example images). In populations of neurons responding selectively to faces, such transient switches elicit differential neural responses that directly reflect category selectivity because they contrast with the continuous stimulation stream of non-face objects. Hence, contrast is maximized, and inference regarding category selectivity can be made without post hoc subtraction. The transient switches to faces are grouped in short bursts containing multiple faces to increase the duration of the BOLD response and consequently improve SNR in fMRI. Critically, within a burst, faces appear only every two stimuli, i.e., alternating with a randomly selected object (Fig. 1b). This way, category-specific adaptation is reduced during the burst, multiple contrasts are measured within a burst, and the temporal separation between two faces is more than 300 ms, leaving sufficient time for occurrence of the bulk of a face-selective neural response (Retter and Rossion 2016). Since the bursts appear at a fixed frequency during a run, the magnitude of the differential neural responses can be measured without HRF modeling, i.e., as the Fourier amplitude of the frequency of the bursts (e.g., Bandettini et al. 1993; Engel et al. 1997; Puce et al. 1995). This model-free approach, therefore, allows fair comparisons across brain regions and individuals and has high SNR, since it is only affected by noise occurring at the exact same frequency

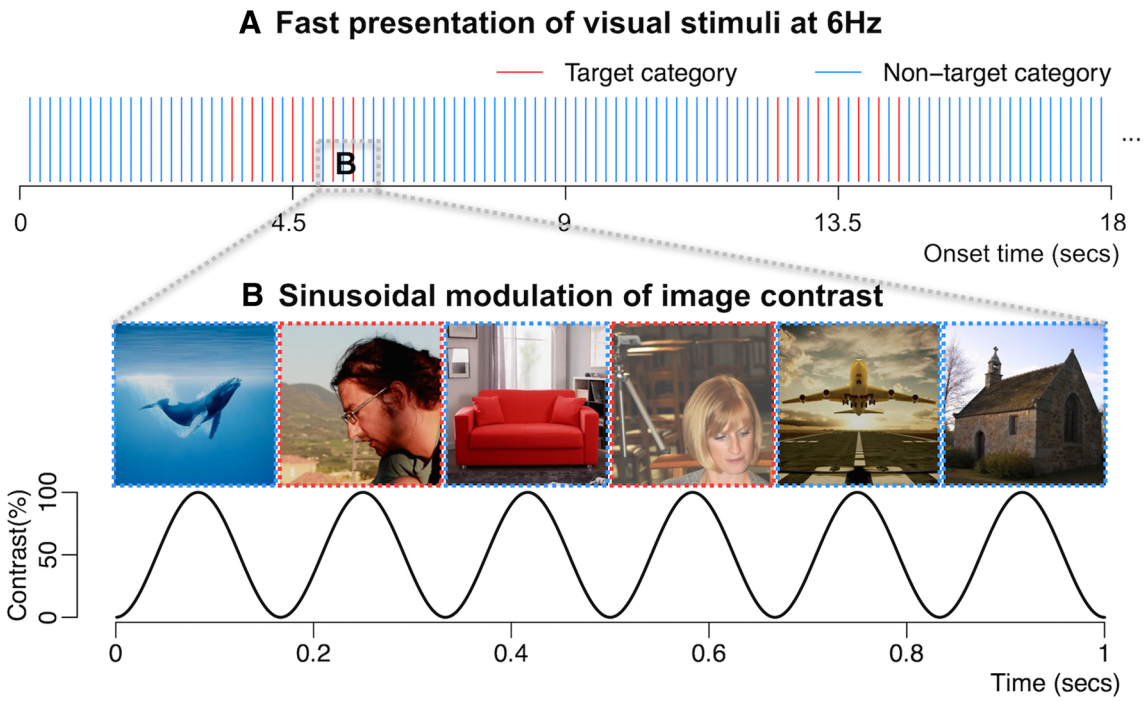

Fig. 1 The fast periodic stimulation (FPS)-fMRI paradigm. **a** Stimuli from non-target categories alternate at a rapid rate (6 Hz, blue bins, see Supplemental Movie S1). Every 9 s (i.e., 54 stimuli), a "burst" of stimuli from the target category (faces, red bins) is presented for 2.167 s. This face burst contains seven faces alternating with six non-face objects to form direct contrasts between faces and objects. Only sections of the sequences are shown in the figure. **b** Example face and object alternations (see Supplemental Figure S1 for more example images) and image contrast modulation (0–100%)

of the bursts, not broadband frequency noise (Regan 1989). Finally, a wide variety of natural images, which have complex statistical properties (Simoncelli and Olshausen 2001), ensure that the succeeding images represent many different types of low-level contrasts, minimizing the contribution of specific low-level visual cues to category-selective responses occurring at the same periodic frequency (Rossion et al. 2015). At the same time, widely variable images of faces are used to ensure that a category-selective response is not tied to specific exemplars (i.e., is generalized). Based on these unique features, we name the current approach as a fast periodic stimulation (FPS)-fMRI paradigm.

We designed a functional "face localizer" based on the above-mentioned principles. After Kanwisher and colleagues (1997), face localizers are arguably the most widely used approach to define category selectivity in neuroimaging, both in human and non-human primates (Kanwisher 2017; Tsao et al. 2008). In humans, the representation of faces differs from other object categories at the level of a large number of distributed regions, or functional clusters, which have been reported by numerous studies primarily not only in the ventral occipito-temporal cortex (VOTC), but also in the superior temporal sulcus (STS), more recently in the ventral anterior temporal lobe (VATL), and to a lesser extent in the parietal and frontal lobes (e.g., Haxby et al. 2000; Duchaine and Yovel 2015; Rossion et al. 2012; Zhen et al. 2015; Collins et al. 2016). Therefore, it provides an excellent model to assess the validity of the FPS-fMRI approach. For comparison, we also ran a conventional face localizer based on a block-design with the exact same stimuli and same duration of scanning.

Our specific goals were to test whether the FPS-fMRI face localizer can (1) identify the well-known core face-selective areas and their right hemispheric dominance at both the group level and the individual level; (2) achieve higher sensitivity (SNR) in detecting face-selective neural response than the conventional approach; (3) achieve higher specificity in localizing face-selective brain areas than a conventional approach, i.e., better isolating brain activity directly related to the high-level perceptual categorization process rather than low-level visual confounds or general factors such as attention; (4) achieve higher test–retest reliability in mapping face-selective brain areas than a conventional approach.

Materials and methods

Compliance with ethical standards

All procedures performed in this study involving human participants were in accordance with the ethical standards

Fig. 2 Example stimuli. Both faces and objects contain a wide range of variation in the composition, color, and lighting. Such high variability reduces the contribution of specific low-level visual cues to perceptual categorization while preserving naturalness of the images (Rossion et al. 2015). No amplitude spectrum equalizing was performed on these images. Fourier phase-scrambled versions of the images (below each face and object image) were also created. The right-most column contains pixel-wise averaged images of all the faces, all the objects, and all the phase scrambled images within each category. No recognizable structure is seen in the average images

of York University (Canada) Research Ethics Board (Certificate #: e2014-155) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. We obtained informed written consent from all the participants prior to the experimental sessions and they received $50 (Canadian) for their participation in the study. This study was funded by European Research Council Grant facessvep 284025 to BR, and an UCL/Marie Curie postdoctoral fellowship to XG. The authors declare that they have no conflict of interest.

Participants

Twelve adults (8 females, mean age = 30.1 ± 5.4 years, age range 24–42 years) from the York University (Canada) community participated in the fMRI experiment. All of the participants had normal or corrected-to-normal vision, and were right-handed (Oldfield 1971). None of the participants reported any history of psychiatric or neurological disorders, or current use of any psychoactive medications.

Stimuli

The stimuli consisted of 100 face images and 200 non-face images (e.g., Fig. 2, the whole set of face and non-face images can be obtained upon request). The face images were digital photographs of 100 different individuals who were non-famous relatives, friends and colleagues of the researchers of the Face Categorization Lab of the University of Louvain, Belgium. Therefore, they were unfamiliar to the participants. Each photograph contained one human face. The photographs were originally taken for personal purposes, and were given to the researchers with a completed consent form to use these photographs for research purposes and display them. They contain a natural range of variation in size, pose, and expression of the faces depicted in the photographs and in lighting and background. The non-face images consist of 200 photographs of scenes, objects, and animals. As in the face images, the non-face images also contain a natural range of variation in the composition and lighting of the images. The face images have a mean gray-scale intensity value of 115.0 ± 1 and a mean contrast value of 0.49 ± 0.11. The non-face object images have a mean

grayscale intensity value of 115.2 ± 0.9 and a mean contrast value of 0.46 ± 0.12. On average, there is no statistical difference between the two sets of images on either the grayscale intensity value or image contrast. Pixel-wise averaging of either all the face images or all the non-face images did not reveal any identifiable structure (Fig. 2, the right most column). We created a Fourier phase scrambled version of each image (e.g., Fig. 2, bottom rows) by randomizing the phase of the original images (Sadr and Sinha 2004), as used in previous EEG studies with similar image sets (e.g., Rossion et al. 2015). At a global level, these images contain the same low-level visual information (i.e., power spectra) of the original images, but without any recognizable structure. The phase-scrambled images were used in a control condition to test whether face-selective voxels activated in the FPS-fMRI paradigm are sensitive to low-level visual cues (i.e., amplitude spectrum).

Stimulation procedure

The images were back-projected in full color onto a projection screen by an MRI-compatible LCD projector and viewed by the participant through a mirror placed within the RF head coil at a viewing distance of 43 cm. They extended 14.6° × 14.6° of visual angle at the viewing distance (or 11 × 11 cm on the screen). The remaining area of the screen was set to a uniform gray background. The whole experiment procedure was controlled through a stimulation program running in Java, which also collected behavioral responses.

Main stimulation: FPS-fMRI

As shown in Fig. 1a, in the FPS sequence, the images were displayed at a base rate of 6 Hz (i.e., 6 natural images/s, the blue bins in the figure), thus with a stimulus onset asynchrony (SOA) of 166.7 ms (ten screen refresh cycles at a refresh rate of 60 Hz). Images were contrast modulated by a sinusoidal function so that each image appeared at 0% contrast, reached 100% contrast at the 6th frame and then dropped its contrast to 9.55% at the 10th frame (Fig. 1b). The same sinusoidal contrast modulation has been used in previous EEG studies (e.g., Rossion and Boremanse 2011), in particular with this paradigm (e.g., Rossion et al. 2015). Although virtually identical amplitude results can be observed in EEG when stimulating with 50% duty-cycle square waves and sine waves (e.g., Retter and Rossion 2016), using sine waves has several advantages. First, since images at low contrast are visible, the visual stimulation is almost always present and gives the observer a continuous perceptual stimulation; second, it is smoother as a visual stimulation than abrupt changes (i.e., square waves).

Every 9 s, a set of seven faces appeared at a rate of 3 Hz, i.e., alternating with non-face images (referred to as a face burst, covering 2.167 s, the red bins in Fig. 1a). Each run had a length of 396 s so that the face burst appeared 44 times at a fixed frequency of 1/9 Hz (i.e., 0.111 Hz, referred to as the face stimulation frequency). For each presentation, an image was drawn from the corresponding image set (face or non-face) according to a random order. When all the images in the respective sets had been presented, a new random order was generated and the images were drawn according to this new random order. So across all the runs and all the subjects, the order of image presentation was always different as it was randomly generated online. In total, in one FPS-fMRI run, face images appeared 7 × 44 = 308 times while non-face images appeared 2068 times.[1] We refer to this condition as FPS-face.

Phase-scrambled image stimulation

In this condition, we used exactly the same design as the FPS-face sequence but all face and non-face images were phase scrambled. We refer this condition to as FPS scrambled.

Conventional face localizer

To test the sensitivity and selectivity of the FPS-fMRI approach, we included a conventional fMRI block design functional localizer ("face localizer", e.g., Kanwisher et al. 1997), and referred it to as CONV-face condition. In this design, faces and non-face images appeared in alternating 18-s blocks with no fixation rest period between blocks. Such a design is thought to be optimal for estimating the contrast between two conditions (Smith et al. 2007; Maus et al. 2010). We matched the conventional face localizer to the FPS-face runs for two critical aspects: run duration and total number of face images presented. Specifically, the run length was exactly the same as in the FPS-face sequence (396 s) with 11 repetitions of the face and non-face blocks. In total, face images appeared for 28 × 11 = 308 times, which

[1] Given this ratio and the respective number of nonface images and face images, nonface images repeat more often during a run than face images (i.e., ~3 times per face image versus ~10 times per object image). Equating the number of repetitions here would require using about 600 object images. Alternatively, one could reduce the number of face images, but at the expense of generalizability. Importantly, human electrophysiological studies using this stimulation mode have shown the same face-selective response with face and nonface images being equated for repetition (e.g., Rossion et al. 2015; Jacques et al., 2016) or not (Retter and Rossion 2016). Most importantly, the latter study directly demonstrated that the face-selective response is immune to large variations in ratios between the number of presented face and nonface images (Retter and Rossion 2016).

matched the number of the appearance of faces in the FPS-face runs. However, in the conventional face localizer, the non-face images also appeared for 308 times, as in typical block design studies. Within each 18 s block there were 28 images displayed at 1.56 Hz. Each image appeared for 643 ms with its contrast ramped up from 80 to 100% and then dropped back to 80% following a sinusoidal function. This way, images were presented successively, without a blank interval. The modulation of the contrast provided a relatively smooth transition between images (Supplemental movie S2 for an example).

Behavioral task and order of conditions

In all the three conditions, the participants performed the same behavioral task, orthogonal to the measure of interest. They were instructed to press a predefined key on an MRI-compatible response pad using the right index finger when they detected color changes of the central crosshairs superimposed on the images (Rossion et al. 2015). The crosshairs extended a visual angle of 1.2° in the center of the screen. During each run, the color of the crosshairs changed from black to white for 200 ms, for a total of 70 times with the interval between two changes randomized while keeping above a minimal interval of 2 s. All participants achieved high accuracy (mean accuracy across conditions range from 0.927 to 0.993) in the behavioral task with no significant difference among the three conditions in either accuracy ($F2,11 = 3.53$, $p = 0.07$, Greenhouse–Geisser corrected) or correct response time ($F2,11 = 1.55$, $p = 0.23$).

Each participant started the first session with one run of each condition (FPS-face, FPS-scrambled, CONV-face) in a random order. A high-resolution anatomic image was obtained after the first session. After the anatomic image, each participant continued with the second session of one run of each of the three conditions in a pseudo-random order with the first run in the second session being a different condition from the last run in the first session. The participants took a short break after finishing the second session. After the break, they continued with the third session with only one run of the FPS-face condition and one run of the CONV-face condition in a pseudo-random order, with the first run in the third session being a different condition from the last run in the second session. In total, each participant had three FPS-face runs, three CONV-face runs, two FPS-scramble runs, and one anatomic run. The whole experiment took about 1 h and 30 min. Only two FPS-scrambled runs were collected to save scanning time and because the FPS-scrambled runs were not compared to the other conditions quantitatively.

MR image acquisition

We acquired the MRI images using a 3T Siemens Magnetom Trio system (Siemens Medical System, Erlangen, Germany) with a 32-channel head coil. Anatomic images were collected using a high-resolution T1-weighted magnetization-prepared gradient-echo image (MP-RAGE) sequence (192 sagittal slices, TR 2300 ms, TE 2.62 ms, voxel size 1 mm isotropic, FA 9°, FoV 256 × 256 mm^2, matrix size 256 × 256, parallel scanning mode GRAPPA, accelerate factor 2). The acquisition time for the anatomic scan was 321 s. Functional images were collected with a T_2*-weighted gradient-echo echoplanar imaging (EPI) sequence (TR 1,500 ms, TE 30 ms, FA 62°, voxel size 3 mm isotropic, FoV 192 × 192 mm^2, matrix size 64 × 64, interleaved, parallel scanning mode GRAPPA, accelerate factor 2), which acquired 25 oblique–axial slices covering the whole occipital lobe and the whole temporal lobe, and most of the parietal and frontal lobes, but missing the superior portion of the parietal and frontal lobes. The acquisition time for each functional run was 414 s.

MRI/fMRI analysis[2]

Preprocessing

The functional runs were motion-corrected in reference to the average image of the first functional run of the experiment using a 6° rigid body translation and rotation via an intra-modal volume linear registration using the FMRIB Software Library (FSL, version 5.0.8, Smith et al. 2004). The motion-corrected images were spatially smoothed with a Gaussian kernel with a moderate size (3 mm FWHM) to increase overlap of regions across runs and reduce noise level while keeping a high spatial resolution (Worsley et al. 1996).

SNR (*z*-score) of category-selective response

For each run with the FPS-fMRI paradigm (FPS-face and FPS-scrambled), we removed linear trends from the preprocessed time series data of each voxel by removing the best straight-line fit to the data and converted the time series data to percentage of BOLD signal change by dividing the time series of each voxel by its mean signal intensity. We then performed fast Fourier transform (FFT) to obtain the amplitude spectrum of the time series. To gauge the strength of the BOLD response at the face stimulation frequency (the signal) relative to the noise, we converted the amplitude of

[2] All the data analysis scripts in the current study are available upon request.

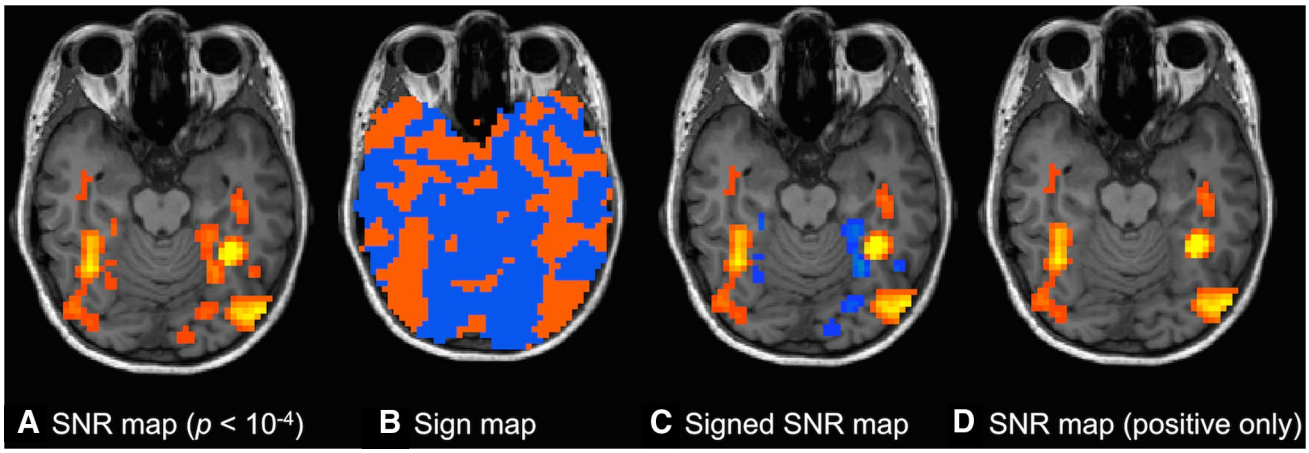

Fig. 3 Sign masking of the FPS-face runs (in one example brain). **a** Unsigned SNR map with voxels that are above threshold of $p < 10^{-4}$, uncorrected. **b** Sign map defined by phase value from FFT analysis, orange for positive sign (activation) and blue for negative sign (deactivation). **c** A signed SNR map created by applying the sign map in **b** to the map in **a. d** SNR map with negative values removed

the face stimulation frequency (0.111 Hz) to a *z*-score as in previous studies (McCarthy et al. 1994; Puce et al. 1995). By definition, a *z*-score is the baseline corrected signal level over noise level measured as the noise standard deviation, and is thus considered as a measure of SNR of the face-selective neural activity (Welvaert and Rosseel 2013):

$$\mathrm{SNR} = (A_S - \mu_N)/\sigma_N \quad (1)$$

where A_S is the amplitude of the face stimulation frequency, μ_N is the mean of 40 neighboring frequency bins (20 on each side, with a bin width of 0.0025 Hz; e.g., Rossion et al. 2015; Jonas et al. 2016) and σ_N is the standard deviation of the amplitude of the 40 neighboring frequency bins. This procedure is applied to each voxel independently. Since the face stimulation frequency (0.111 Hz) is set by the experimenter, the SNR of the face stimulation frequency provides an objective measure of face-selective neural responses.

Defining activation/deactivation using response phase

A voxel having a high SNR value at the face stimulation frequency responded differentially to faces and objects, but not necessarily more to faces than to objects. Indeed, it is possible for a voxel to achieve a high SNR at the face stimulation frequency if the BOLD response systematically decreases (i.e., deactivates) to the appearance of faces relative to the objects. We used the phase value from the FFT analysis to define the direction of category-selective responses in the FPS-face condition for each voxel. In general, a phase value of zero means reaching maximum BOLD response amplitude at the onset of faces. A positive phase value means increasing BOLD response amplitude after the onset of faces, while a negative phase value means decreasing BOLD response amplitude after the onset of faces. To account for individual differences in the time to reach maximum BOLD response amplitude after the onset of faces, we plotted the histogram (20 bins) of phase values of all the voxels with a *z*-score above 3 and with only a positive phase value. We used the phase value of the histogram bin that has the largest number as the center phase (φ) and defined all the voxels with their phase values within the window of ($\varphi \pm \pi/2$) as activation (+sign) and voxels with their phase values outside of this window as deactivation (−sign). We created a sign map and applied this map to a thresholded SNR (*z*-score) map (Fig. 3). We obtained the final category-selective response map containing only voxels that have increased BOLD response to the presence of faces.

Conventional face localizer

For each CONV-face run, following a standard procedure (e.g., Weiner and Grill-Spector 2010), we estimated the voxel-wise BOLD response amplitudes to face blocks and non-face blocks by fitting the preprocessed time series data with a general linear model (GLM), convolved with a hemodynamic response function (canonical HRF, SPM8) with temporal derivatives, autocorrelation model type AR(1), and with nuisance regressors including six movement parameters. We conducted a linear contrast to obtain the *t* statistical map where the BOLD responses to faces were greater than to non-face images. By nature, the *t* values are a measure of SNR, since the *t* values are calculated by dividing the difference of activation amplitudes between faces and non-face images (signal) by the estimated standard error of the activation (noise). To make the SNR directly comparable between conditions, we converted the *t* values to *z*-scores by calculating the corresponding *p* values from the *t* distributions.

Using data from all runs

In the FPS-face condition, across runs, the responses to the periodic face stimulations in a given population of neurons should have the same phase, while any noise from a periodic source (e.g., pulse, breathing) could have different phases across runs. Therefore, we averaged the time series across the three FPS-face runs to increase the signal-to-noise ratio, similar to the use of this approach in electrophysiology (Regan 1989). Since the same principle applies to the FPS-scrambled condition, we averaged time series across the two FPS-scrambled runs. For the conventional localizer, we ran the same GLM analysis as with individual runs, but with the time series from all three runs. In this case, run is added to the GLM analysis as a fixed effect. Using data from all three runs in the GLM analysis, we increased the degree of freedom in the face > non-face contrast from 241 to 723. Therefore, we increased power to detect any differences between neural responses to faces and non-face objects. With the same procedure as with individual runs, we created a z-map for each condition with data from all runs for each participant.

Analysis procedures

We performed qualitative and quantitative analyses to evaluate the effectiveness of the FPS-fMRI approach. We first identified all the brain areas with face-selective activity above a statistical threshold level. Importantly, we expected to disclose activation in the core face processing network, including the OFA, FFA, and pSTS at both the group level and in individual brains, while we expected little or no activation of these areas in the FPS-scrambled condition. The group level analysis reveals the most consistent areas across participants, while the individual level analysis provides complementary information with regard to individual variability in functional localization. We also assessed the specificity of the FPS-fMRI paradigm by measuring putative activation in low-level visual areas as a result of face stimulation as well as activation of face-selective areas in the FPS-scrambled condition. Within anatomically defined ROIs, we quantified the magnitude (as peak SNR) and extent (as number of super-threshold voxels) of face-selective activity. With such quantitative measures, we compared the level of face-selective activity between hemispheres and between the FPS-fMRI approach and the conventional approach. Since we collected multiple scanning runs for each paradigm, we calculated the test–retest reliability in defining the spatial activation map and compared the reliability between the two paradigms. Finally, we tested the respective contribution of the stimulation paradigm (FPS vs. conventional) and analysis procedure (Fourier analysis vs. GLM) in detecting face-selective activity.

Group analysis

We used the individual z-maps derived from data based on all runs for each condition to calculate an average group z-map for each condition. We used cortical surface-based averaging algorithms in FreeSurfer, which has been show to yield a better alignment across subjects than volume-based normalization methods (Fischl et al. 1999). The surface-based approach morphs each participant's cortical surface reconstructed from high-resolution anatomical scan to an averaged spherical surface of all the 12 participants using a best-fit sulci alignment. Individual z-maps were then interpolated onto the average sphere and averaged across participants. The averaged z-scores are no longer from a standard normal distribution. Instead, they have a standard deviation of $1/\sqrt{n}$ ($n = 12$, number of participants). Therefore, we adjusted the z threshold with the same factor ($1/\sqrt{n}$). Because group averaging can be affected by extreme values, we applied a conservative threshold ($p < 10^{-6}$, uncorrected, equivalent to a Bonferroni corrected $p < 0.05$) to identify areas that are face-selective at a group level.

Region of interest (ROI)-based individual analysis

Previous studies have shown a considerable amount of variability across individuals in both brain structure and functional localization (Frost and Goebel 2012; Zilles and Amunts 2013), particularly in face localizers (Rossion et al. 2012; Zhen et al. 2015). Therefore, it is critical to also analyze the face-selective responses at an individual level. We first created an average cortical surface and mapped the individual face-selective activity to the average cortical surface as in the group analysis. We then created 20 anatomically defined ROIs (10 in each hemisphere) on the average cortical surface. The ROIs were selected based on the results of the group analysis and were defined by an automatical parcellation scheme by Freesurfer. The parcellation algorithms are based on anatomical rules and have good concordance with manual labels (Destrieux et al. 2010). The selected ROIs include: inferior occipital gyrus, calcarine sulcus, cuneus, parieto-occipital sulcus, precuneus, fusiform gyrus, lingual gyrus, anterior occipito-temporal sulcus, superior temporal sulcus, and inferior frontal sulcus. In addition, we manually drew on the cortical surface a ROI covering the anterior collateral sulcus. For the superior temporal sulcus, we further divided it into an anterior portion and a posterior portion with the boundary defined by the posterior tip of the hippocampus (Kim et al. 2000). We combined the fusiform gyrus and the lateral occipito-temporal sulcus into a single ROI, since the activation in the fusiform gyrus extends laterally into the sulcus. We also combined the anterior occipito-temporal sulcus and the anterior collateral sulcus into a

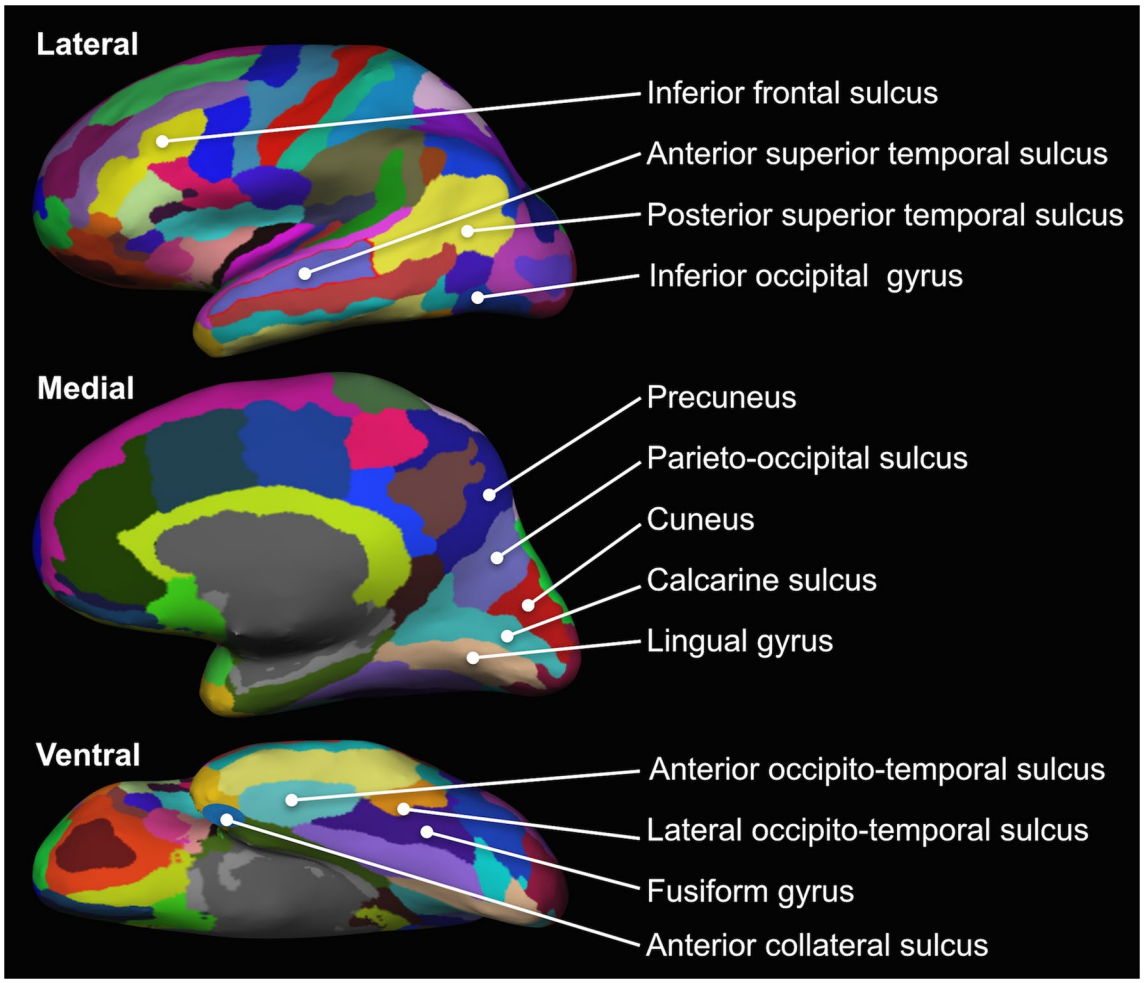

Fig. 4 Anatomically defined regions of interests (ROIs). The ROIs were selected based on the results from the group analysis. The ROIs were automatically parcellated by Freesurfer based on reconstructed cortical surface from high-resolution anatomic scans. We divided the superior temporal sulcus into an anterior portion and a posterior portion with the boundary defined by the posterior tip of the hippocampus. We combined fusiform gyrus and lateral occipito-temporal sulcus to be one ROI since the activation in the fusiform gyrus extends laterally into the sulcus. We manually labeled anterior collateral sulcus and combined it with anterior occipito-temporal sulcus to form the ATL ROI

single ROI, representing a face-selective area in the ATL. Figure 4 shows the ROIs on an inflated cortical surface.

Test–retest reliability in functional localization

To assess the reliability of the FPS paradigm, we analyzed data of each individual run in the FPS-face condition. In comparison, we also analyzed the individual runs in the CONV-face condition. We quantified the overlap between face-selective voxels identified in different runs, using the Dice coefficient:

$$O_{ij} = 2 \times V_{ij}/(V_i + V_j) \quad (2)$$

where O_{ij} is the consistency score between run i and run j, V_{ij} is the number of super-threshold voxels in both runs, V_i is the number of super-threshold voxels in run i, and V_j is the number of super-threshold voxels in run j.

For each individual participant, within each condition, we calculated the consistency scores between run 1 and run 2, between run 1 and run 3, and between run 2 and run 3. We then averaged the three consistency scores to obtain a consistency score for that condition for every individual and compute an average score of reliability at the group level. We ran this analysis at the whole brain level, within the anatomically defined right fusiform gyrus, and within the functionally defined right FFA. For the analysis within functionally defined right FFA, since the data from the first run were used to define the ROI, we calculated the dice coefficient with data from run 2 and run 3.

Respective contribution of analysis procedure and stimulation paradigm

One advantage of the FPS paradigm is that it does not rely on the GLM framework. Instead, neural response amplitude is measured by applying FFT to the neural response time course, as in early fMRI face-localizer studies (Puce et al. 1995). As explained earlier, the FFT approach has high SNR since the signal is only affected by narrow-band noise. To investigate the potential impact of the two data analysis procedures (GLM vs. FFT), we performed GLM analysis to the FPS-face data. Correspondingly, we performed FFT analysis to the data from the CONV-face runs.

GLM analysis of the FPS-face data

As with the GLM analysis with the data from the CONV-face runs, we modeled BOLD response with two event types, faces and non-face objects, the two events alternating 44 times within each scan run. Each face event lasted for 2.167 s, while each non-face event lasted for 6.833 s with no gap in between events. We included all three FPS-face runs in the model data and modeled run as a fixed effect. A linear contrast (t test) was constructed to compare the response amplitudes to faces and to non-face objects. The resulting t values were converted to z-scores for further comparison.

FFT analysis of the CONV-face data

The conventional localizer included 18-s face blocks and 18-s non-face object blocks alternating 11 times within each scanning run. Such a periodic presentation of the blocks allowed to analyze the neural responses in the frequency domain. Specifically, both face-selective and non-face selective neural responses should have high amplitudes at 1/36 Hz. Similar to the FPS-face condition, we defined signs for each voxel using the phase value from the FFT analysis.

Results

FPS-fMRI effectively defines category-selective BOLD responses

To illustrate the effectiveness of the FPS-fMRI paradigm in defining category-selective BOLD responses, we first consider a representative voxel, identified by showing the largest faces > objects contrast in a single brain in the conventional face localizer (Fig. 5a). This voxel is located in the region showing the largest face-selective response in the human brain, the lateral section of the right fusiform gyrus, often defined as the fusiform face area (FFA, Kanwisher et al. 1997). Fast Fourier transform (FFT) applied to the BOLD response time course of this voxel (Fig. 5b) reveals a high signal at the face stimulation frequency (0.111 Hz) in the FPS-face condition (Fig. 5c). Given the very high frequency resolution (1/396 s = 0.0025 Hz), the signal is concentrated on a tiny frequency bin in the amplitude spectrum of the BOLD response. For this example voxel, all of the responses of interest concentrate on the fundamental frequency, i.e., 0.111 Hz. In the same region in a few other individual brains, there were also negligible amplitude increases at the second harmonic (0.222 Hz) (Supplemental Figure S2).

In all individual brains, the FPS-face condition achieved very high SNR (z-score ranged from 13.9 to 31.0, mean 20.8 ± 5.2) at the face stimulation frequency in the peak face-selective voxel in the right FFA identified by the conventional face localizer (z-score ranged from 7.9 to 17.3, mean 12.4 ± 2.9). Therefore, the FPS-fMRI paradigm can effectively modulate the BOLD signal representing a differential neural response to faces vs. non-face objects. Strikingly, there was no signal at the face stimulation frequency in the FPS-scrambled condition in any of these voxels [z-score mean 0.25 ± 1.04, $t(11)$ 0.83, $p = 0.42$, two-tailed test against 0]. The absence of response in the BOLD amplitude spectrum of the same voxels for phase-scrambled images implies that low-level image properties contained in the power spectrum did not contribute to the peak of the face-selective BOLD response at the face stimulation frequency in the right FFA, as well as in other face-selective regions (Supplemental Figure S3).

Mapping face-selective cortical areas (group analysis)

In the group averaged activation map of the FPS-face condition, as shown in Fig. 6a, with a conservative threshold level of uncorrected $p < 10^{-6}$ (equivalent to a Bonferroni corrected $p < 0.05$), in both hemispheres, we identified face-selective areas in the well-known core face processing network (Haxby et al. 2000; Duchaine and Yovel 2015) consisting of three cortical areas, namely the FFA in the middle section of the lateral fusiform gyrus (FG), the occipital face area ("OFA") in the inferior occipital gyrus (IOG), and the posterior superior temporal sulcus (pSTS). In both hemispheres, the most significant response, i.e., peak of face-selectivity, corresponds to the FFA, and the highest average z-score is found in the right FFA, with a typical 3D coordinate in a normalized brain (42, −54, −14 in Talairach coordinates, compare, for instance, to the right FFA coordinate in Kanwisher et al. (1997): 40, −55, −10; Zhen et al. (2015), for pFFA: 42, −51, −14 (converted from MNI to Talairach); Jonas et al. (2016), FFA identified in human intracerebral recording: 41, −45, −16).

Importantly, the group averaged map of the FPS-face condition also revealed consistent face-selective activity in the

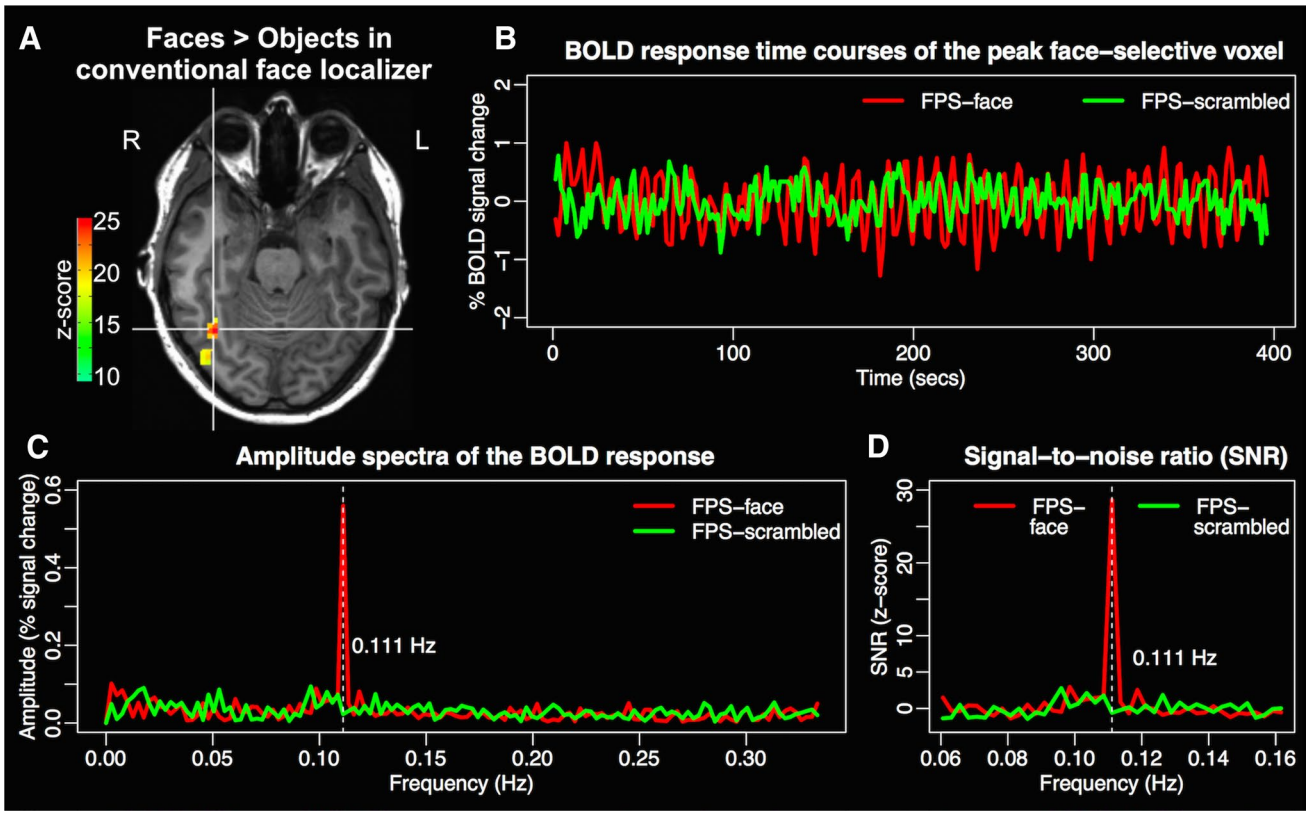

Fig. 5 BOLD response in an example face-selective voxel. **a** Peak face-selective voxel (cross hair, right lateral fusiform gyrus) defined by faces > objects contrast in the conventional face localizer. **b** BOLD response time courses of the peak face-selective voxel in the FPS paradigm with natural images (FPS-face, in red) or with phase-scrambled images (FPS-scrambled, in green). **c** Amplitude spectra of the BOLD response. The white dashed line indicates the face stimulation frequency (0.111 Hz). **d** Signal-to-noise ratio (z-score) of the amplitudes

anterior temporal lobe (ATL) in both hemispheres (Fig. 6a). The role of ATL in face-selective processing is known from intracerebral recordings (Puce 1999; Jonas et al. 2016) and human lesion studies causing individual face recognition impairments (e.g., Busigny et al. 2014) but has only been recently brought to attention in fMRI (Rajimehr et al. 2009; Nasr and Tootell 2012; Collins and Olson 2014; Von Der Heide et al. 2013; Collins et al. 2016), due to the difficulty in identifying these regions associated with a drop of SNR and large magnetic susceptibility artifacts (Axelrod and Yovel 2013; Jonas et al. 2015; Rajimehr et al. 2009; Wandell 2011; Lafer-Sousa et al. 2016; see Rossion et al. 2018 for a discussion of this issue) (see Supplemental Figure S4 for an example of MRI signal drop in the ATL).

Besides the core face processing network and the ATL, the group averaged map of the FPS-face condition also identified consistent face-selective activity in areas previously reported to be involved in face processing in neuroimaging studies. These areas include the inferior frontal sulcus (IFS; e.g., Fox et al. 2009; Ishai et al. 2005; Chan and Downing 2011), lingual gyrus (LG; e.g., Gobbini and Haxby 2006; Rossion et al. 2012), Cuneus (e.g., Benuzzi et al. 2007; Rossion et al. 2012), and parieto-occipital sulcus (POS; e.g., Tuladhar et al. 2007; Jokish and; Jensen 2007).

In the group-averaged map of the CONV-face condition, at the same threshold level, we also identified the core face-selective network including bilateral FFA, OFA, and pSTS. However, we noticed three marked differences in comparison to the FPS-face condition: (1) the magnitude of face-selective activation in the FPS-face condition is much higher than in the CONV-face condition. The average maximum SNR (z-score) in the FPS-face condition (25.0 ± 5.7) is about twofold higher than in the CONV-face condition (13.1 ± 3.2, $t(11) = 6.2$, $p < 0.001$, Cohen's $d = 2.57$, Fig. 6b). Notably, such an increase in the peak face-selective neural responses in the FPS approach compared to the conventional approach is found in every individual brain tested (range of the ratio of increase 1.1–3.4). At the same time, there is no significant difference between the total volumes of face-selective activity identified in the two conditions ($p = 0.26$, Fig. 6c); (2) at this threshold level ($p < 10^{-6}$, uncorrected). This is because of the lack of face-selective activity in ATL in the group-averaged map (Fig. 6a); (3) in the CONV-face condition is compensated by extensive activity in the low-level

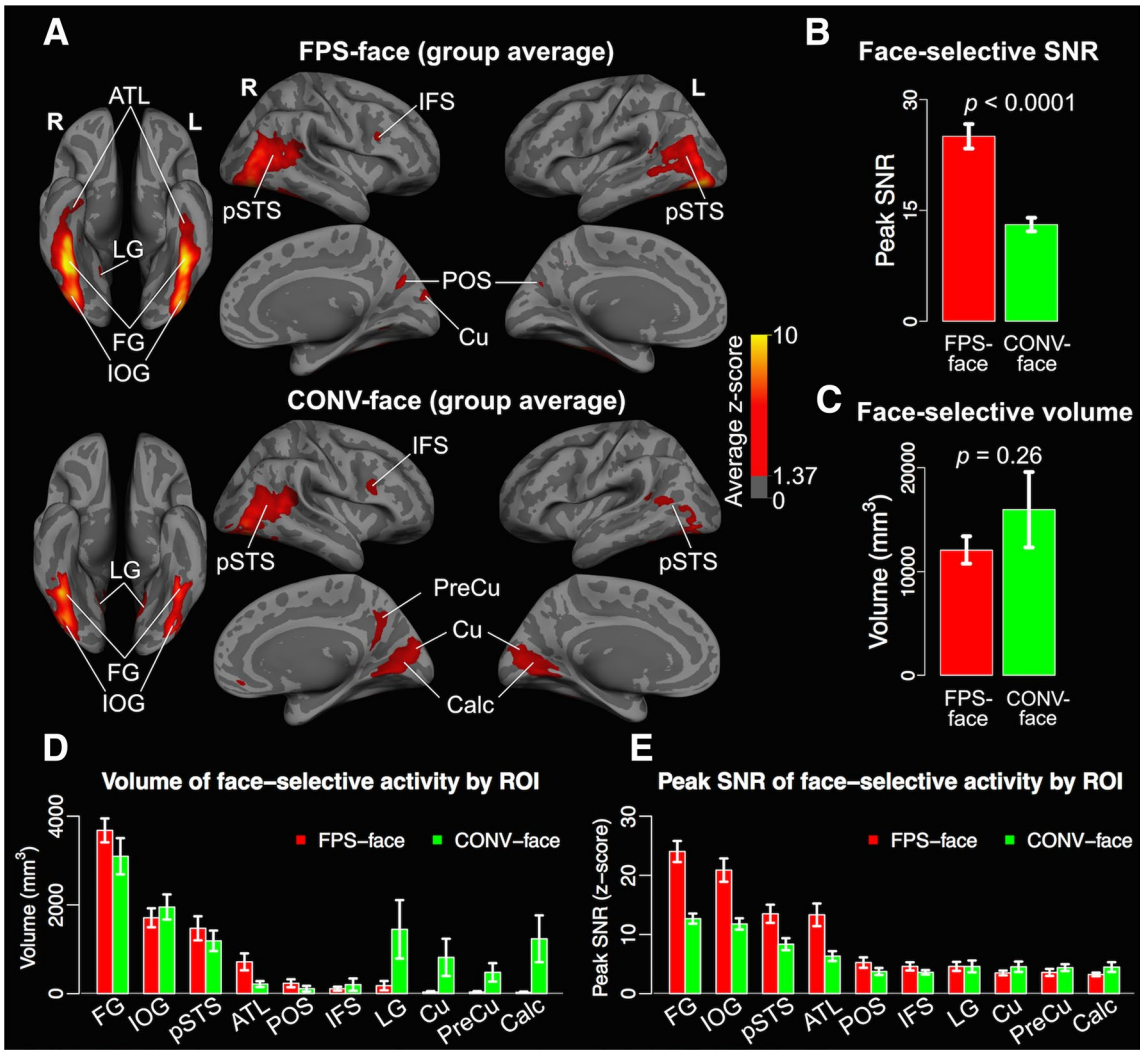

Fig. 6 Face-selective activity in the whole cortex and in anatomic ROIs. **a** Face-selective cortical areas identified based on group averaged z maps ($n=12$). The maps are thresholded at $p<10^{-6}$, uncorrected, and are projected onto an inflated average cortical surface of all the participants. **b** Peak SNR of face-selective activity in the whole cortex in the FPS-face condition and CONV-face condition, averaged across participants. **c** Total volume of face-selective responses in the whole cortex averaged over individual brains with a threshold level of $p<10^{-4}$, uncorrected. **d** Volume of face-selective activity in anatomical ROIs averaged over individual brains with a threshold level of $p<10^{-4}$, uncorrected. **e** Peak SNR of face-selective activity in anatomical ROIs, averaged across participants. Data are represented as mean ± SEM. In **d** and **e**, we combined homologous ROIs across hemispheres, as the results were similar for both hemispheres. *FG* fusiform gyrus, *IOG* inferior occipital gyrus, *pSTS* posterior superior temporal sulcus, *ATL* anterior temporal lobe, *POS* parieto-occipital sulcus, *IFS* inferior frontal sulcus, *LG* lingual gyrus, *Cu* cuneus, *PreCu* precuneus, *Calc* calcarine sulcus

visual regions (e.g., calcarine sulcus, Fig. 6a, d) only in this CONV-face condition.

Face-selective cortical areas in individual brains

Complementary to the group activation maps, we quantified face-selective neural activity in each individual brain with a threshold level of $p<10^{-4}$, uncorrected. This is a necessary step because substantial variability exists in brain structure as well as in functional localization (Frost and Goebel 2012; Rossion et al. 2012; Zhen et al. 2015; Zilles and Amunts 2013). As can be seen in Fig. 7, the face-selective activations identified in the FPS-face condition and the CONV-face condition are more focal in individual brains than in the averaged maps (Fig. 6a).

In both paradigms, consistent with previous neuroimaging studies (e.g., Kanwisher et al. 1997; McCarthy et al. 1997; Sergent et al. 1992; Rossion et al. 2012; Zhen et al. 2015), we found right hemisphere dominance, i.e., larger volume of face-selective activity in the right than in the left hemisphere, for the whole cortex ($p=0.0001$ for FPS-face condition; $p=0.001$ for CONV-face condition), as well as

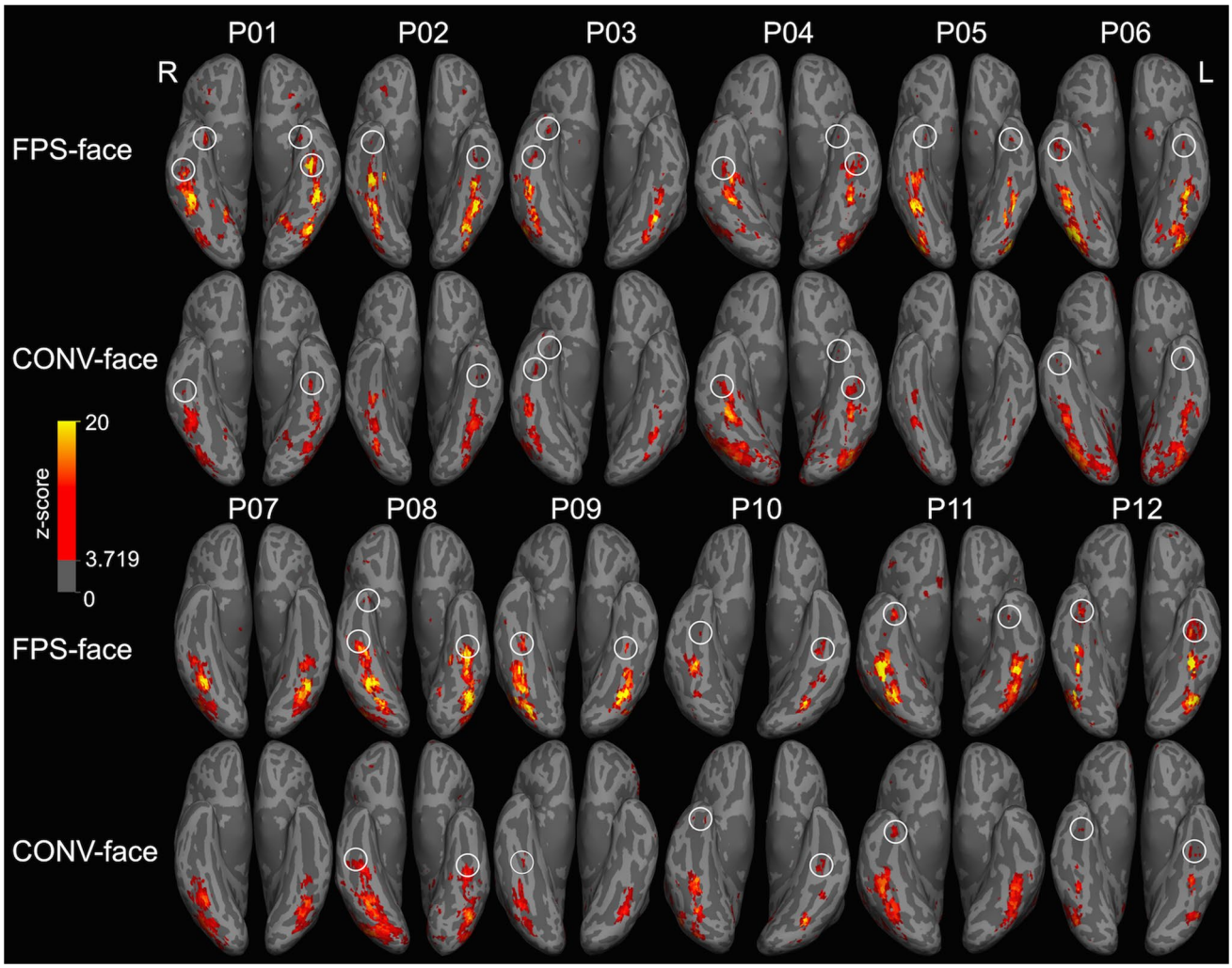

Fig. 7 Individual maps of face-selective activity in VOTC. Magnitude (*z*-score) of face-selective activity was projected to an inflated cortical surface of each individual brain for the FPS-face and the CONV-face conditions, with a threshold level of $p < 10^{-4}$, uncorrected. Note the overlap between the face-selective regions identified by the two conditions, but with much higher magnitude reached in the FPS paradigm, leading to extra activation clusters in the anterior temporal lobe (highlighted with white circles). See Supplemental Figure S5 for lateral and medial views of face-selective activity in all the individual brains

within ROIs: FG ($p = 0.0001$), pSTS ($p = 0.004$), and POS ($p = 0.009$) for the FPS-face condition; FG ($p = 0.0003$), pSTS ($p = 0.001$), and IOG ($p = 0.04$) for the CONV-face condition, and the location associated with high response magnitude is fairly consistent across the two paradigms for all individual brains (Fig. 7).

The FPS-face conditions identified $67 \pm 27\%$ of all the voxels that are identified as face-selective in the CONV-face condition, while the CONV-face condition identified $71 \pm 21\%$ of the face-selective voxels in the FPS-face condition. While the percentage of brain activation overlap may appear relatively low in regard to the visual similarity between brain maps in the two conditions (e.g., Fig. 6), in a single high-level region, e.g., the right FG, the FPS-face condition identified $91 \pm 8\%$ of the total number of face-selective voxels identified by the CONV-face condition. However, this percentage is lower for the CONV-face condition ($75 \pm 22\%$).

Where are the voxels that are activated specifically in each condition? To answer this question, we quantified the volume of above-threshold voxels in ten anatomically defined brain areas (combining homologous areas across hemispheres) selected based on results of the group analysis. As shown in Fig. 6d, while the highest proportion of volume activated was found in the three core regions of the face network (FG, IOG, pSTS) plus ATL in the FPS-face condition, this proportion was slightly lower in the CONV-face condition. Also, importantly, the volume of face-selective area in ATL was significantly larger in the FPS-face condition than in the CONV-face condition ($p = 0.011$). This is

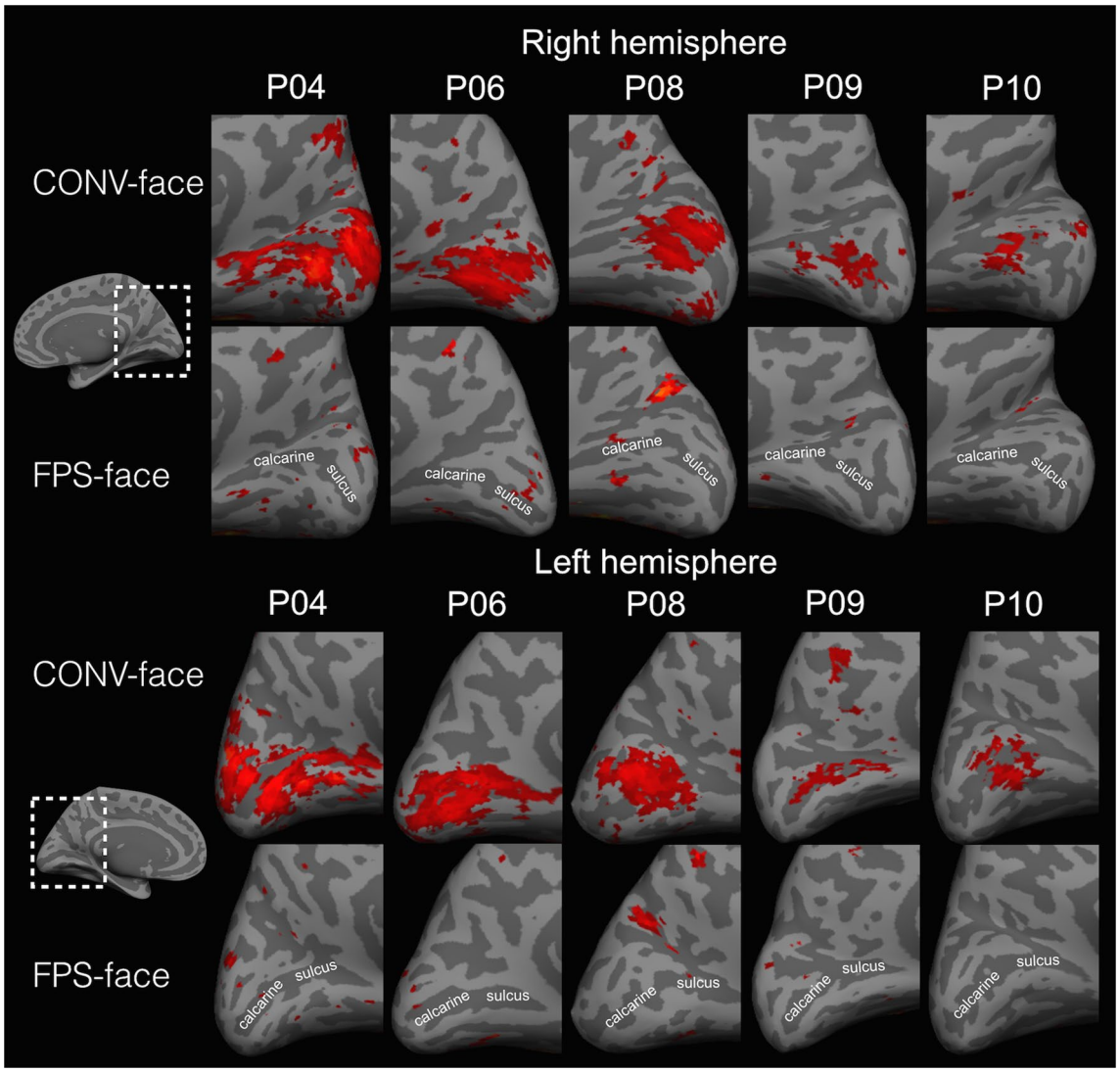

Fig. 8 "Face-selective" activity in calcarine sulcus. Inflated surfaces of the medial occipital cortex of five example participants show "face-selective" activity in bilateral calcarine sulcus in the CONV-face condition but not in the FPS-face condition. The maps are thresholded at $p < 10^{-4}$, uncorrected

particularly important in the context of the relatively recent focus on ATL in face processing and the general difficulty in identifying these regions in fMRI due to magnetic susceptibility artifacts.

Considering individual brains, the FPS approach was able to identify regions in 11 (Fig. 7) individual brains out of a total of 12, without using specific fMRI sequences to reduce magnetic susceptibility artifacts in the temporal lobe (Devlin et al. 2000; Embleton et al. 2010; Visser et al. 2010). In contrast, a much larger volume of "face-selective" activation was found in low-level visual regions such as the calcarine sulcus and the precuneus in the CONV-face condition than in the FPS-face condition in both hemispheres ($p = 0.039$, $p = 0.038$ in calcarine sulcus and precuneus, respectively, see Figs. 6d, 8). This finding suggests that even with widely variable natural images the conventional face localizer is more affected by low-level visual differences between faces and non-face object categories, as predicted.

Besides volume, the FPS paradigm achieved much higher face-selective response magnitude in higher-level areas than the CONV-face condition (Fig. 6e; FG, $p = 0.0001$; IOG, $p = 0.001$; pSTS, $p = 0.008$; ATL, $p = 0.0005$).

In summary, the FPS approach is much more sensitive to detect face-selective neural activity in typical high-level visual areas and the anterior temporal lobe, while being less contaminated by low-level confounds (i.e., more specific) than a conventional localizer procedure in fMRI.

Test–retest reliability

How reliable, or replicable, are these observations, is a key issue in the current fMRI research (Bennett and Miller 2010; Nichols et al. 2017). Figure 9a provides a visual illustration of the consistency across runs in one representative participant: there is a good correspondence of face-selective voxels across different runs within both the FPS-face condition and the CONV-face condition.

First, we confined the analysis within an anatomically defined area containing only voxels in the right fusiform gyrus (the anatomical ROI or aROI approach) to make the current measure comparable to values reported in previous studies (Berman et al. 2010; Duncan et al. 2009; Duncan and Devlin 2011). At a threshold level of uncorrected $p < 10^{-4}$, the face-selective voxels identified in different runs in the CONV-face condition reached an average dice coefficient of 0.58 ± 0.25, which is in the high end of the typical range as reported in previous studies (Berman et al. 2010; Duncan et al. 2009; Duncan and Devlin 2011). However, the FPS-face condition reached a significantly higher consistency [0.80 ± 0.07, range 0.67–0.90, $t(11) = 3.08$, $p = 0.01$], reaching 80% overlap of face-selective voxels between runs on average. Note that this consistency value is obtained in a ROI, the right fusiform gyrus, containing only about 30% of voxels above threshold, so that brain activity in the vast majority of voxels below threshold may vary randomly from run to run and decrease this consistency value. Another approach is to confine the analysis to the right FFA in every individual as defined on the first run of each condition (the functional ROI, or fROI approach), and calculate the Dice index using data from the other two runs. Here, consistency values reach 0.91 ± 0.07 (0.79–0.99, see Fig. 9b for a visual demonstration of consistency in the FFA in individual participants) in the FPS-face condition, significantly higher than in the CONV-face condition (0.67 ± 0.37; $t(11) = 2.3$, $p = 0.04$). To our knowledge, this is the highest consistency value ever reported in localizing right FFA across different imaging runs within the same session.

Since test–retest reliability may decrease with increasing threshold (Berman et al. 2010; Duncan et al. 2009; Duncan and Devlin 2011), we calculated the Dice coefficients at five levels from a liberal (uncorrected $p < 10^{-2}$) to a conservative (uncorrected $p < 10^{-6}$, equivalent to a Bonferroni corrected $p < 0.05$) threshold. We performed the calculation separately within the anatomically defined right fusiform gyrus and the whole cortex (aROI approach). We also calculated the reliability score within functionally defined right FFA based on the first run of each condition, and all the super-threshold voxels in the whole cortex in the first run (fROI approach).

In the CONV-face condition, reliability decreased with increasing threshold (Fig. 9c) in both the anatomically defined right fusiform gyrus (slope = −0.04, 95% CI −0.054 to −0.030) and the functionally defined right FFA (slope = −0.04, 95% CI −0.067 to −0.019), as previously reported (Berman et al. 2010; Duncan et al. 2009; Duncan and Devlin 2011). However, the reliability decreased less steeply with increasing threshold in the FPS-face condition (slope = −0.016 for aROI approach and −0.012 for fROI approach), achieving Dice coefficients of around 0.9 at all threshold levels in the right FFA. To the best of our knowledge, this is the highest test–retest reliability of spatial overlap of activation maps across fMRI scanning runs yet reported in a high-level functional area.

In the whole cortex (Fig. 9c), reliability was lower in general comparing to corresponding values from the right fusiform gyrus (aROI) or right FFA (fROI). Nevertheless, in the FPS-face condition, within a functional mask defined by the super-threshold voxels from the first run, the reliability score reached above 0.8 at all threshold levels. In contrast, the highest average reliability score in the CONV-face condition was below 0.6.

In summary, while the conventional face localizer used in this study is at least as valid and reliable as those used in previous studies, the FPS-fMRI approach greatly enhances test–retest reliability of category-selective neural activation.

The respective contribution of FFT analysis and stimulation mode

Finally, we investigated the respective contribution of stimulation paradigm and analysis procedure to the maximum face-selective response in the whole brain and in three face-selective areas showing maximal differences between the FPS-face condition and CONV-face condition: the FG, ATL, and calcarine sulcus.

At the whole brain level (Fig. 10a), both stimulation paradigm [$F(1, 11) = 4.83$, $p = 0.05$, $\eta = 0.31$] and analysis procedure [$F(1,11) = 99.2$, $p < 0.001$, $\eta = 0.90$] contributed to peak face-selective activity, without significant interaction. We found similar patterns within FG (Fig. 10b) and ATL (Fig. 10c). There were significant main effects of stimulation paradigm [FG: $F(1, 11) = 5.53$, $p = 0.038$, $\eta = 0.33$; ATL, $F(1, 11) = 4.01$, $p = 0.07$, $\eta = 0.27$] and analysis procedure [FG: $F(1,11) = 67.6$, $p < 0.001$, $\eta = 0.86$; ATL: $F(1,11) = 11.4$, $p = 0.002$, $\eta = 0.58$] without any interaction. Hence, the results suggest that both factors contribute to the highest face-selective responses observed in key regions. The fast and periodic presentation of visual stimuli increases face-selective SNR over the slow and blocked presentation. At the same time, analyzing the data using FFT also increases face-selective SNR over the conventional GLM approach.

In the calcarine sulcus (Fig. 10d), the main effects of the stimulation paradigm [$F(1,11) = 11.1$, $p = 0.007$, $\eta = 0.50$] and analysis procedure [$F(1,11) = 17.8$,

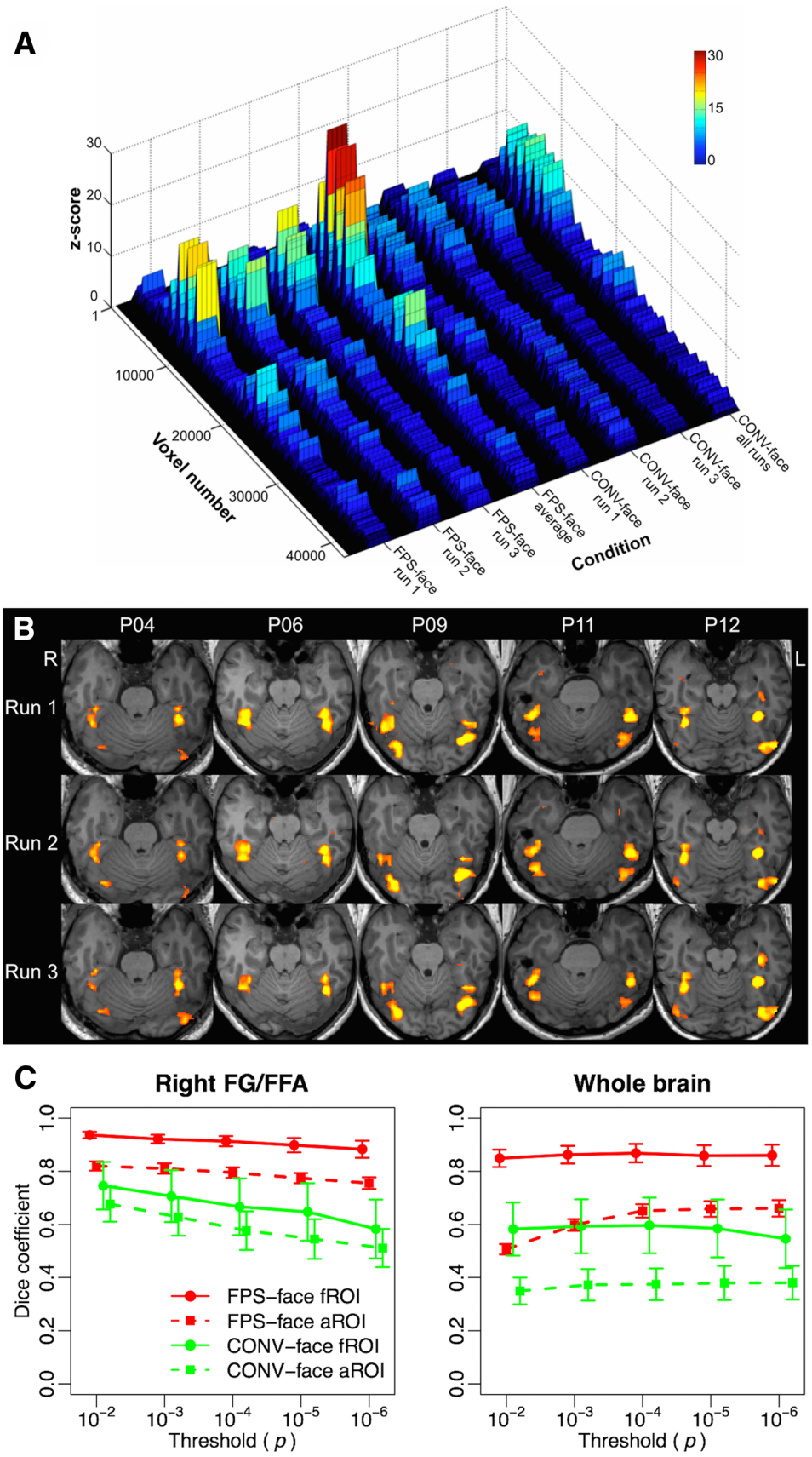

Fig. 9 Test–retest reliability of face-selective spatial activation maps. **a** An example of face-selective neural response in each run of the FPS-face condition (FPS-face runs 1, 2, 3) and the CONV-face condition (CONV-face runs 1, 2, 3) in one participant (see Supplemental Figure S6 for all the participants). "FPS-face run average" was calculated based on the average BOLD response time course of the three FPS-face runs. "CONV-face all runs" was calculated based on data from all the three CONV-face runs. Voxels in the whole brain were listed on a one-dimensional axis (voxel number), so only spatial correspondence between different conditions was maintained. **b** Example spatial activation maps of three FPS-face runs in five participants (threshold level, $p < 10^{-4}$, uncorrected). **c** Dice coefficient of spatial overlap of super-threshold voxels across three runs in the FPS-face and CONV-face conditions at five levels of thresholds. Data are represented as mean ± SEM. Left panel: reliability analyses confined within anatomically defined right fusiform gyrus (aROI) or the functionally defined right FFA based on the first run (fROI). Right panel: reliability analyses in the whole cortex (aROI) or within the super-threshold voxels based on the first run (fROI)

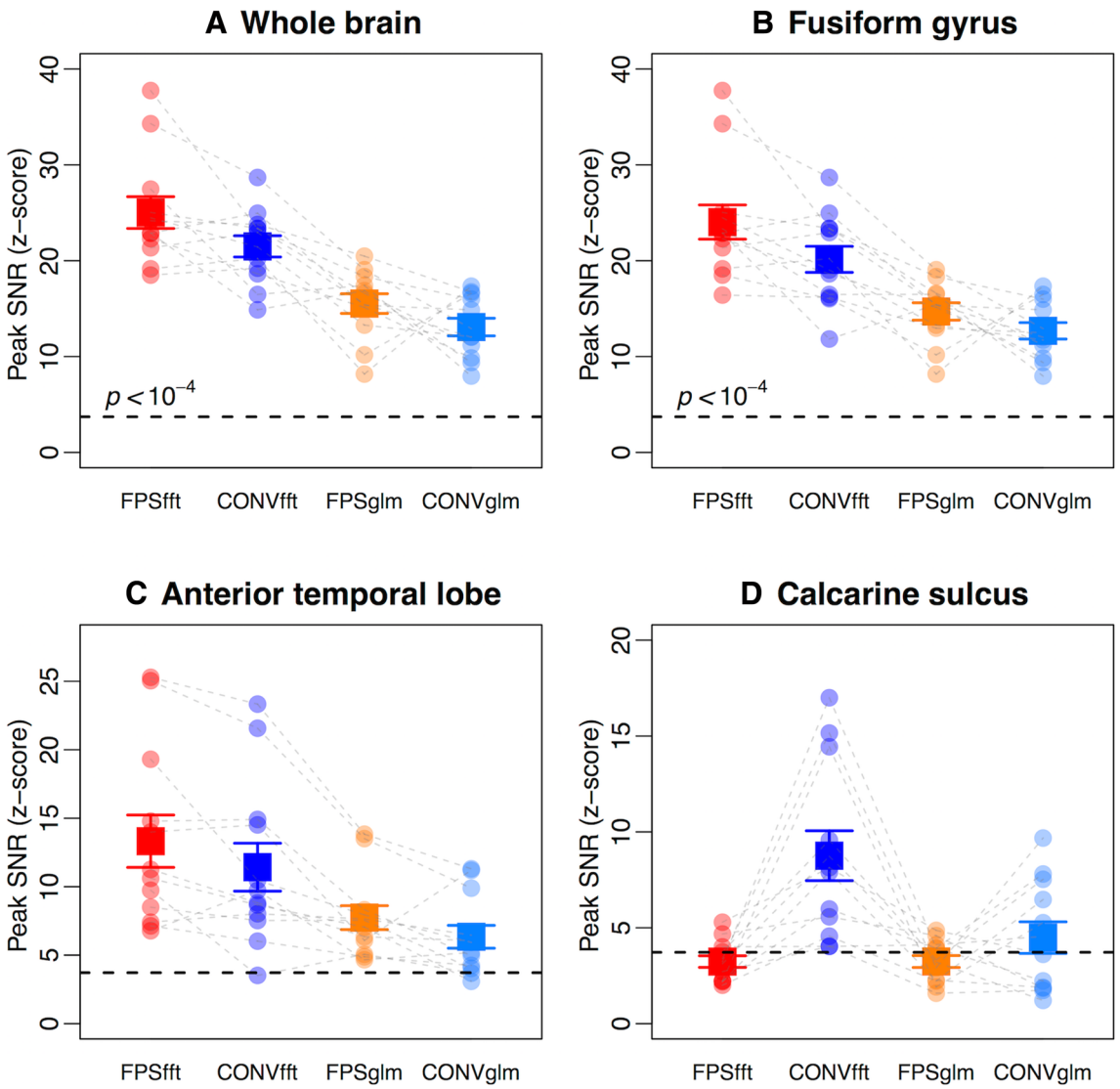

Fig. 10 Respective contribution of analysis procedure and stimulation paradigm on peak face-selective response. **a** Peak face-selective response in the whole cortex. **b** Peak face-selective response in the fusiform gyrus. **c** Peak face-selective response in the anterior temporal lobe. **d** Peak face-selective response in the calcarine sulcus. *FPSfft* FPS-face condition with FFT-based data analysis, *CONVfft* CONV-face condition with FFT-based data analysis, *FPSglm* FPS-face condition with GLM-based analysis, *CONVglm* CONV-face condition with GLM-based analysis. Each dot represents the value of one participant. The black dashed lines represent a threshold level of $p < 10^{-4}$, uncorrected. The gray dashed lines connect the values from the same participant in different conditions. The solid squares represent the group means ± SEM. We combined homologous ROIs across hemispheres for this analysis

$p = 0.001$, $\eta = 0.62$] were modulated by a significant interaction [$F(1,11) = 16.8$, $p = 0.002$, $\eta = 0.60$]. The pattern of results show that in the FPS-face condition, activity in the calcarine sulcus was low, and analysis procedure did not make any difference [$t(11) = -0.02$, $p = 0.98$]. However, the FFT analysis was more sensitive than the GLM analysis in measuring "face-selective" activity in the calcarine sulcus in the CONV-face condition [$t(11) = 4.24$, $p = 0.001$]. The results indicate that the removal of low-level visual confounds is due to the specific stimulation parameters of the FPS-fMRI paradigm (i.e., periodicity constraints, brief masked stimulation, direct contrasts).

Discussion

Many neuroscientists seek to understand brain function by characterizing and comparing neural responses within different (cortical) regions and brain networks. Understanding the nature of the processes occurring in different brain regions

and networks requires well-justified and reproducible methods for defining these regions and networks. In this context, functional localizers in fMRI have become invaluable tools, the most widely used comparing faces—arguably the most significant signal for human ecology—to non-face visual objects (Kanwisher et al. 1997; Kanwisher 2017; Grill-Spector et al. 2017). An optimal face localizer is critical to define regions in the human adult brain to explore further functions such as facial identity recognition or facial emotion decoding, but also to explore the neurodevelopment of this function (de Heering and Rossion 2015; Golarai et al. 2007; Gomez et al. 2017; Scherf et al. 2007) and its impairment following sudden onset or long-life brain dysfunction (e.g., Rossion et al. 2003; Avidan et al. 2014; Yang et al. 2016). However, it is fair to say that even after 20 years of use and development of functional fMRI localizers, severe limitations in sensitivity, objectivity, and reliability remain, explaining a large fraction of the great variability across studies in this field of research. Most studies compare neural responses to well-segmented full-front faces to responses to segmented images of a single category, or a few categories of interest only (see Berman et al. 2010). Moreover, there is a trade-off in fMRI face localizers between, on the one hand, the tight control of systematic low-level stimulus confounds, in which face and nonface stimuli can be equalized for low-level visual cues at the expense of sensitivity and generalizability, and, on the other hand, the use of more natural stimuli at the expense of specificity (e.g., "the dynamic face localizer", Fox et al. 2009; Yang et al. 2016; Zhen et al. 2015). The present FPS-fMRI approach, partly inspired from frequency tagging in human electroencephalography (Norcia et al. 2015; Rossion et al. 2015), goes a long way towards resolving these issues.

Regarding sensitivity, the FPS-fMRI face localizer achieves much (i.e., twice) higher SNR in detecting face-selective neural responses overall than a conventional face localizer matched for duration and number of face stimuli presented. This advantage was found in every individual brain tested in the study. Such a high sensitivity is especially important for studies with rare participants, such as brain-damaged patients, where the data have to be interpreted at the individual level (e.g., Fox et al. 2011; Freud et al. 2017; Rossion et al. 2003; Weiner et al. 2016). The increased SNR is due to both the stimulation paradigm and the model-free FFT analysis (Fig. 10). The stimulation rate is optimized to give brain regions just enough time to selectively process each stimulus and to record a full category-selective response before the next stimulus from the same category appears (Retter and Rossion 2016). It also maximizes the direct contrast between the category of interest, faces, and other stimuli, each face being forward- and backward-masked by non-face objects in the continuous visual stimulation stream. Due to the long run duration, frequency resolution of the fMRI signal spectrum following FFT is extremely high (0.0025 Hz). Hence, while the broadband noise spreads across numerous frequency bins, all the signal of interest falls in a tiny frequency bin (0.111 Hz), providing extremely high SNR (Regan 1989). Increased sensitivity allows reducing acquisition time by a factor of two (see Supplemental Figure S7), providing much more room for further exploration of specific functions of these regions, and to disclose all or most relevant face-selective regions in individual brains. The significant clusters in the ATL in the vast majority of individual brains with FPS-fMRI, despite signal dropout making it difficult to localize face-selective responses with conventional fMRI paradigms in this region, further illustrate this point.

The conventional way of estimating the BOLD response amplitude relies on the same HRF model being used for different regions of the brain and for different individuals. However, since there is a substantial amount of variation of BOLD hemodynamic response across brain regions as well as across participants, a uniform HRF inevitably leads to erroneous estimates (Handwerker et al. 2004). This matter is even more serious for modeling hemodynamic response in higher level regions, since commonly used HRF models are usually validated from measurements of specific regions and stimulation, such as the primary visual cortex to simple visual stimuli (e.g., Boynton et al. 1996) or the primary motor cortex to finger-tapping (Buxton et al. 1998). In contrast, the model-free FFT analysis afforded by periodicity in the FPS-fMRI design leads not only to a more objective definition and quantification of a response of interest (i.e., amplitude exactly at the pre-defined frequency), but also a fair comparison of this response across different brain regions and participants. Note that a key difference with previous frequency-based fMRI analyses (Koenig-Robert et al. 2015; McCarthy et al. 1994; Puce et al. 1995; Wang et al. 2014, 2015) is that category-selective neural responses are directly reflected as amplitude of the stimulation frequency here, without post hoc subtraction procedure.

Importantly, our approach also achieves high test–retest reliability in localizing face-selective areas. Although reliability is the pedestal of any scientific research, the results of fMRI research are generally less reliable than many researchers implicitly believe (Bennett and Miller 2010). Here, across the whole brain or within functionally defined face-selective regions, we achieve an average of 80–90% overlap of spatial activation across runs, which is to our knowledge the highest test–retest reliability observed in this field of research, and beyond (Bennett and Miller 2010; Berman et al. 2010; Duncan et al. 2009; Duncan and Devlin 2011). This high reliability makes the current paradigm desirable not only for fundamental scientific research, but, importantly, for clinical purposes, such as presurgical functional localization (Weiner et al. 2016) and/or for

intervention studies where the participants or patients are tested multiple times at different time points (Dormal et al. 2015; Weiner et al. 2016). This reliability is even more impressive considering that the specific exemplar images are selected randomly for each presentation, and thus not presented in the same order, or repeated exactly the same number of times in each run. Precisely, rather than attempting to independently define and compare abstract neural "representations" for faces and nonface objects, we attribute the high reliability of the approach to the measure of a process, perceptual categorization of faces, which is repeatedly sampled across a wide variety of stimuli and contexts (i.e., non-face stimuli) within each run, increasing invariance, i.e., robustness, of the response of interest.

Critically, the increase in sensitivity afforded by the current paradigm does not come at the expense of a decrease of specificity, which is often the case when natural stimuli vary substantially in terms of low-level properties (Fox et al. 2009; see also Rice et al. 2014 and; Andrews et al. 2015) or in block designs with one-back tasks differing in difficulty and attentional resources between face and object images (see discussion in Rossion et al. 2012). In fact, overall, there are no more voxels activated in the FPS-fMRI paradigm than in the conventional paradigm with the same stimuli: the significant activation increase in higher level regions such as the ATL, defined as face-selective in previous fMRI studies of face selectivity (Rajimehr et al. 2009; Nasr and Tootell 2012; Collins et al. 2016) and human intracranial recordings (Puce 1999; Jonas et al. 2016) and considered as a key player in face perception, is mainly compensated by the lack of "face-selective" responses in the primary visual cortex with the FPS approach compared to the conventional localizer. Note that the exact same stimuli are used in both designs and FFT analysis per se does not account for the removal of low-level visual cortex activation. Rather, removal of low-level visual confounds is due to the fast periodic stimulation paradigm. This can be tentatively explained by the following. In a conventional block design, a small minority—say 10%—of images of the same category with specific low-level visual attributes (e.g., increased power in low spatial frequencies) could be sufficient to lead to average differences between (face and object) stimulation blocks. However, a random distribution of 10% of these stimuli across the various frequency bursts of the stimulation sequence should not lead to a periodic response at 0.111 Hz. Moreover, with natural images as used here, populations of neurons responding to living things in general, for instance, would contribute to "face-selectivity" in a block design (or in studies comparing faces to houses or cars only, e.g., Berman et al. 2010; Loffler et al. 2005; Rossion et al. 2012). However, the periodicity constraint in FPS-fMRI minimizes or even eliminates such confounds (Rossion et al. 2015) because each face image contrasts with different object images throughout the sequence, and the face bursts are made of a few (different) exemplars only. Hence, a subset of identical biased image contrasts would have to be systematically present in each face burst to lead to a periodic response due to such biases. Therefore, a major strength of the present approach is the ability to preserve naturalness of the images while removing low-level visual contributions, as demonstrated here by the absence of face-selective response when these images are phase-scrambled.

The principles of FPS-fMRI validated here with a functional face localizer could be applied straightforwardly to refine our knowledge of the functions of these regions with respect to face identity and facial expression processing for instance (Calder and Young 2005; Haxby et al. 2000; Duchaine and Yovel 2015), to map cortical areas selectively coding for visual linguistic material (e.g., letter strings and words, Lochy et al. 2015) in the VOTC (Wandell 2011), or to other modalities such as auditory stimulation (Zatorre et al. 2002). More generally, our approach is able to objectively track fast processes occurring in a dynamic visual stream even with a slow temporal resolution method such as fMRI. This is important for providing direct comparisons with neural signals obtained in similar paradigms used with other imaging modalities, such as (intracranial) electroencephalography (Jonas et al. 2016), and also open new possibilities for defining functions with other low temporal resolution methods such as calcium imaging (Grienberger and Konnerth 2012) in human and nonhuman brains.

Acknowledgements We thank Valérie Goffaux, Corentin Jacques, Jacques Jonas, Kirsten Petras, and Talia Retter and two anonymous reviewers for their helpful comments on an earlier version of this paper. We also thank Talia Retter for editing the manuscript.

References

Aguirre GK, D'Esposito M (1999) Experimental design for brain fMRI. Functional MRI. In: Moonen CTW, Bandettini PA (eds) Functional MRI. Springer, Berlin, pp 369–380

Amunts K, Zilles K (2015) Architectonic mapping of the human brain beyond Brodmann. Neuron 88:1086–1107

Andrews TJ, Watson DM, Rice GE, Hartley T (2015) Low-level properties of natural images predict topographic patterns of neural response in the ventral visual pathway. J Vis 15(7):3

Avidan G, Tanzer M, Hadj-Bouziane F, Liu N, Ungerleider LG, Behrmann M (2014) Selective dissociation between core and extended regions of the face processing network in congenital prosopagnosia. Cereb Cortex 24:1565–1578

Axelrod V, Yovel G (2013) The challenge of localizing the anterior temporal face area: a possible solution. NeuroImage 81:371–380

Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS (1993) Processing strategies for time-course data sets in functional MRI of the human brain. Magn Reson Med 30:161–173

Bennett CM, Miller MB (2010) How reliable are the results from functional magnetic resonance imaging? Ann N Y Acad Sci 1191:133–155

Benuzzi F, Pugnaghi M, Meletti S, Lui F, Serafini M, Baraldi P, Nichelli P (2007) Processing the socially relevant parts of faces. Brain Res Bull 74:344–356

Berman MG, Park J, Gonzalez R, Polk TA, Gehrke A, Knaffla S, Jonides J (2010) Evaluating functional localizers: the case of the FFA. NeuroImage 50:56–71

Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. J Neurosci 16:4207–4221

Brodmann K (1909) Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues. Barth JA, Leipzig

Busigny T, van Belle G, Jemel B, Hosein A, Joubert S, Rossion B (2014) Face-specific impairment in holistic perception following focal lesion of the right anterior temporal lobe. Neuropsychologia 56:312–333

Buxton RB, Wong EC, Frank LR (1998) Dynamics of blood flow and oxygenation changes during brain activation: the balloon model. Magn Reson Med 39:855–864

Buxton RB, Uludağ K, Dubowitz DJ, Liu TT (2004) Modeling the hemodynamic response to brain activation. NeuroImage 23:S220–S233

Calder AJ, Young AW (2005) Understanding the recognition of facial identity and facial expression. Nat Rev Neurosci 6:641–651

Chan AWY, Downing PE (2011) Faces and eyes in human lateral prefrontal cortex. Front Hum Neurosci 5(51):1–10

Collins JA, Olson IR (2014) Beyond the FFA: the role of the ventral anterior temporal lobes in face processing. Neuropsychologia 61:65–79

Collins JA, Koski JE, Olson IR (2016) More than meets the eye: the merging of perceptual and conceptual knowledge in the anterior temporal face area. Front Hum Neurosci 10:189. https://doi.org/10.3389/fnhum.2016.00189

Crouzet SM, Thorpe SJ (2011) Low-level cues and ultra-fast face detection. Front Psychol 2:342. https://doi.org/10.3389/fpsyg.2011.00342

D'Esposito M (2010) Why methods matter in the study of the biological basis of the mind: A behavioral neurologist's perspective. In: Reuter-Lorenz PA, Baynes K, Mangun GR, Phelps EA (eds) The cognitive neuroscience of mind: a tribute to Michael S. Gazzaniga. MIT Press, Cambridge, pp 203–221

Dale AM (1999) Optimal experimental design for event-related fMRI. Hum Brain Mapp 8:109–114

de Heering A, Rossion B (2015) Rapid categorization of natural face images in the infant right hemisphere. eLife 4:1–14

Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. NeuroImage 53:1–15

Devlin JT, Russell RP, Davis MH, Price CJ, Wilson J, Moss HE, Tyler LK et al (2000) Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. NeuroImage 11:589–600

Dormal G, Lepore F, Harissi-Dagher M, Albouy G, Bertone A, Rossion B, Collignon O (2015) Tracking the evolution of crossmodal plasticity and visual functions before and after sight restoration. J Neurophysiol 113:1727–1742

Duchaine B, Yovel G (2015) A revised neural framework for face processing. Annu Rev Vis Sci 1:393–416

Duncan KJ, Devlin JT (2011) Improving the reliability of functional localizers. NeuroImage 57(3):1022–1030

Duncan KJ, Pattamadilok C, Knierim I, Devlin JT (2009) Consistency and variability in functional localisers. NeuroImage 46:1018–1026

Embleton KV, Haroon HA, Morris DM, Ralph MA, Parker GJ (2010) Distortion correction for diffusion-weighted MRI tractography and fMRI in the temporal lobes. Hum Brain Mapp 31:1570–1587

Engel SA, Glover GH, Wandell BA (1997) Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cereb Cortex 7:181–192

Fischl B, Sereno MI, Tootell RBH, Dale AM (1999) High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum Brain Mapp 8:272–284

Fox C, Moon S, Iaria G, Barton J (2009) The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study. NeuroImage 44:569–580

Fox CJ, Hanif HM, Iaria G, Duchaine BC, Barton JJ (2011) Perceptual and anatomic patterns of selective deficits in facial identity and expression processing. Neuropsychologia 49:3188–3200

Freud E, Ganel T, Shelef I, Hammer MD, Avidan G, Behrmann M (2017) Three-dimensional representations of objects in dorsal cortex are dissociable from those in ventral cortex. Cereb Cortex 27:422–434

Friston KJ, Price CJ, Fletcher P, Moore C, Frackowiak RS, Dolan RJ (1996) The trouble with cognitive subtraction. NeuroImage 4:97–104

Frost MA, Goebel R (2012) Measuring structural–functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment. NeuroImage 59:1369–1381

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

Gentile F, Rossion B (2014) Temporal frequency tuning of cortical face-sensitive areas for individual face perception. NeuroImage 90:256–265

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, … Van Essen DC (2016) A multi-modal parcellation of human cerebral cortex. Nature 536:171–178

Gobbini MI, Haxby JV (2006) Neural response to the visual familiarity of faces. Brain Res Bull 71:76–82

Golarai G, Ghahremani DG, Whitfield-Gabrieli S, Reiss A, Eberhardt JL, Gabrieli JDE, Grill-Spector K (2007) Differential development of high-level visual cortex correlates with category-specific recognition memory. Nat Neurosci 10:512–522

Gomez J, Barnett MA, Natu V, Mezer A, Palomero-Gallagher N, Weiner KS, Grill-Spector K et al (2017) Microstructural proliferation in human cortex is coupled with the development of face processing. Science 355:68–71

Grienberger C, Konnerth A (2012) Imaging calcium in neurons. Neuron 73:862–885

Grill-Spector K, Weiner KS (2014) The functional architecture of the ventral temporal cortex and its role in categorization. Nat Rev Neurosci 15:536–548

Grill-Spector K, Weiner KS, Kay K, Gomez J (2017) The functional neuroanatomy of human face perception. Annu Rev Vis Sci 3:167–196

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

Haxby JV, Hoffman EA, Gobbini MI (2000) The distributed human neural system for face perception. Trends Cogn Sci 4:223–233

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425–2430

Huth AG, Nishimoto S, Vu AT, Gallant JL (2012) A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76:1210–1224

Ishai A, Schmidt CF, Boesiger P (2005) Face perception is mediated by a distributed cortical network. Brain Res Bull 67:87–93

Jokisch D, Jensen O (2007) Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. J Neurosci 27:3244–3251

Jonas J, Rossion B, Brissart H, Frismand S, Jacques C, Hossu G, Maillard L et al (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, Jacques C, Liu-Shuang J, Brissart H, Colnat-Coulbois S, Maillard L, Rossion B (2016) A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. Proc Natl Acad Sci USA 113:E4088–E4097

Kanwisher N (2017) The quest for the FFA and where it led. J Neurosci 37:1056–1061

Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 7:4302–4311

Kim J-J, Crespo-Facorro B, Andreasen NC, O'Leary DS, Zhang B, Harris G, Magnotta VA (2000) An MRI-based parcellation method for the temporal lobe. NeuroImage 11:271–288

Koenig-Robert R, VanRullen R, Tsuchiya N (2015) Semantic wavelet-induced frequency-tagging (SWIFT) periodically activates category selective areas while steadily activating early visual areas. PLoS One 10:e0144858

Kovacs G (2005) Electrophysiological correlates of visual adaptation to faces and body parts in humans. Cereb Cortex 16:742–753

Kriegeskorte N, Formisano E, Sorger B, Goebel R (2007) Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc Natl Acad Sci USA 104:20600–20605

Krüger G, Glover GH (2001) Physiological noise in oxygenation-sensitive magnetic resonance imaging. Magn Reson Med 46:631–637

Lafer-sousa R, Conway BR, Kanwisher NG (2016) Color-biased regions of the ventral visual pathway lie between face- and place-selective regions in humans, as in macaques. J Neurosci 36:1682–1697

Lochy A, van Belle G, Rossion B (2015) A robust index of lexical representation in the left occipito-temporal cortex as evidenced by EEG responses to fast periodic visual stimulation. Neuropsychologia 66:18–31

Loffler G, Yourganov G, Wilkinson F, Wilson HR (2005) fMRI evidence for the neural representation of faces. Nat Neurosci 8:1386–1391

Maus B, van Breukelen GJP, Goebel R, Berger MPF (2010) Optimization of blocked designs in fMRI studies. Psychometrika 75:373–390

McCarthy G, Spicer M, Adrignolo A, Luby M, Gore JC, Allison T (1994) Brain activation associated with visual motion studied by functional magnetic resonance imaging in humans. Hum Brain Mapp 2:234–243

McCarthy G, Puce A, Gore JC, Allison T (1997) Face-specific processing in the human fusiform gyrus. J Cogn Neurosci 9:605–610

McKeefry D, Zeki S (1997) The position and topography of the human color centre as revealed by functional magnetic resonance imaging. Brain 120:2229–2242

Murphy K, Bodurka J, Bandettini PA (2007) How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. NeuroImage 34:565–574

Nasr S, Tootell RB (2012) Role of fusiform and anterior temporal cortical areas in facial recognition. NeuroImage 63:1743–1753

Nichols TE, Das S, Eickhoff SB, Evans AC, Glatard T, Hanke M … Yeo BTT (2017) Best practices in data analysis and sharing in neuroimaging using MRI. Nat Neurosci 20:299–303

Norcia AM, Appelbaum LG, Ales JM, Cottereau B, Rossion B (2015) The steady-state visual evoked potential in vision research: a review. J Vis 15(6):4:1–46

Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci USA 87:9868–9872

Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K (1992) Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci USA 89:5951–5955

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113

Potter MC (2012) Recognition and memory for briefly presented scenes. Front Psychol 3:1–9

Puce A (1999) Electrophysiological studies of human face perception III: effects of top-down processing on face-specific potentials. Cereb Cortex 9:445–458

Puce A, Allison T, Gore JC, McCarthy G (1995) Face-sensitive regions in human extrastriate cortex studied by functional MRI. J Neurophysiol 74:1192–1199

Rajimehr R, Young JC, Tootell RB (2009) An anterior temporal face patch in human cortex predicted by macaque maps. Proc Natl Acad Sci USA 106:1995–2000

Regan D (1989) Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine. Elsevier, New York

Retter TL, Rossion B (2016) Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. Neuropsychologia 91:9–28

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(26):8837–8844

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):1–21

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

Rossion B, Torfs K, Jacques C, Liu-Shuang J (2015) Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. J Vis 15(18):1–18

Rossion B, Jacques C, Jonas J (2018) Mapping face categorization in the human ventral occipito-temporal cortex with direct neural intracranial recordings. Ann N Y Acad Sci

Rousselet GA, Husk JS, Bennett PJ, Sekuler AB (2008) Time course and robustness of ERP object and face differences. J Vis 8(3):1–18

Sadr J, Sinha P (2004) Object recognition and random image structure evolution. Cogn Sci 28:259–287

Scherf KS, Behrmann M, Humphreys K, Luna B (2007) Visual category-selectivity for faces places and objects emerges along different developmental trajectories. Dev Sci 10:F15–F30

Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 268(5212):889–893

Sergent J, Ohta S, MacDonald B (1992) Functional neuroanatomy of face and object processing. Brain 115:15–36

Simoncelli EP, Olshausen BA (2001) Natural image statistics and neural representation. Annu Rev Neurosci 24:1193–1216

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Matthews PM et al (2004) Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23:S208–S219

Smith SM, Jenkinson M, Beckmann C, Miller K, Woolrich M (2007) Meaningful design and contrast estimability in FMRI. NeuroImage 34:127–136

Susilo T, Duchaine B (2013) Advances in developmental prosopagnosia research. Curr Opin Neurobiol 23:423–429
Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. Nature 381:520–522
Tootell RB, Reppas JB, Dale AM, Look RB, Sereno MI, Malach R, Rosen BR et al (1995) Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. Nature 375:139–141
Tsao DY, Moeller S, Freiwald WA (2008) Comparing face patch systems in macaques and humans. Proc Natl Acad Sci USA 105:19514–19519
Tuladhar AM, Huurne N, ter Schoffelen JM, Maris E, Oostenveld R, Jensen O (2007) Parieto-occipital sources account for the increase in alpha activity with working memory load. Hum Brain Mapp 28:785–792
VanRullen R (2006) On second glance: still no high-level pop-out effect for faces. Vis Res 46:3017–3027
Visser M, Embleton KV, Jefferies E, Parker GJ, Ralph MA (2010) The inferior, anterior temporal lobes and semantic memory clarified: novel evidence from distortion-corrected fMRI. Neuropsychologia 48:1689–1696
Von Der Heide RJ, Skipper LM, Olson IR (2013) Anterior temporal face patches: a meta-analysis and empirical study. Front Hum Neurosci 7:17. https://doi.org/10.3389/fnhum.2013.00017
Wandell BA (2011) The neurobiological basis of seeing words. Ann N Y Acad Sci 1224:63–80
Wandell BA, Winawer J (2011) Imaging retinotopic maps in the human brain. Vision Res 51:718–737
Winawer J, Witthoft N (2015) Human V4 and ventral occipital retinotopic maps. Vis Neurosci 32:(E020)
Wang Y-F, Liu F, Long Z-L, Duan X-J, Cui Q, Yan JH, Chen H-F (2014) Steady-state BOLD response modulates low frequency neural oscillations. Sci Rep 4(7376):1–7
Wang Y-F, Dai G-S, Liu F, Long Z-L, Yan JH, Chen H-F (2015) Steady-state BOLD response to higher-order cognition modulates low-frequency neural oscillations. J Cogn Neurosci 27:2406–2415
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, Jonas J, Gomez J, Maillard L, Brissart H, Hossu G, Rossion B et al (2016) The face-processing network is resilient to focal resection of human visual cortex. J Neurosci 36:8425–8440
Welvaert M, Rosseel Y (2013) On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data. PLoS One 8:e77089
Worsley KJ, Marrett S, Neelin P, Evans AC (1996) Searching scale space for activation in PET images. Hum Brain Mapp 4:74–90
Yang H, Susilo T, Duchaine B (2016) The anterior temporal face area contains invariant representations of face identity that can persist despite the loss of right FFA and OFA. Cereb Cortex 26:1096–1107
Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. Trends Cogn Sci 6:37–46
Zhen Z, Yang Z, Huang L, Kong X, 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
Zilles K, Amunts K (2013) Individual variability is not noise. Trends Cogn Sci 17:153