



# A single glance at natural face images generate larger and qualitatively different category-selective spatio-temporal signatures than other ecologically-relevant categories in the human brain



Corentin Jacques<sup>a,b,1</sup>, Talia L. Retter<sup>a,b,c,1</sup>, Bruno Rossion<sup>a,b,\*</sup>

<sup>a</sup> Psychological Sciences Research Institute (IPSY), University of Louvain, 1348 Louvain-la-Neuve, Belgium

<sup>b</sup> Institute of Neuroscience, University of Louvain, 1348 Louvain-la-Neuve, Belgium

<sup>c</sup> Department of Psychology/296, University of Nevada, Reno, NV 89557, United States

## ARTICLE INFO

### Article history:

Received 13 January 2016

Revised 5 April 2016

Accepted 19 April 2016

Available online 30 April 2016

### Keywords:

Visual perception

Face perception

Electroencephalogram

Occipito-temporal cortex

Body part perception

Scene perception

## ABSTRACT

Although humans discriminate natural images of faces from other categories at a single glance, clarifying the neural specificity and spatio-temporal dynamics of this process without low-level visual confounds remains a challenge. We recorded high-density scalp electroencephalogram while presenting natural images of various objects at a fast periodic rate (5.88 images/s). In different stimulation sequences, numerous variable exemplars of three categories associated with cortical specialization in neuroimaging – faces, body parts, or houses – appeared every five images (5.88 Hz/5 = 1.18 Hz). In these fast periodic visual stimulation (FPVS) sequences, common low- and high-level visual processes between these categories and other objects are captured at the 5.88 Hz frequency, while high-level category-selective responses are objectively quantified at the 1.18 Hz frequency and harmonics. Category-selective responses differed quantitatively and qualitatively between faces, body parts and houses. First, they were much larger (2–4 times) for faces over the whole scalp. Second, specific and reliable scalp topographical maps of category-selective responses pointed to distinct principle neural sources for faces (ventral occipito-temporal), body parts (lateral occipito-temporal) and houses (dorso-medial occipital). Category-selective EEG responses were found at multiple time-windows from 110 to 600 ms post-stimulus onset. Faces elicited the most complex spatio-temporal profile with up to four selective responses, although body parts and houses also elicited selective responses more complex than previously described. These observations indicate that a single glance at natural face images inserted in a rapid stream of natural objects generates a quantitatively and qualitatively unique category-selective spatio-temporal signature in occipito-temporal cortical areas of the human brain.

© 2016 Elsevier Inc. All rights reserved.

## Introduction

The human brain distinguishes between natural (i.e., unsegmented) images of faces and non-faces accurately and rapidly, such that gaze immediately and automatically shifts to a face picture in forced choice paradigms for example (Crouzet and Thorpe, 2010, 2011 see also Lewis and Edmonds, 2003; Rousselet et al., 2003). This impressive face categorization ability owes to extensive experience of the human brain with faces throughout development, as well as the particular relevance and ubiquity of face stimuli in the environment. The ease and simplicity of face categorization masks the complexity of the neural processes subtending this function, which requires discrimination of

faces from competing shapes and generalization across face exemplars varying widely in size, orientation, viewpoint, etc.

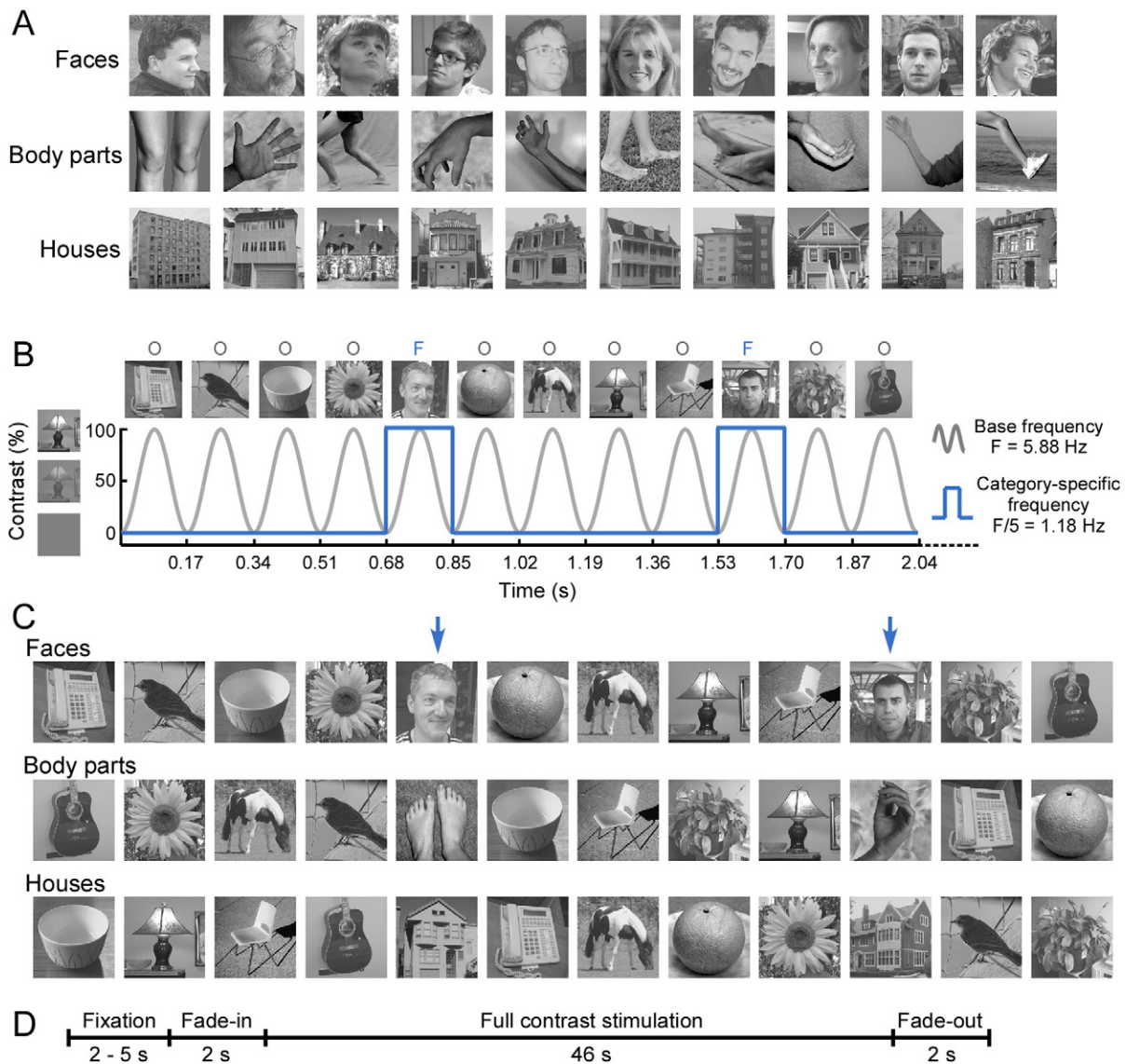
A recently developed approach to understand how the human brain categorizes natural face images at a glance is to record brain activity with high temporal resolution techniques, such as electro/magnetoencephalography (EEG/MEG), while presenting stimuli at a fast periodic rate (Rossion et al., 2015). For instance, natural images of non-face object categories are presented at a rate of 5.88 images/s (170 ms stimulus onset asynchrony, SOA), with variable exemplars of faces appearing every five stimuli (i.e., 5.88 Hz/5: 1.18 Hz; Fig. 1, Supplementary movie 1).

This fast periodic visual stimulation (FPVS) approach increases understanding of face categorization by identifying responses to natural face images objectively (i.e., at a pre-defined 1.18 Hz frequency), rapidly (i.e., in a few minutes), and directly (i.e., without post-hoc subtraction of responses elicited by other categories) (de Heering and Rossion, 2015; Rossion et al., 2015). Importantly, each image is presented briefly, such that responses are captured within a single gaze fixation. As in

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

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

<sup>1</sup> CJ and TR made equal contributions to this work.



**Fig. 1.** Fast periodic visual stimulation (FPVS) paradigm. **A.** Example of natural images of faces, body parts and houses used in the experiment. **B.** Images were presented by sinusoidal contrast modulation at a rate of 5.88 stimuli per second (5.88 Hz). Natural stimuli were randomly selected from a large pool of images of various man-made and natural (living and non-living) objects and different images of either faces (as illustrated here), body parts or houses were presented every 5 stimuli (i.e., appearing at the frequency of  $5.88 \text{ Hz}/5 = 1.18 \text{ Hz}$ ). **C.** Examples of twelve 5.88 Hz cycles for the three experimental conditions where images of either faces, body parts or houses were used as category-specific stimuli in different sequences. **D.** Timeline of a whole stimulation sequence, which started by a fixation cross displayed on a gray background for 2 to 5 s (random duration), followed by a 50 s sequence of image presentation. The image presentation sequence started by a 2 s face-in period and ended with a 2 s fade-out period (only the 46 s of full-contrast stimulation were analyzed).

rapid serial visual presentation (RSVP, [Potter and Levy, 1969](#)), each image is also forward- and backward-masked by other images in the stimulation stream. Moreover, thanks to periodicity constraints and the wide variety of natural images used, low-level property differences between faces and non-faces, such as information contained in the Fourier 2D amplitude spectrum which can account for early saccades towards natural face images ([Crouzet and Thorpe, 2011](#); [Honey et al., 2008](#)) and early EEG/MEG category-selective responses in standard stimulation paradigms (e.g., [Carlson et al., 2013](#); [Cauchoux et al., 2014](#)), do not appear to contribute to the EEG face-selective response in FPVS ([Rossion et al., 2015](#); [de Heering and Rossion, 2015](#)).

Here, we investigate to what extent the category-selectivity of responses to faces found with the FPVS-EEG approach is unique, quantitatively and qualitatively, to faces. To do so we compare rapid periodic presentation of faces to two ecologically-relevant categories, body parts and houses/visual scenes, which are both associated with selective responses in the ventral occipito-temporal cortex (VOTC) as identified in functional magnetic resonance imaging (fMRI) studies ([Downing](#)

[et al., 2006](#); [Grill-Spector and Weiner, 2014](#)). Providing that we observe a category-selective response to each of these categories relative to other (living and non-living) objects, we examine whether responses to faces differ in terms of their overall amplitude, as well as to what extent they can be distinguished in terms of their spatio-temporal signature.

Using FPVS, a category-selective response, i.e. a contrast between responses to one visual category and other objects, can be identified objectively in the frequency domain exactly at the frequency of the category-specific stimulation and harmonics. Moreover, this response can be quantified directly by measuring signal amplitude only at these frequency bins. Such quantification is difficult with more traditional event-related (ERP) approaches and time-domain analyses because differences between conditions may be spread over several ERP components which vary in timing and polarity, making identification and quantification of these responses ambiguous. For instance, consistent evidence accumulated for about 20 years indicates that an ERP component occurring between 130 and 200 ms following a transient and

abrupt visual stimulus presentation, the N170, is larger in amplitude in response to images of faces compared to images of other objects (e.g., Bentin et al., 1996; Cauchoix et al., 2014; Eimer, 2000; Ganis et al., 2012; Itier and Taylor, 2004; Jeffreys, 1989; Rossion et al., 2000; Rousselet et al., 2008; Rossion and Jacques, 2008 for review). However, this amplitude difference can vary substantially across studies depending on stimulation, recording and analysis parameters and, to our knowledge, has not been quantified. Moreover, in traditional ERP paradigms, face-selective responses may be present at other time-windows (e.g. the P1 component peaking at about 100 ms post-stimulus onset, i.e. before the N170; Goffaux et al., 2003; Halgren et al., 2000; Itier and Taylor, 2004; Liu et al., 2002), but they are not consistent across individuals and studies and may be masked by dominant low- and high-level general processes that follow the abrupt onset of visual stimuli. Here we quantify category-selective responses as the sum of harmonics of the 1.18 Hz response in the frequency-domain (e.g., Dzhelyova and Rossion, 2014; Liu-Shuang et al., 2016). This allows testing the hypothesis that category-selectivity, objectively quantified and assessed with a direct measurement of brain activity over the whole scalp, is larger for faces than for other ecologically-relevant categories.

In addition, we aim to characterize in space and time the face-selective response in contrast to body part- and house-selective responses. While we previously identified three novel face-selective components in this paradigm, “P1-faces” (160 ms peak), “N1-faces” (220 ms) and “P2-faces” (410 ms) (Rossion et al., 2015), these components were only visually described at a maximally responding right occipito-temporal site. Here, we characterize category-selective responses across the entire scalp and time-course following the brief presentation of natural face images, and statistically compare these responses and putative category-selective responses to body parts and houses.

## Material and methods

### Participants

Eleven volunteers (7 females, all right handed, mean age = 24 years, SD = 4.32 years) received financial compensation in exchange for their participation in the experiment. All participants gave written informed consent and the experiment was approved by the Biomedical Ethical Committee of the University of Louvain. All participants reported normal or corrected-to-normal vision.

### Stimuli

We used 200 natural images of various objects (from 14 non-face categories: cats, dogs, horses, birds, flowers, fruits, vegetables, houseplants, phones, chairs, cameras, dishes, guitars and lamps), 50 natural images of faces, 50 natural images of upper and lower human body parts, and 50 natural images of houses (Fig. 1A; see also Rossion et al., 2015; for the full set of face stimuli: <http://face-categorization-lab.webnode.com/resources/natural-face-stimuli/>). Importantly, all objects and faces/body parts/houses were unsegmented from the background, i.e., embedded in their original visual scene. Each image contained a single item, which was depicted at the center of the image. Across images, objects/faces differed in terms of size, viewpoint, lighting conditions and background. The stimuli were converted to grayscale, resized to 200 × 200 pixels, and equalized for mean pixel luminance and contrast (across-pixel standard deviation) in Matlab (The Mathworks).

### FPVS procedure

During EEG recording, participants were seated comfortably at a distance of 1 m from the CRT computer screen with a refresh rate of 100 Hz. They were asked to fixate sequences of natural images

appearing on the screen. Images sustained  $5.2 \times 5.2^\circ$  of visual angle. In each stimulus sequence, which lasted 50 s, stimuli were presented through sinusoidal contrast modulation at a rate of 5.8791 Hz (rounded here to 5.88 Hz) (Fig. 1B) using a custom Matlab software running over Psychtoolbox. Within a sequence, each stimulation cycle lasted 170 ms (i.e., 1000 ms/5.88) and began with a uniform gray background from which an image appeared as its contrast increased following a sinusoidal function. Maximal contrast was reached at cycles between 80 and 90 ms and then decreased at the same rate. At this rate, and with images being visible at low-contrast (e.g., 20–30%, which respectively describes 56–67% of each stimulation cycle here; Schneider et al., 2007), the stimulation appeared as continuous to the subjects (Supplementary movies 1 to 3). A basic stimulation unit consisted of five stimulation cycles (i.e.,  $5 \times 170 \text{ ms} = 850 \text{ ms}$ ), in which images of four objects taken from the set of 14 categories were followed by an image of either a face, a body part, or a house (Fig. 1C). Thus, in a sequence, the category-specific stimuli (i.e. faces/body parts/houses) were presented at a frequency of  $5.88 \text{ Hz}/5 = 1.18 \text{ Hz}$  (actual frequency: 1.1758 Hz) (Liu-Shuang et al., 2014). All images were randomly selected from the pool of images of their respective categories.

The experiment therefore consisted of three types of sequences, which formed three conditions: (1) face, (2) body part, and (3) house (Fig. 1C; Supplementary movies 1 to 3). Each of the three main conditions was repeated two times for the first five subjects and three times for the six remaining subjects. Conditions were presented in random order.

Before each 50 s sequence, a fixation cross was displayed against the gray background for 2–5 s (duration randomly jittered between sequences) in order to stabilize participants' fixation. The stimulation sequence started by 2 s of fade-in and ended by two seconds of fade-out. During the fade-in, the contrast modulation depth of the periodic stimulation progressively increased from 0 to 100% (full-contrast), while the opposite manipulation was applied during the fade-out. The fading time aimed at reducing blinks and abrupt eye-movements due to the sudden appearance or disappearance of flickering stimuli. Therefore, a full-contrast sequence lasted for 46 s. A long sequence duration produced a high frequency resolution, which improved the isolation of the periodic response of interest in a discrete frequency bin and enhanced its signal-to-noise ratio relative to the background EEG noise distributed throughout the spectrum (Norcia et al., 2015; Regan, 1989; Rossion et al., 2012b).

During the presentation of sequences, subjects were instructed to fixate on a small red cross located in the center of the stimulus sequence while continuously monitoring the stimuli. Subjects' task was to detect 500 ms duration color-changes of this fixation-cross (red to blue). Color-changes randomly occurred 10 times within every sequence. This task was orthogonal to the manipulation of interest in the study and was used to encourage participants to maintain a constant level of attention throughout the experiment.

The sequence was thus composed of two distinct frequencies: (1) a base frequency that corresponded to the frequency of appearance of all images and reflected visual processing common to all images (i.e. a general visual response), (2) a category-specific frequency that corresponded to the frequency of appearance of the stimuli from the three categories and reflected the processes specifically recruited by each of these categories (i.e. a category-selective response). Hence, the FPVS approach used here allowed identifying and separating two distinct types of responses in the EEG signal: (1) a *general visual response* occurring at the base stimulation frequency (5.88 Hz) and its harmonics and (2) a *category-selective response* to face, body part or house images at the category-specific frequency (1.18 Hz) and harmonics. Specifically, in this approach, changes in low-level visual properties such as in spatial frequency amplitude spectrum, local contrast, boundaries, texture, etc. projected to the base frequency (i.e. 5.88 Hz). Moreover, responses to shape attributes that are not specific to a particular category were also captured in the EEG at the base

frequency. Importantly, the numerous natural images presented vary widely in terms of shape and low-level properties. Since these properties did not vary at the specific periodic frequency of face/body part/house presentation (1.18 Hz), they did not contribute to the category-selective response measured at this frequency and harmonics. In contrast, in a given sequence, because the category of interest was the only one that appeared periodically, populations of neurons specific to this category were activated at the category-specific frequency (i.e. 1.18 Hz), and this response was captured at this frequency only. Therefore, as mentioned in the introduction in FPVS a category-selective response may be generated and measured directly in the EEG signal at the 1.18 Hz frequency and harmonics, i.e. without requiring post-hoc subtraction of neural responses to different categories (Liu-Shuang et al., 2014; Rossion et al., 2015).

#### EEG recording

EEG was recorded using a 160-channel Biosemi ActiveTwo system (Biosemi, Amsterdam, Netherlands) ([http://www.biosemi.com/pics/cap\\_160\\_layout\\_medium.jpg](http://www.biosemi.com/pics/cap_160_layout_medium.jpg)). For sake of clarity, Biosemi channel labels were converted to 10–5 system channel labels. During recording, the system uses two additional electrodes for reference and ground (CMS, common mode sense, and DRL, driven right leg). EEG analog signal was digitized at a 1024 Hz sampling rate over 24 bits. During recording setup, electrode offset was reduced between  $\pm 20 \mu\text{V}$  for each individual electrode by softly abrading the scalp underneath with a blunt plastic needle and injecting the electrode with saline gel. Eye movements were monitored using four electrodes placed at the outer canthi of the eyes and above and below the right orbit. Recordings were manually initiated when participants showed an artifact-free EEG signal and included at least 5 s of resting-state EEG before stimulation.

#### EEG analyses

##### Preprocessing

EEG analyses were carried out using Letswave 5 (<http://www.nocions.org/letswave5/>) and custom scripts running on Matlab (The Mathworks). Continuous EEG data were low-pass filtered (0–246 Hz using an FFT filter with a Hanning window providing full attenuation over a 10 Hz window), downsampled to 512 Hz, and epoched in 56 s segments (from –4 before stimulation to 2 s after the end of stimulation). Each epoch was then DC corrected, and 50 Hz line-noise was removed at 50, 100 and 150 Hz using an FFT filter (Hanning window, 1 Hz width). Noisy electrodes were then linearly interpolated from 4 immediately surrounding clean channels (1 channel each for 3 participants), and independent component analyses (ICA) with a square matrix was applied to the EEG data to isolate and remove large artifacts generated by eye blinks (which were captured by one component in each subject). Epochs were then re-referenced to a common average reference computed using all channels excluding ocular channels.

##### Frequency domain analyses

Epochs were further segmented so as to contain an exact integer number of category-specific 1.18 Hz cycles beginning 2 s after the onset of the sequence (at the end of the fade-in period) until approximately 48 s after sequence onset, before stimulus fade-out (54 cycles, 23,514 time bins in total  $\approx 46$  s). The resulting epochs were averaged per condition (faces, body parts, houses) to increase signal-to-noise ratio, transformed into the frequency domain using a Fast Fourier Transform (FFT) and the amplitude spectra were computed by taking the modulus of the Fourier coefficients at each frequency bin. The long sequence resulted in spectra with a high frequency resolution of 0.0217 Hz (1/46 s).

Next we determined the range of harmonics of category-specific (1.18 Hz) and base frequency (5.88 Hz) responses to consider for

further analyses, based on group-level data. We first averaged the amplitude spectra across subjects separately for each condition (i.e. grand averages), and then averaged the resulting grand average spectra across all channels. We estimated signal-to-noise (SNR) in all frequency bins of these across-channels grand-averaged amplitude spectra by taking the amplitude at a given frequency divided by the mean amplitude in the 20 surrounding bins (excluding the immediately adjacent bins) (Liu-Shuang et al., 2014; Rossion et al., 2012b, 2015; Srinivasan et al., 1999). Z-scores spectra were computed by taking the difference between the amplitude at a given frequency bin and the mean amplitude in the 20 surrounding bins (excluding the immediately adjacent bins), divided by the standard deviation of amplitudes in these 20 surrounding bins (Liu-Shuang et al., 2014; Rossion et al., 2012b, 2015). The number of harmonic responses considered was constrained by the highest harmonic that was significant (z-score  $> 3.1$ ,  $p < 0.001$ , 1-tailed) in at least one condition and in a continuous block of harmonics. For category-specific frequency harmonics, we considered the first 12 harmonics of the category-specific frequency (i.e. 1.1758 Hz to 14.11 Hz), excluding the 5th and 10th harmonics which were confounded with the base frequency (5.879 Hz and 11.76 Hz). For base frequency harmonics we considered the first 4 harmonics of the base frequency (5.879 Hz, 11.76 Hz, 17.54 Hz, and 23.52 Hz).

We then used this range of harmonics to quantify category-specific and base frequencies of responses across conditions in 3 different analyses. First, we determined whether responses for each condition at relevant harmonics and channels were significantly above noise by computing z-scores on grand averaged amplitude spectra (z-score  $> 3.1$ ,  $p < 0.001$ , 1-tailed, i.e. signal  $>$  noise). Second, to quantify the overall amplitudes of category-specific and base frequency responses and provide a compact description of the data, we summed baseline corrected amplitudes (i.e. for each frequency bin we subtract the mean amplitude in the 20 surrounding bins excluding the immediately adjacent bins, and the local minimum and maximum; Rossion et al., 2015) over relevant harmonics (i.e. for category-specific frequency: first 12 harmonics of 1.18 Hz excluding 5.88 Hz and 11.76 Hz; for base frequency: first 4 harmonics of the base frequency; the same range of harmonics was used across the three conditions) separately in each subject and each condition. Then we used a percentile bootstrap approach (sampling subjects with replacement) to statistically compare conditions two-by-two (10,000 bootstrap samples,  $p < 0.01$ , 2-tailed).

Finally, we used a decoding approach to highlight potential differences across categories in the spatial organization of the neural sources generating scalp responses. Decoding was performed in each subject and classification performances were averaged across subjects. For each subject we recorded at least 2 sequences of EEG data per category. The first sequence was used as the training set and the second sequence was used as the test set for the classifier. For subjects in which we recorded 3 sequences per condition, we averaged two sequences for the training set and used one sequence for the test set, running the decoding analyses on the three possible combinations of sequences in the training and test sets and averaging decoding performance across these three iterations. To avoid general amplitude differences across electrodes driving decoding performance, the training and test sets were separately normalized by z-scoring across categories separately at each electrode. We then use a winner-take-all maximum correlation classifier which predicted the category of the test sequence based on the highest correlation between the topography of the test sequence and the three training topographies. Across subjects, chance level was 33.3%. However, at the single subject level, when only two sequences of data were available ( $N = 5$  subjects), performance was either 100% or 0%, which tended to artificially increase standard deviation across subjects and therefore reduced statistical significance of the decoding performance. In addition to decoding, we visualized topographical differences by first normalizing the topographies in each subject using McCarthy and Wood's method to remove general amplitude differences

across conditions (McCarthy and Wood, 1985; similar results were obtained by normalizing topographies using z-scores).

#### Time-domain analyses

Periodic category-selective responses were also investigated in the time domain in the following way. After re-referencing to a common average reference (see Section 2.5.2), epochs were bandpass filtered (0.1 to 30 Hz, zero phase shift Butterworth filter, order 4). In a separate analysis path, these epochs were further notch-filtered to selectively remove the contribution of the base stimulation frequency and its first four harmonics (5.88 Hz to 23.52 Hz, FFT filter with a Hanning window of 0.1 Hz width) from the time-domain waveforms. Epochs with and without notch filtering were then segmented in shorter epochs centered on the onset latency of the category-specific event and of exactly 1 category-specific cycle duration ( $\sim 170$  to  $\sim 680$  ms). These short epochs were baseline corrected by subtracting the mean amplitude in the  $-170$  to  $0$  ms time-window (corresponding to the duration of one cycle at 5.88 Hz), and then averaged for each subject and each condition separately. Next, we used the averaged 5.88 Hz notched-filtered epochs to determine time-windows where electrophysiological responses were significantly different from zero at the group level. These time-windows were identified by running a percentile bootstrap approach (sampling subjects with replacement, 10,000 bootstrap samples,  $p < 0.01$ , 2-tailed) on every time-sample ( $-170$  to  $680$  ms) and electrode. To minimize the probability of false positives due to the large number of comparisons performed, only significant differences lasting for at least 25 consecutive milliseconds and including a cluster of at least two neighboring electrodes were considered.

## Results

#### Frequency domain

We first explored the data in the frequency domain (Fig. 2), examining the magnitude and scalp distribution of base and category-specific frequency responses at each relevant harmonic (determined on spectra averaged across subjects and across channels, see the Material and methods section) in the face, body part, and house conditions. For each harmonic and each channel, we assessed whether these responses

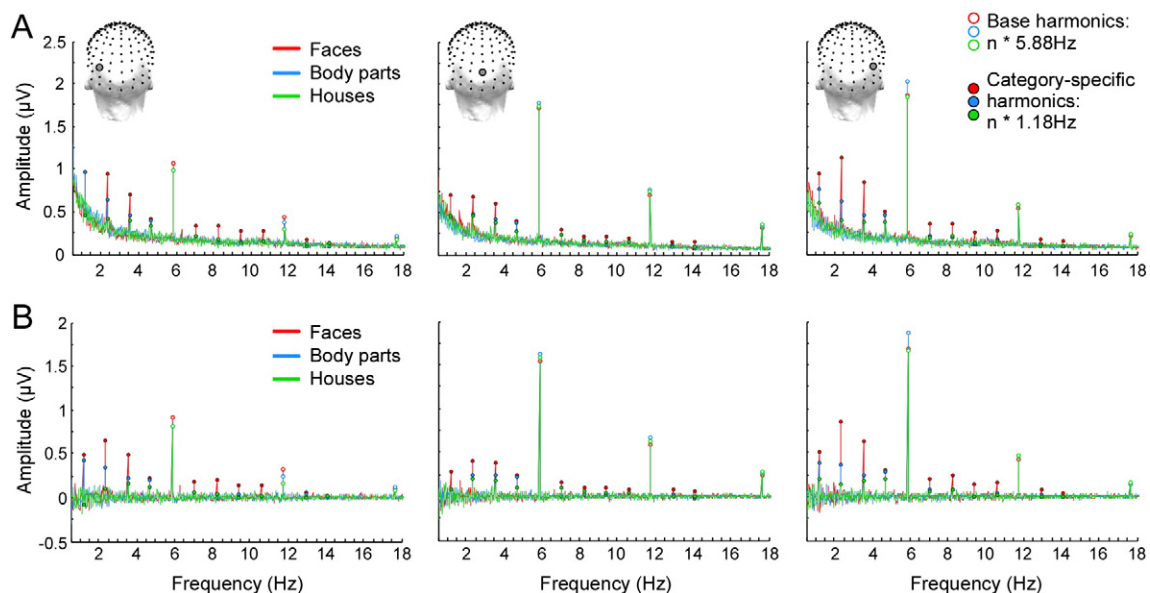
were significantly higher than noise at the group-level using z-score transformed grand-averaged amplitude spectra. Because stimuli are presented periodically, this analysis allowed quantifying EEG responses that were objectively defined in the frequency domain at the exact stimulation frequency and harmonics.

#### Base frequency

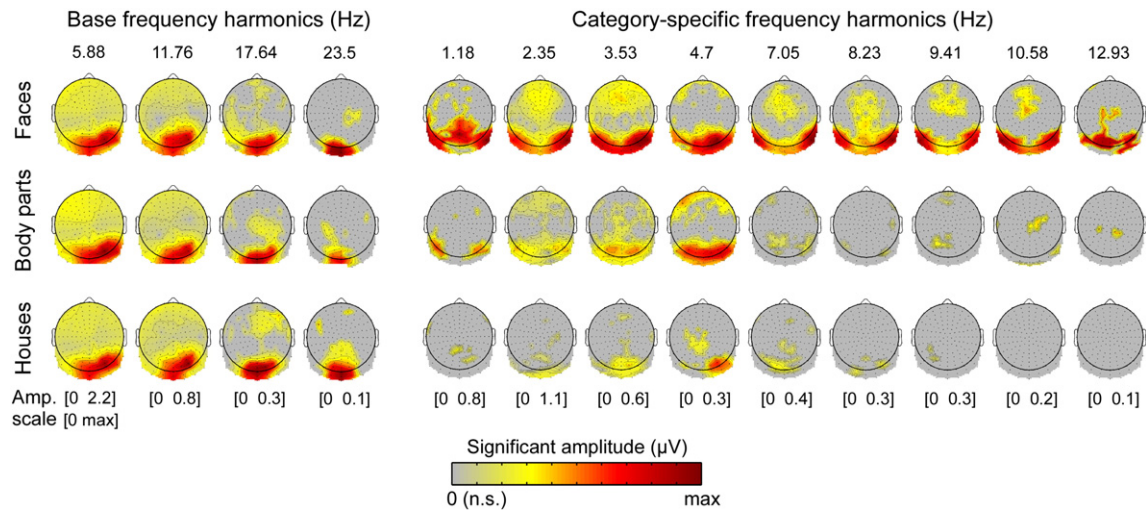
High signal-to-noise ratio (SNR) electrophysiological responses at the four significant harmonics of the *base frequency* (5.88 Hz to 23.52 Hz) were found (SNR  $> 5$  on an average of all 160 channels at the first harmonic for all three conditions; SNR  $> 12$  on the maximal channel, PO6). SNR across individual base frequency harmonics, measured at a medial occipital electrode (Oz), ranged from around 11 to 3 at 5.88 Hz and 23.52 Hz, respectively. These responses did not exhibit amplitude or scalp topography differences across face, body part, and house categories (Figs. 2, 3). However, we observed a dissociation in the scalp distribution of these responses between the lower (i.e. 5.88 Hz and 11.76 Hz) and the higher (i.e. 17.54 Hz, and 23.52 Hz) harmonics. Specifically, while responses at lower harmonics of the base frequency were focused both on medial occipital (maximal at channel Oz) and lateral occipital channels with a right hemispheric dominance (maximal at PO6), responses at higher harmonics were restricted to medial occipital regions. This pattern was also observed for faces in a previous study with the same stimulation frequencies (Rossion et al., 2015). It suggests that the fundamental frequency of the base rate response, at 5.88 Hz, captures common shape-related processes that may occur in the lateral occipital cortex (Grill-Spector et al., 2001), while subsequent higher harmonic frequency responses essentially reflect low-level visual processes taking place in early visual cortex around the occipital pole. This likely occurs because brain regions in higher-level visual cortex tend to respond maximally at lower stimulation frequencies compared to regions in lower-level visual cortex (McKeeff et al., 2007).

#### Category-selective responses

Examining responses at category-specific frequency revealed significant category-selective responses to faces, body parts, and houses distributed over occipital and occipito-temporal regions at multiple harmonics (Figs. 2, 3). We highlight two main observations. First, overall,



**Fig. 2.** Frequency domain representation of EEG signal during FPVS. Grand average frequency domain representations of the EEG recorded in the three different conditions (faces, body parts, and houses) at three occipito-temporal electrodes, of which the position on the scalp is shown on the upper left inserts (back view of the head) as a gray circle (left: P07; center: Oz; right: P08). A. Amplitude spectra. The response at the base frequency (5.88 Hz) and harmonics, as well as category-selective responses at the category-specific frequency (1.18 Hz) and harmonics are visible at all three electrodes and separately highlighted. B. Baseline-subtracted amplitude spectra, used for response quantification.

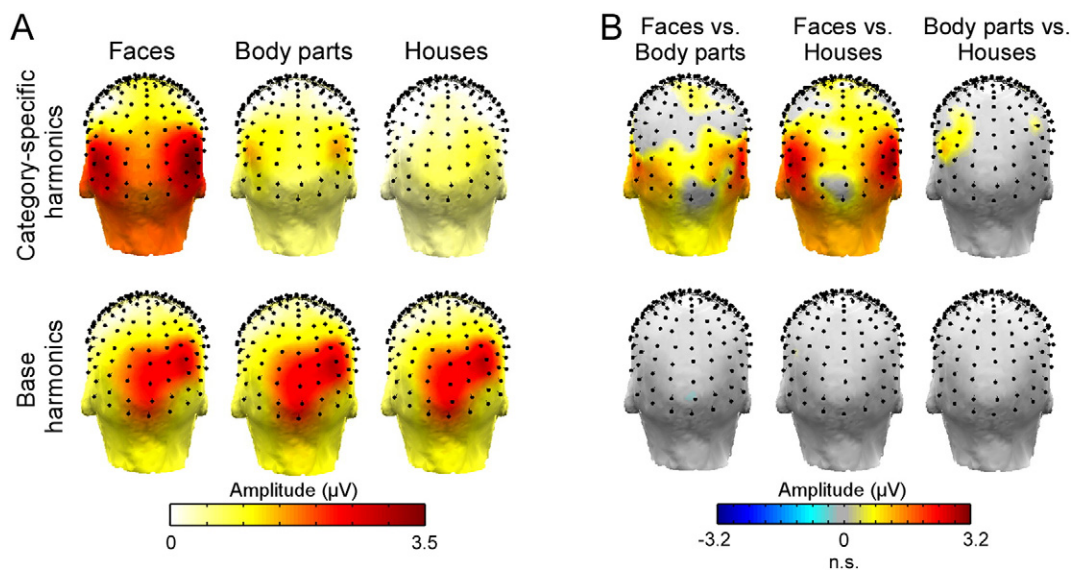


**Fig. 3.** Scalp topographical maps of significant category-specific and base frequency responses for each condition (faces, body parts, and houses) and each harmonic. Maps show the two-dimensional scalp distribution (viewed from above the head, 160 channels) of grand-average significant EEG responses (baseline-subtracted spectra) for each harmonic of the base (left: 5.88–23.5 Hz) and category-specific (right: 1.18–12.93 Hz) frequencies. Conditions are displayed in separate rows and harmonics in separate columns. Significant responses are color-coded as a function of the amplitude of the EEG response. Gray indicates no significant response. The common amplitude scale used for each harmonic is reported for each column below the bottom row of maps.

responses were largest for faces and smallest for houses across all category-specific frequency harmonics. The SNR for the first harmonic at the channel with maximal amplitude for each condition varied accordingly (faces: 5.5 (P10h); body parts: 3.5 (P7); houses: 2 (PO4)), although this was even more evident across harmonics. In fact, responses to body parts and houses were concentrated at lower harmonics with barely any significant response beyond the 4th harmonic (significant harmonics for body parts: 1–4, 6, 8, 9 and for houses: 2–4, 6), whereas faces generated consistent significant responses until the 11th harmonic (significant harmonics for faces: 1–4, 6–9, 11). Second, for the first 4 harmonics, scalp distribution appeared different across categories, with faces generating ventral occipito-temporal responses with right hemispheric dominance, body parts eliciting more bilateral occipito-temporal responses at slightly more dorsal and posterior locations relative to faces, and response to houses being clearly more lateral to medial occipital with a right hemispheric dominance.

*Larger response to faces over occipito-temporal regions.* To quantify differences in the amplitude of category-specific and base frequency responses across categories, we summed baseline-subtracted response amplitudes (see the [Material and methods](#) section) over significant category-specific harmonics ([Fig. 4A](#)) and tested for significant differences between pairs of categories (i.e. faces vs. body parts, faces vs. houses, body parts vs. houses) at each channel ([Fig. 4B](#)).

These analyses confirmed that the magnitude of category-selective response over the occipito-temporal scalp regions was different across categories, with by far the largest response measured for face images ([Fig. 4A](#), upper row). When considering all 160 channels together, the category-selective response to faces (1.15  $\mu\text{V}$ ) was 2.1 times larger than to body parts (0.56  $\mu\text{V}$ ;  $p < 0.001$ , one-tailed paired t-test) and 4.1 times larger than to houses (0.28  $\mu\text{V}$ ;  $p < 0.0001$ ) (body parts vs. houses: body parts 1.0 times larger,  $p < 0.01$ ). The differential category-selective responses (i.e. faces vs. body parts and faces vs.



**Fig. 4.** Scalp topographies of category-selective and general visual response amplitudes. A. Upper row: Scalp topographical maps (back-view of the head) of the amplitude of category-selective responses for each condition (faces, body parts and houses) summed over category-specific frequency harmonics. Lower row: Scalp topographies of the amplitude of general visual response for each condition summed over base frequency harmonics. B. Pairwise statistical comparisons ( $p < 0.01$ , two-tailed percentile bootstrap test) showing the amplitude of the difference across conditions when significant for each category-selective responses and general visual response. Gray indicates no significant difference.

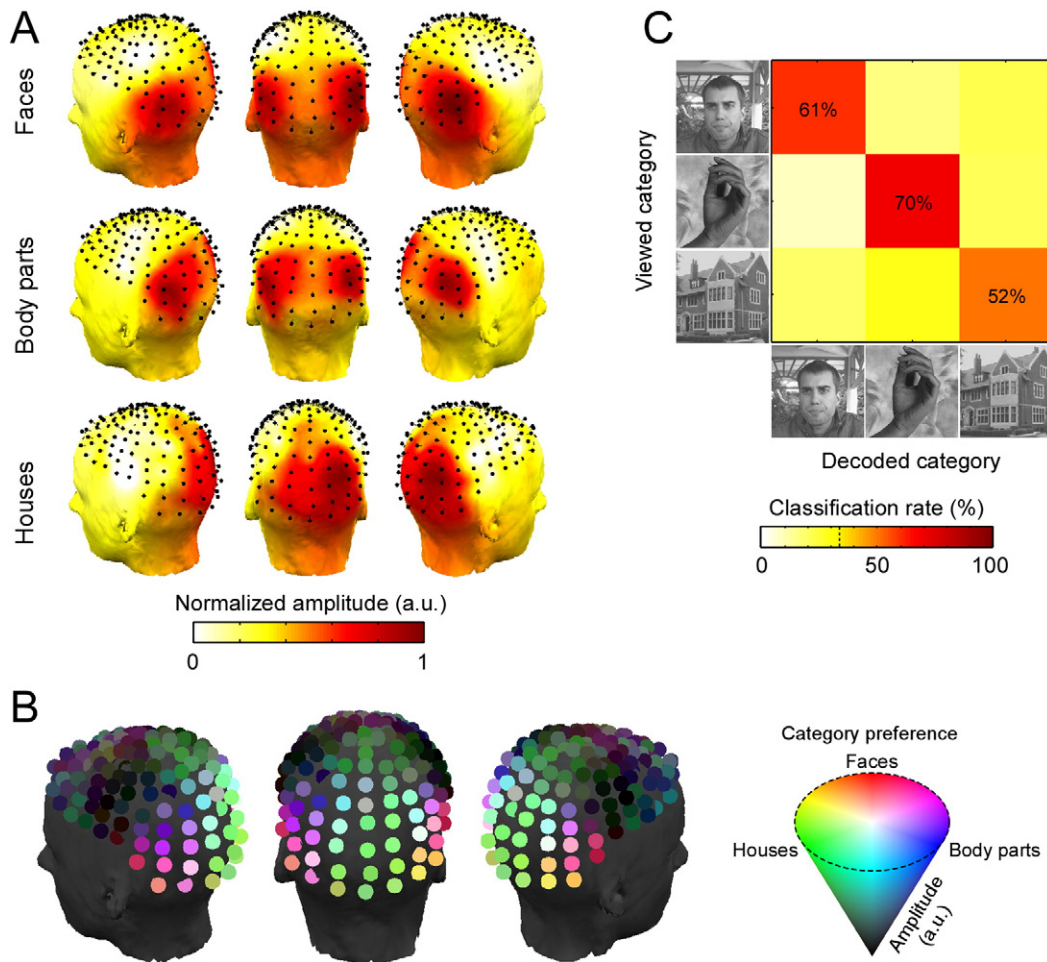
houses) were largest at occipito-temporal channels (i.e., P10/P10h/P8 over the right hemisphere and P9/P9h/P7 over the left hemisphere) (Fig. 4B, upper row). The magnitude of this differential category-selective response was larger over the right hemisphere at ventral channels (PO10/P10,  $p < 0.05$ ) when comparing face to body part, but did not differ between hemispheres when comparing face to house responses. Importantly, the selective response to faces was significantly higher than the selective response to both body parts and houses in every individual subject at right occipito-temporal channels ( $z > 3.1$ ,  $p < 0.001$ , 1-tailed). Finally, comparing responses generated by body part and house stimuli indicates a significantly larger response for body parts mainly over more dorsal occipito-temporal regions in the left hemisphere. This was due to body part stimuli generating medial occipital as well as bilateral occipito-temporal responses, whereas the response to houses was mostly medial occipital and right-lateralized (Fig. 4A, upper row).

Importantly, these categorical differences were restricted to the processing of the category-specific stimuli, as no significant difference was found across categories when comparing base frequency responses summed across the first 4 harmonics (5.88 Hz to 23.52 Hz) (Fig. 4A, B, lower row). As outlined above for individual harmonics, base frequency responses were maximal at medial occipital and right dorsal occipito-temporal sites (Fig. 4A, lower row).

*Images of face, body parts, and houses generate reliable and dissociable category-selective topographical responses. To explore differences in*

scalp topography distribution across categories we computed normalized topographies (McCarthy and Wood, 1985) based on the sum of baseline-subtracted amplitudes across category-specific frequency harmonics. Visual inspection of the normalized topographies (Fig. 5A) indicates clear differences in the scalp distribution of category-selective responses across categories. Face responses were largest at bilateral ventral occipito-temporal channels P8/P10h over the right hemisphere and P7/P9h over the left hemisphere (Fig. 5A, top-row). At these channels, response amplitude was larger over the right compared to the left hemisphere, but this lateralization effect did not reach significance ( $p = .10$ ). Responses to body parts were bilateral (Fig. 5A, middle-row) and, compared to faces, the focus of activation was both more dorsal (i.e. shifted upward) and posterior. Body part responses were maximal at channels PO6/PO8/P6/P8 over the right hemisphere and homologous channels PO5/PO7/P5/P7 over the left hemisphere. For houses, compared to faces and body parts, responses were more dorsal and medial occipital, with a focus on right hemispheric medio-lateral occipital electrodes (O2/PO4/PO6/PO8), and a weak broader distribution over several medial occipital electrodes (Fig. 5A, bottom-row).

These differences in topographies were further highlighted when viewing the relative amplitude for each category across all electrodes (Fig. 5B). They indicated that the spatial organization of the neural sources generating these category-selective topographies are different across categories. To quantify these topographical differences and the underlying neural sources organization, we used a decoding approach.



**Fig. 5.** Faces, body parts and houses generate distinct scalp topographical responses. A. Scalp topographies of the normalized response amplitude (sum over category-specific harmonics) for the three conditions (faces, body parts and houses). B. Category preference for each scalp electrode is color-coded as a function of its relative response amplitude to the three categories. Lighter color indicates higher overall amplitude. C. Confusion matrix showing three-way category decoding based on topographical distribution of category-selective responses during FPVS. Classification rates are averaged across subjects. Rates for correct on-diagonal comparisons are highlighted; chance level is 33%.

We reasoned that if the distributions of amplitudes across the scalp are reliably different across categories, these topographical patterns would allow decoding the category viewed by the subjects using independent data sets. The decoding analysis indicates a significantly above chance category classification for faces (mean hit rate  $\pm$  std.:  $61\% \pm 36\%$ ,  $p < 0.05$ , Fig. 5C) and body parts ( $70\% \pm 41\%$ ,  $p < 0.05$ ) but not significantly above chance for houses ( $52\% \pm 40\%$ ). However, when considering individual category-specific harmonics, we found that only harmonics two to four (2.35 Hz to 4.7 Hz) allowed consistent decoding of the stimulus category. When performing the decoding analysis on the sum of these 3 harmonics, decoding performance increased for faces ( $85\% \pm 31\%$ ,  $p < 0.001$ ) and houses ( $73\% \pm 42\%$ ,  $p < 0.001$ ), and slightly decreased for body parts ( $61\% \pm 49\%$ ,  $p = 0.07$ ). This indicates that, although higher harmonics displayed significant responses, their scalp distribution does not allow to reliably distinguish between category-selective responses.

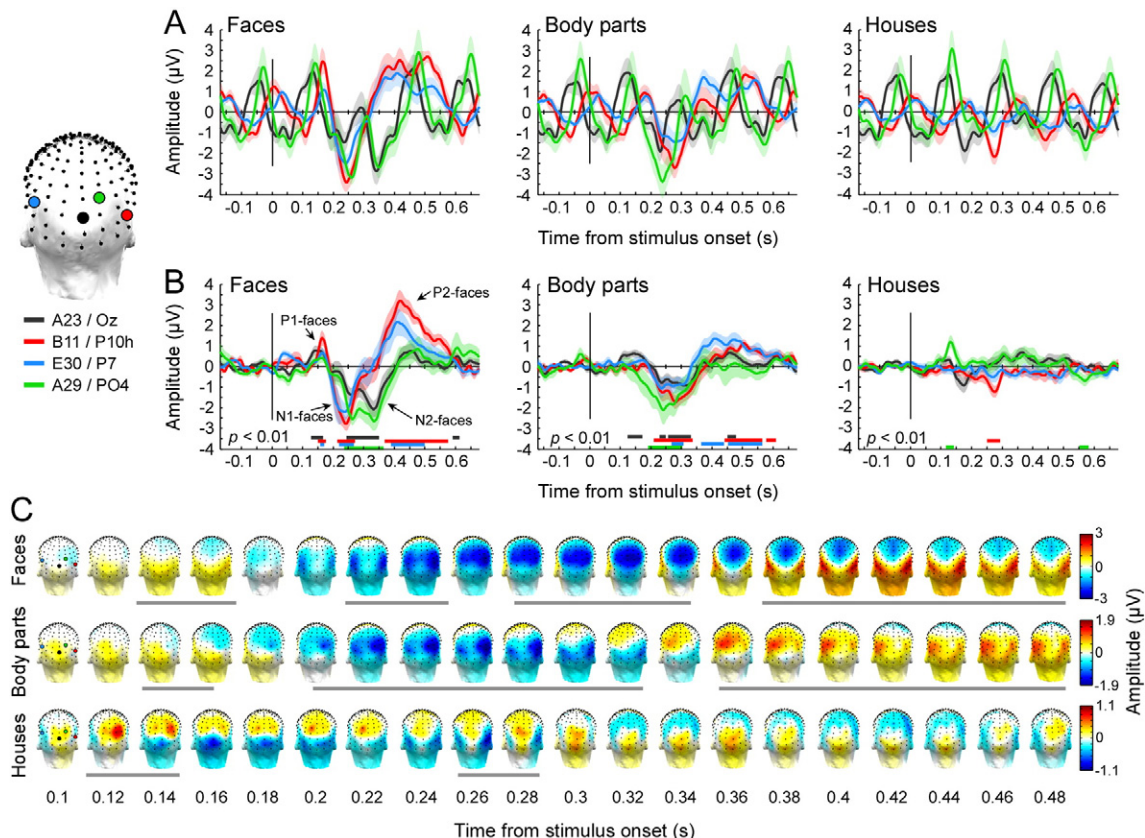
*Time-domain analyses: dissociable spatio-temporal signatures across faces, body parts, and houses*

Next we explored the EEG data in the temporal domain to visualize and investigate the temporal and spatial evolution of the category-selective periodic signals which differentiates across categories.

As expected from the frequency-domain analyses, segments of EEG data in which we did not filter out the signal at base frequency displayed

a strong periodic signal which cycle duration matches the duration of the 5.88 Hz frequency (i.e. 170 ms). These responses were maximal over medial and lateral occipital sites (Fig. 6A, channels Oz and PO4) and much reduced at anterior occipito-temporal sites (channel P10h and P7). In addition, these waveforms revealed responses which are dissociable from the main base frequency responses, and which are time-locked to the onset of the category-specific stimulus. This is exemplified most clearly for faces at ventral occipito-temporal electrode P10h.

To isolate these electrophysiological responses specifically related to the processing of the category-specific stimuli we filtered-out the responses at the base frequency and harmonics (Fig. 6B). In Fig. 6B we display three electrodes that exhibited the highest category-selective responses in the frequency-domain for each category (Faces: P10h, Body parts: P7, Houses: PO4, Fig. 4) as well as an additional midline occipital channel (Oz) to highlight the response at the base stimulation frequency. These waveforms revealed dissociable spatio-temporal dynamics across categories, with the largest and most complex responses being observed for faces (Fig. 6B). For faces we found four main time-windows of interest, which correspond to deflections in the EEG signal (Fig. 6B, C): (1) a brief positive response at around 130–160 ms after stimulus onset, peaking at 150 ms over medial and ventral right occipito-temporal sites; (2) a large negative deflection over bilateral occipito-temporal channels peaking around 230 ms; (3) a second negative deflection maximal from 250 ms to 350 ms with



**Fig. 6.** Time-domain representation of category-selective responses during FVPV. A. Grand averages of the EEG responses relative to the onset of category-specific stimuli (at 0 ms, vertical bar), for the three conditions (columns, faces, body parts, and houses) at four occipito-temporal electrodes. Electrodes' scalp locations are shown on the left with BioSemi labels and corresponding labels from the 10–5 electrode system. Shaded areas are the across-subjects standard error of the mean. The waveforms display a clear periodic signal at the base rate stimulation frequency (5.88 Hz) which is similar across categories. Additional voltage fluctuations occur after the onset of the category-specific stimulus and are distinct across categories. B. The same data is shown after filtering-out the signal at the 5.88 Hz base frequency and harmonics (5.88–23.52 Hz) and highlights the electrophysiological responses specifically elicited by the appearance of the category-specific stimulus. Colored lines below the waveforms indicate time-points at which the filtered signal significantly deviates from baseline ( $p < 0.01$ , two-tailed percentile bootstrap test) for each of the four electrodes shown. C. Spatio-temporal evolution of the category-selective responses triggered by the category-specific stimulus shown in B. Topographical maps (back of the head) are shown for each category (rows) from 0.1 to 0.5 s averaged over successive 20 ms windows and with a separate amplitude color-scale by condition. Gray horizontal lines beneath topographies indicate different components of the responses based on spatial and temporal properties.

a broad medial parieto-occipital scalp distribution; (4) a large positive component peaking at around 420 ms over more anterior and ventral occipito-temporal regions, also showing a right hemisphere advantage.

For *body parts*, although the overall response was much weaker than for faces, we identified three main time-windows with significant responses: (1) a positivity at around 130–160 ms after stimulus onset, with a medial occipital distribution; (2) a negative deflection over bilateral occipito-temporal channels starting around 200 ms, maximal around 250 ms, and finishing by about 320 ms, i.e. slightly later than the first negative component observed for faces; (3) a long duration positive component with a parieto-temporal distribution which started at around 360 ms and persisted until approximately 480 ms.

For *houses*, responses were very small, with two significant time-windows: (1) at around 115–150 ms, peaking at about 130 ms, a very focal positive component which was specifically observed for house stimuli at dorso-lateral occipital sites and restricted to the right hemisphere (the location of this response is similar to the response to house stimuli identified in the frequency analysis, Fig. 5A); (2) at about 260–280 ms, a relatively weak and brief negative component over ventral anterior occipito-temporal channels with a right hemispheric dominance.

## Discussion

Briefly presented natural images of faces generate a large electrophysiological response over the occipito-temporal cortex, with a right hemispheric lateralization, as in the recent study introducing this FPVS paradigm (Rossion et al., 2015). This response reflects discrimination of faces from other objects (living and non-living) and generalization across face exemplars varying substantially in size, viewpoint, orientation, etc., so that low-level contributions to the face-selective response can be eliminated while the naturalness of the images is preserved (i.e., face-selective responses are not present for phase-scrambled images: de Heering and Rossion, 2015; Rossion et al., 2015). Here, the selectivity of the response to faces is further investigated by replacing the face images in the sequence with natural images of either body parts or houses, two categories consistently associated with distinct cortical regions compared to faces in fMRI.

Importantly, we again emphasize that the response identified at the category-specific stimulation frequency and harmonics ( $n1.18$  Hz) is not an absolute response to faces, houses, or body parts (e.g., Meeren et al., 2013), but already a *differential* response to each of these categories relative to a multitude of other object categories. The brain response(s) common with other objects, a mixture of low- and high-level (i.e., shape) visual processes, is projected to the base rate and its harmonics ( $n5.88$  Hz), which did not differ between conditions. While it was previously assumed that the response at  $n1.18$  Hz was unique to faces appearing at that frequency (de Heering and Rossion, 2015; Rossion et al., 2015), it could have been, at least partly observed for other stimuli appearing every five images in the rapid stimulation stream. In addition, while faces images were presented less than non-face object images (20% of images) they were more prevalently represented than any other single object category in their stimulation sequences. However, the present results show that the selective response observed for faces within a FPVS stream is not due to these general factors, and rather genuinely reflects face-selectivity: indeed, despite all other aspects of the stimulation sequences being identical across conditions, this response is distinguishable from selective responses to body parts or houses by its larger amplitude (i.e. a quantitative difference), as well as its distinctive topography and spatio-temporal dynamics (a qualitative difference).

### *Faces generate the largest category-selective response*

Overall, faces generated a much larger (2–4 times) category-selective response than body parts or houses. Since this response

reflects a categorization process, involving visual discrimination and generalization, this finding is in line with, and supports, behavioral evidence that faces are categorized as such more easily and rapidly than other ecologically-relevant categories (e.g., Crouzet and Thorpe, 2010; Hershler and Hochstein, 2005, 2006; Hershler et al., 2010).

This finding of a larger category-selective response to face images appears to be robust: it is a particularly large difference (i.e., more than 2 times the response to body parts, more than 4 times the response to houses when considering all channels), and this difference is found for every participant tested in the study. Moreover, even though this category-selective response reflects generalization across images, the larger response to faces is unlikely to depend on factors such as the visual homogeneity of the image set used, a factor that does not account for face-specific EEG responses such as the N170 (Rossion and Jacques, 2008; Ganis et al., 2012) or face-specific visual recognition impairments following brain-damage (prosopagnosia, Busigny et al., 2010). Here, the face within each image purposely varied in terms of size, viewpoint, lighting, background, etc. across the set of images (see Fig. 1A). While the body part images, including hands, arms, legs, and feet, were even more variable than the face images and produced a weaker response, images of houses were comparable to the face images in terms of heterogeneity, yet they produced the weakest response by far.

The larger category-selective response found for faces is also unlikely to be explained by the fast rate of visual stimulation used (5.88 Hz, 170 ms SOA). While each stimulus of the three categories of interest was sandwiched in between images of other objects, with a SOA allowing only one gaze/fixation per image, this duration does not prevent clearly perceiving each image, in agreement with behavioral evidence for perceptual discrimination of complex object or scenes in such 6 Hz rapid serial visual presentation (RSVP) sequences (Potter, 2012). Moreover, the stimuli of interest (faces/body parts/houses) were presented at a frequency low enough to avoid overlap between category-selective responses (i.e., 1.18 Hz, 850 ms SOA). Indeed, clear baseline activity periods were observed in all conditions (Fig. 6B), so that temporally complete category-selective responses appear to be captured in our paradigm. Thus, while we cannot formally exclude that the stimulation rates used here may be optimal to capture face-selective responses but less so for house- or body part-selective responses, these parameters are unlikely to account for our finding of a larger category-selective response for faces. Nevertheless, future studies may build upon this paradigm to define the presentation rates that elicit maximal category-selective responses, and whether these rates are different for faces than for other categories.

The larger category-selective response for faces is not so surprising, given that a human face is a particularly relevant and ubiquitous stimulus in our visual environment, carrying a wide range of information (gender, identity, age, expression, eye gaze direction, etc.), the processing of which calls upon many different visual and memory functions. fMRI and intracranial electrophysiological recordings in humans have revealed that face-selective responses are widely distributed across the whole human OTC, extending from the inferior occipital gyrus to anterior temporal lobe (Allison et al., 1999; Haxby et al., 2000; Jacques et al., 2016; Rossion et al., 2012a; Sergent et al., 1992; Tsao et al., 2008; Grill-Spector and Weiner, 2014; Zhen et al., 2015). Up to six spatially distinct face-selective regions can be found in single individuals in the OTC (Tsao et al., 2008; Rossion et al., 2012a), this number probably being underestimated considering magnetic susceptibility artifacts affecting the anterior half of the temporal lobe in fMRI (Axelrod and Yovel, 2013; Jonas et al., 2015). Moreover, face-selective responses found along the human superior temporal sulcus (STS; Puce et al., 1998; Zhen et al., 2015) may also contribute to the selective electrophysiological response to faces observed on the scalp here.

Relative to faces, the visual representation of body parts and house stimuli on the cortical surface is smaller, which may partly explain the smaller electrophysiological response measured on the scalp. Body parts selectively activate only two clusters of regions in the OTC in

fMRI: the extrastriate body area (EBA) in the lateral OTC (Downing et al., 2001), which may include several smaller subregions (Weiner and Grill-Spector, 2011) and the fusiform body area (FBA), a small region in the posterior fusiform and occipito-temporal sulcus (Peelen and Downing, 2005; Schwarzlose et al., 2005). Similarly, houses selectively activate two main regions in the OTC: the parahippocampal place area PPA in the collateral sulcus and parahippocampal gyrus (PPA, Epstein and Kanwisher, 1998; Nasr et al., 2014) and a region in the transverse occipital sulcus (TOS) in the lateral occipital cortex (Grill-Spector, 2003).

Additionally, brain regions selective for body parts and houses defined in fMRI may be involved in more general processes less restricted to a particular stimulus category. For instance, a distinctive response to body parts might arise from occipito-temporal and parietal regions involved in interpreting and performing biological actions, rather than coding information about people per se (Astaiev et al., 2004). Similarly, house stimuli are generally used as a diagnostic exemplar category for “places” or “scenes” in fMRI, inducing responses likely to be less specific to “houses” per se but having more to do with coding of scene layout or navigability (Epstein et al., 1999; Bastin et al., 2013). Therefore, the selectivity in regions responding to body parts or houses for the set of images used here might not be as sharp as that observed for face images in face-selective regions, resulting in a smaller response to body parts or houses than to faces presented within the context of the set of other objects in our FPVS sequence.

*The selective representations of faces, body parts, and houses differ in space and time*

A pattern classification analysis performed on a compact description of the different category-selective responses (i.e. sum across category-specific harmonics) revealed distinct scalp spatial distributions of responses to the different categories. Selective responses to faces, body parts and houses were respectively maximal at ventral occipito-temporal, lateral occipito-temporal, and dorso-medial occipital channels (Fig. 5). Thanks to the high SNR from FPVS, we were able to show these significant differences despite using a simple correlation-based classifier with only 2–3 stimulation sequences of data per subject. To our knowledge, this is the first evidence of significant and reliable scalp topographical differences between *category-selective* responses to faces, body parts and houses.

Spatio-temporal analyses of the EEG further revealed that faces generated at least four distinct category-selective responses over an extended period of time from 130 ms up to about 600 ms after face onset. Three of these differential components peaking over (right) occipito-temporal sites were described in our previous study (Rossion et al., 2015) as *P1-faces* (120–160 ms), *N1-faces* (200–260 ms) and *P2-faces* (350–600 ms) (Fig. 6B, C). A fourth component (*N2-faces*) emerges significantly between 280 and 340 ms over medial occipito-parietal sites. This rich spatio-temporal signature of face-selective responses contrasts with standard EEG/MEG studies of face and object categorization which record the change of EEG activity to transient presentations of these stimuli (i.e., the abrupt onset of a face or object stimulus preceded and followed by a uniform background). These studies have consistently identified only a single component showing a larger amplitude to images of faces than to images of non-face objects: the N170/M170 (Bentin et al., 1996; Cauchois et al., 2014; Eimer, 2000; Ganis et al., 2012; Itier and Taylor, 2004; Rossion et al., 2000; Rousselet et al., 2008; Rossion and Jacques, 2011 for review) and its positive counter-part at the vertex (the VPP: Jeffreys, 1996; Joyce and Rossion, 2005). This is because early EEG components to slow transient stimulation such as the P1, preceding the N170, are dominated by low-level visual responses (Rossion and Caharel, 2011), while components recorded after 200 ms in this stimulation mode often overlap with eye movements or decisional/attentional and motor processes, making it difficult to isolate the full complexity of face-selective responses with a conventional (i.e., slow, interrupted and non-periodic) stimulation

mode in EEG/MEG. Here, importantly, each image is presented so briefly that is associated with a single fixation, and late confounding processes are not frequency-locked to the stimulation rate.

Interestingly, with its complexity and temporal extent of several hundreds of milliseconds, the face-selective response measured with FPVS fits much better to the complexity and extent of the face-selective responses measured in fMRI and intracranial research than the face-selective response measured in transient ERPs on the scalp, which isolates a single consistent spatio-temporal window of face-selectivity around the N170. As such, the FPVS approach allows to better bridge the gap between measurements of face-selective responses made using scalp electrophysiology and fMRI.

The first face-selective component recorded here (*P1-faces*) starts at about 130 ms and peaks around 150 ms over medial and ventral right occipito-temporal sites, reflecting the onset of the face-selective response in OTC. Considering that the sinusoidal increase of contrast in the visual stimulation peaks over 80–90 ms (Fig. 1B) and that 20 to 30% contrast might be necessary to evoke a face EEG response (Schneider et al., 2007), the earliest face-selective response should be shifted forward by about 30 ms, corresponding to an onset latency of about 100–110 ms and peak around 120 ms.

Despite this early onset, it is important to distinguish this *P1-face* response to a sensitivity to faces sometimes observed at the level of the standard P1 ERP component following transient stimulation (e.g., Goffaux et al., 2003; Halgren et al., 2000; Itier and Taylor, 2004; Liu et al., 2002), an effect which emerges at around 80 ms, i.e. before the peak of the P1, and is largely inconsistent across studies (Rossion and Jacques, 2008). To our knowledge, this P1 sensitivity to faces is not lateralized and appears to be driven entirely by low-level visual cues, in particular the differential spatial frequency amplitude spectra between faces and other stimuli (Tanskanen et al., 2005; Rousselet et al., 2008; Ganis et al., 2012; see Rossion and Caharel, 2011 for P1 face effect obtained with phase-scrambled stimuli).

The *P1-faces* onset latency reported here is earlier than the typical onset of face-selective responses on the scalp at the N170 (i.e., at about 130 ms) usually estimated using a two-by-two subtraction procedure (e.g., faces vs. houses). However, this later latency remains ambiguous as the latency measurement is performed using segmented stimuli rendered artificially similar to each other to control for low-level visual differences across categories (e.g. Rousselet et al., 2008). Here, since we used a large number of variable natural images, obtaining a consistent selective response at the face stimulation frequency means that the response to most face exemplars had to differ from the response to most exemplars of non-face object categories in a sequence. In this respect, our approach is more similar to a decoding approach using multiple categories, which has revealed discrimination of various categories including faces from about 100 ms in scalp recordings (Carlson et al., 2013; Cichy et al., 2014) or in intracranial field potentials recorded in human high-level visual areas (Liu et al., 2009; see also Jacques et al., 2016).

The second face-selective negative component (*N1-faces*) peaks bilaterally around 230 ms over occipito-temporal sites with a right hemisphere advantage. Although it resembles a N170, we should state again that this deflection is an inherent contrast response between faces and other visual stimuli (i.e., a face-selective response), not an event-related potential (ERP) component to the sudden onset of a face from a uniform background. Both of these early face-selective responses (*P1-faces* and *N1-faces*) could be generated by populations of neurons in the inferior lateral occipital cortex, in the posterior superior temporal sulcus and regions in the posterior and middle fusiform gyrus which show larger response to faces over non-faces in fMRI (Rossion et al., 2012a; Tsao et al., 2008; Grill-Spector and Weiner, 2014; Zhen et al., 2015).

The third (negative) face-selective response (*N2-faces*) is observed broadly over medial occipito-parietal regions between 280 and 340 ms. This response may be generated in mesial cortical regions

such as the precuneus or parieto-occipital sulcus, which are involved in recognition of known faces (Avidan and Behrmann, 2009; Gobbi and Haxby, 2007; Jonas et al., 2014) or in regions of the intraparietal sulcus and superior parietal cortex possibly reflecting attentional capture by the saliency of the face category (Kincade, 2005).

Last, we observe a bilateral positive component (P2-faces) over more anterior and ventral occipito-temporal regions, also showing a right hemisphere advantage. The dipolar configuration of this component with a polarity reversal over centro-parietal regions points to cortical sources in the ventral temporal cortex. This component might be similar to the P350 component measured with intracranial EEG mostly over the anterior half of the temporal lobe (Allison et al., 1999).

Body part images generate three distinct category-selective components, including a very early response at 130 ms over medial occipital sites, and two subsequent prolonged components. The initial 130–160 ms positive response is relatively weak and measured at electrodes sites typically associated with low-level visual responses (medial occipital regions). Although such low-level confounds do not contribute to the responses observed for faces in our paradigm (de Heering and Rossion, 2015; Rossion et al., 2015), we cannot yet exclude systematic differences between the natural images of objects and the images of body parts (Fig. 1). The second body part-selective negative response occurs over bilateral occipito-temporal channels and resembles the N1-faces but with a slightly more dorso-lateral focus and later latency. This scalp topography is in agreement with a negative event-related potential selective for human bodies and body parts, peaking around 190 ms after stimulus onset over occipito-temporal regions (the N190; Minnebusch et al., 2009; Kovacs et al., 2006; Taylor et al., 2010; Thierry et al., 2006). It is thought to arise from regions in the lateral occipito-temporal cortex where fMRI activations for human bodies and body parts have been consistently identified (i.e. the extrastriate body area; EBA; Downing et al., 2001; Orlov et al., 2010; Weiner and Grill-Spector, 2011). Here, this negative component progressively shifts to more ventral and anterior regions, over right inferior temporal regions. Such a scalp distribution agrees with the cortical location of the body part-selective region in the lateral fusiform gyrus and occipito-temporal sulcus described both in fMRI (i.e. the fusiform body area; FBA; Peelen and Downing, 2005; Schwarzlouse et al., 2005; Weiner and Grill-Spector, 2010) and intracranial EEG (Engell and McCarthy, 2014; McCarthy et al., 1999). Finally, body parts generate a third, dorsal parieto-temporal component which started with a left hemisphere dominance over parietal and middle/superior temporal regions (at around 360–400 ms), and spreads over both hemispheres at approximately 460–480 ms. This possibly corresponds to activations of cortical regions involved in integrating information to process biological motion and action in the posterior parietal cortex (Grefkes and Fink, 2005) or the posterior superior temporal sulcus and superior temporal gyrus (Puce and Perrett, 2003; Thompson et al., 2005).

In sum, relative to previous reports describing only a single body part-selective ERP component over lateral OTC (the N190; Minnebusch et al., 2009; Taylor et al., 2010; Thierry et al., 2006), we find that body part images presented within FPVS sequences activate a more comprehensive part of the body part-processing brain network spread over the occipito-temporal and posterior parietal cortex. These regions are involved in processing action- and motion-related aspects (Puce and Perrett, 2003; Grefkes and Fink, 2005; Thompson et al., 2005). This is likely because the various natural images presented within an FPVS sequence provide both a form of contextual background in which the images of body parts appear and an impression of motion similar to that experienced when viewing rapidly-cut movie clips (Supplementary movie 2).

Finally, pictures of houses only generate responses in two significant time windows. The first house-selective response is a positive component peaking at about 130 ms after stimulus onset over a focal region of the right posterior lateral occipital cortex. This scalp location is compatible with the cortical location of a house/scene-selective region

in the TOS (Grill-Spector, 2003; Nakamura et al., 2000), which has been causally linked to scene categorization and discrimination using TMS (Dilks et al., 2013). Our data indicate that this region may be activated very early following stimulus onset. The early latency of this response, however, suggests it could also be influenced by low-level cues typical of buildings such as a higher proportion of rectilinear shapes (Nasr et al., 2014) or higher energy at cardinal orientations (Keil and Cristobal, 2000).

The second house-selective time-window contains a relatively weak and brief negative component over ventral anterior occipito-temporal channels with a right hemispheric dominance. The scalp distribution of this component is compatible with a source in the PPA in the COS and PHG regions. In line with this idea, the latency of this component is somewhat similar to that at which the PPA distinguishes images of buildings from images of other objects (onset: ~170 ms, peak: ~340 ms, Bastin et al., 2013). While it is yet unclear whether visually evoked responses arising from such deep and medial brain regions can be measured at the scalp, recent evidence combining scalp and intracranial EEG recordings in mesial brain structures during epileptiform activity indicate that these medial sources can contribute to EEG recorded at the scalp (Koessler et al., 2014).

### Summary and perspectives

With FPVS, we found that briefly presented natural images of variable faces elicit a much larger and more complex category-selective response at the system level in the human brain than images of body parts or houses presented under the exact same stimulation conditions. Given the richness, complexity and prolonged duration of this face-selective response, it is likely that it does not only reflect face categorization, a rapid process, but also further processes recruited automatically. Contrary to other objects, faces are rapidly individualized and they convey a wide range of cues important for social categorization. Although these processes may be subtended by vastly different brain regions, the richness of the information that can be extracted from faces certainly contributes to the large magnitude of this face-selective response and its complex spatio-temporal pattern. The sensitivity of the FPVS approach used here opens an avenue for future research aiming at understanding the functional processes reflected by these multiple face-selective responses. For instance, this approach could be easily used to test how variations in the stimulus space affect each of the face-selective responses (e.g., color vs. grayscale stimuli, full front vs. profile stimuli; occluded vs. full view stimuli), and thus at which point in time these stimulus properties are coded during face-selective processes. Moreover, the impact of memory representations (e.g., familiar vs. unfamiliar faces) on these responses could be explored. Finally, providing slight changes to the paradigm, task effects (i.e., explicit face detection, gender decision, etc.) could also be tested with this FPVS approach, for instance by presenting stimuli at a fast 5.88 Hz rate, but with target categories appearing at a slower rate than 1.18 Hz, or even using a non-periodic presentation of the target stimuli. While a non-periodic presentation would not allow a quantification of the category-selective response in the frequency domain, it would allow characterizing the different spatio-temporal components and their putative task modulations, as performed here.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2016.04.045>.

### Acknowledgments

The authors thank two anonymous reviewers for their constructive critical comments on a previous version of this manuscript. CJ is supported by the Belgian Federal Science Policy Office (BELSPO, return grant 2012). BR and TR are supported by the Belgian National Fund for Scientific Research (Fonds de la Recherche Scientifique; FNRS). This

work was supported by a grant from the European Research Council (ERC, facessvnp 284025). Authors declare no competing interest.

## References

- Allison, T., Puce, A., Spencer, D., McCarthy, G., Belger, A., 1999. Electrophysiological studies of human face perception. I: potential generated in occipitotemporal cortex by face and non-face stimuli. *Cereb. Cortex* 9, 415–430.
- Astafiev, S.V., Stanley, C.M., Shulman, G.L., Corbetta, M., 2004. Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nat. Neurosci.* 7, 542–548. <http://dx.doi.org/10.1038/nn1241>.
- Avidan, G., Behrmann, M., 2009. Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. *Curr. Biol.* 19, 1146–1150. <http://dx.doi.org/10.1016/j.cub.2009.04.060>.
- Axelrod, V., Yovel, G., 2013. The challenge of localizing the anterior temporal face area: a possible solution. *NeuroImage* 81, 371–380. <http://dx.doi.org/10.1016/j.neuroimage.2013.05.015>.
- Bastin, J., Vidal, J.R., Bouvier, S., Perrone-Bertolotti, M., Bénis, D., Kahane, P., David, O., Lachaux, J.-P., Epstein, R.A., 2013. Temporal components in the parahippocampal place area revealed by human intracerebral recordings. *J. Neurosci.* 33, 10123–10131. <http://dx.doi.org/10.1523/JNEUROSCI.4646-12.2013>.
- Bentin, S., McCarthy, G., Perez, E., Puce, A., Allison, T., 1996. Electrophysiological studies of face perception in humans. *J. Cogn. Neurosci.* 8, 551–565.
- Busigny, T., Graf, M., Mayer, E., Rossion, B., 2010. Acquired prosopagnosia as a face-specific disorder: ruling out the general visual similarity account. *Neuropsychologia* 48, 2051–2067. <http://dx.doi.org/10.1016/j.neuropsychologia.2010.03.026>.
- Carlson, T., Tovar, D.A., Kriegeskorte, N., 2013. Representational dynamics of object vision: the first 1000 ms. *J. Vis.* 13 (10), 1–19. <http://dx.doi.org/10.1167/13.10.1> (1).
- Cauchois, M., Barragan-Jason, G., Serre, T., Barbeau, E.J., 2014. The neural dynamics of face detection in the wild revealed by MVPA. *J. Neurosci.* 34, 846–854. <http://dx.doi.org/10.1523/JNEUROSCI.3030-13.2014>.
- Cichy, R.M., Pantazis, D., Oliva, A., 2014. Resolving human object recognition in space and time. *Nat. Neurosci.* 17, 455–462. <http://dx.doi.org/10.1038/nn.3635>.
- Crouzet, S.M., Thorpe, S.J., 2010. Fast saccades toward faces: face detection in just 100 ms. *J. Vis.* 10 (4), 1–17. <http://dx.doi.org/10.1167/10.4.16> (16).
- Crouzet, S.M., Thorpe, S.J., 2011. Low-level cues and ultra-fast face detection. *Front. Psychol.* 2, 342. <http://dx.doi.org/10.3389/fpsyg.2011.00342>.
- de Heering, A., Rossion, B., 2015. Rapid categorization of natural face images in the infant right hemisphere. *Elife* 4, e06564. <http://dx.doi.org/10.7554/eLife.06564>.
- Dilks, D.D., Julian, J.B., Paunov, A.M., Kanwisher, N., 2013. The occipital place area is causally and selectively involved in scene perception. *J. Neurosci.* 33, 1331–1336. <http://dx.doi.org/10.1523/JNEUROSCI.4081-12.2013>.
- Downing, P.E., Jiang, Y., Shuman, M., Kanwisher, N., 2001. A cortical area selective for visual processing of the human body. *Science* 293, 2470–2473. <http://dx.doi.org/10.1126/science.1063414>.
- Downing, P.E., Chan, A.W.-Y., Peelen, M.V., Dodds, C.M., Kanwisher, N., 2006. Domain specificity in visual cortex. *Cereb. Cortex* 16, 1453–1461. <http://dx.doi.org/10.1093/cercor/bhj086>.
- Dzhelyova, M., Rossion, B., 2014. Supra-additive contribution of shape and surface information to individual face discrimination as revealed by fast periodic visual stimulation. *J. Vis.* 14 (14), 15. <http://dx.doi.org/10.1167/14.14.15>.
- Eimer, M., 2000. The face-specific N170 component reflects late stages in the structural encoding of faces. *Neuroreport* 11, 2319–2324.
- Engell, A.D., McCarthy, G., 2014. Face, eye, and body selective responses in fusiform gyrus and adjacent cortex: an intracranial EEG study. *Front. Hum. Neurosci.* 8, 642. <http://dx.doi.org/10.3389/fnhum.2014.00642>.
- Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature* 392, 598–601. <http://dx.doi.org/10.1038/33402>.
- Epstein, R., Harris, A., Stanley, D., Kanwisher, N., 1999. The parahippocampal place area: recognition, navigation, or encoding? *Neuron* 23, 115–125.
- Ganis, G., Smith, D., Schendan, H.E., 2012. The N170, not the P1, indexes the earliest time for categorical perception of faces, regardless of interstimulus variance. *NeuroImage* 62, 1563–1574. <http://dx.doi.org/10.1016/j.neuroimage.2012.05.043>.
- Gobbini, M.I., Haxby, J.V., 2007. Neural systems for recognition of familiar faces. *Neuropsychologia* 45, 32–41. <http://dx.doi.org/10.1016/j.neuropsychologia.2006.04.015>.
- Goffaux, V., Gauthier, I., Rossion, B., 2003. Spatial scale contribution to early visual differences between face and object processing. *Cogn. Brain Res.* 16, 416–424.
- Grefkes, C., Fink, G.R., 2005. The functional organization of the intraparietal sulcus in humans and monkeys. *J. Anat.* 207, 3–17. <http://dx.doi.org/10.1111/j.1469-7580.2005.00426.x>.
- Grill-Spector, K., 2003. The neural basis of object perception. *Curr. Opin. Neurobiol.* 13, 159–166. [http://dx.doi.org/10.1016/S0959-4388\(03\)00040-0](http://dx.doi.org/10.1016/S0959-4388(03)00040-0).
- Grill-Spector, K., Weiner, K.S., 2014. The functional architecture of the ventral temporal cortex and its role in categorization. *Nat. Rev. Neurosci.* 15, 536–548. <http://dx.doi.org/10.1038/nrn3747>.
- Grill-Spector, K., Kourtzi, Z., Kanwisher, N., 2001. The lateral occipital complex and its role in object recognition. *Vis. Res.* 41, 1409–1422.
- Halgren, E., Raji, T., Marinkovic, K., Jousmaki, V., Hari, R., 2000. Cognitive response profile of the human fusiform face area as determined by MEG. *Cereb. Cortex* 10, 69–81.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Hershler, O., Hochstein, S., 2005. At first sight: a high-level pop out effect for faces. *Vis. Res.* 45, 1707–1724.
- Hershler, O., Hochstein, S., 2006. With a careful look: still no low-level confound to face pop-out. *Vis. Res.* 46, 3028–3035. <http://dx.doi.org/10.1016/j.visres.2006.03.023>.
- Hershler, O., Golan, T., Bentin, S., Hochstein, S., 2010. The wide window of face detection. *J. Vis.* 10, 21. <http://dx.doi.org/10.1167/10.10.21>.
- Honey, C., Kirchner, H., VanRullen, R., 2008. Faces in the cloud: Fourier power spectrum biases ultrarapid face detection. *J. Vis.* 8 (12), 1–13. <http://dx.doi.org/10.1167/8.12.9> (9).
- Itier, R.J., Taylor, M.J., 2004. N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cereb. Cortex* 14, 132–142. <http://dx.doi.org/10.1093/cercor/bhg111>.
- Jacques, C., Withoft, N., Weiner, K.S., Foster, B.L., Rangarajan, V., Hermes, D., Miller, K.J., Parvizi, J., Grill-Spector, K., 2016. Corresponding ECoG and fMRI category-selective signals in human ventral temporal cortex. *Neuropsychologia* 83, 14–28. <http://dx.doi.org/10.1016/j.neuropsychologia.2015.07.024>.
- Jeffreys, D.A., 1989. A face-responsive potential recorded from the human scalp. *Exp. Brain Res.* 78, 193–202.
- Jeffreys, D.A., 1996. Evoked potential studies of face and object processing. *Vis. Cogn.* 3, 1–38.
- Jonas, J., Maillard, L., Frismand, S., Colnat-Coulbois, S., Vespignani, H., Rossion, B., Vignal, J.-P., 2014. Self-face hallucination evoked by electrical stimulation of the human brain. *Neurology* 83, 336–338. <http://dx.doi.org/10.1212/WNL.0000000000000628>.
- Jonas, J., Rossion, B., Brissart, H., Frismand, S., Jacques, C., Hossu, G., Colnat-Coulbois, S., Vespignani, H., Vignal, J.-P., Maillard, L., 2015. Beyond the core face-processing network: intracerebral stimulation of a face-selective area in the right anterior fusiform gyrus elicits transient prosopagnosia. *Cortex*. <http://dx.doi.org/10.1016/j.cortex.2015.05.026>.
- Joyce, C., Rossion, B., 2005. The face-sensitive N170 and VPP components manifest the same brain processes: the effect of reference electrode site. *Clin. Neurophysiol.* 116, 2613–2631.
- Keil, M.S., Cristobal, G., 2000. Separating the chaff from the wheat: possible origins of the oblique effect. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* 17, 697–710.
- Kincaid, J.M., 2005. An Event-Related Functional Magnetic Resonance Imaging Study of Voluntary and Stimulus-Driven Orienting of Attention. *J. Neurosci.* 25, 4593–4604. <http://dx.doi.org/10.1523/JNEUROSCI.0236-05.2005>.
- Koessler, L., Cecchin, T., Colnat-Coulbois, S., Vignal, J.-P., Jonas, J., Vespignani, H., Ramantani, G., Maillard, L.G., 2014. Catching the invisible: mesial temporal source contribution to simultaneous EEG and SEEG recordings. *Brain Topogr.* 28, 5–20. <http://dx.doi.org/10.1007/s10548-014-0417-z>.
- Kovacs, G., Zimmer, M., Banko, E., Harza, I., Antal, A., Vidnyanszky, Z., 2006. Electrophysiological correlates of visual adaptation to faces and body parts in humans. *Cereb. Cortex* 16, 742–753.
- Lewis, M.B., Edmonds, A.J., 2003. Face detection: mapping human performance. *Perception* 32, 903–920.
- Liu, J., Harris, A., Kanwisher, N., 2002. Stages of processing in face perception: an MEG study. *Nat. Neurosci.* 5, 910–916.
- Liu, H., Agam, Y., Madsen, J.R., Kreiman, G., 2009. Timing, Timing, Timing: Fast Decoding of Object Information from Intracranial Field Potentials in Human Visual Cortex. *Neuron* 62, 281–290. <http://dx.doi.org/10.1016/j.neuron.2009.02.025>.
- Liu-Shuang, J., Norcia, A.M., Rossion, B., 2014. An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia* 52, 57–72. <http://dx.doi.org/10.1016/j.neuropsychologia.2013.10.022>.
- Liu-Shuang, J., Torfs, K., Rossion, B., 2016. An objective electrophysiological marker of face individualisation impairment in acquired prosopagnosia with fast periodic visual stimulation. *Neuropsychologia* 83, 100–113. <http://dx.doi.org/10.1016/j.neuropsychologia.2015.08.023>.
- McCarthy, G., Wood, C.C., 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. *Electroencephalogr. Clin. Neurophysiol. Potentials* Sect. 62, 203–208. [http://dx.doi.org/10.1016/0168-5597\(85\)90015-2](http://dx.doi.org/10.1016/0168-5597(85)90015-2).
- McCarthy, G., Puce, A., Belger, A., Allison, T., 1999. Electrophysiological studies of human face perception. II: response properties of face-specific potentials generated in occipitotemporal cortex. *Cereb. Cortex* 9, 431–444.
- McKeeff, T.J., Remus, D.A., Tong, F., 2007. Temporal limitations in object processing across the human ventral visual pathway. *J. Neurophysiol.* 98, 382–393.
- Meeren, H.K.M., de Gelder, B., Ahlfors, S.P., Hämäläinen, M.S., Hadjikhani, N., 2013. Differential cortical dynamics in face and body perception: an MEG study. *PLoS One* 8, e71408. <http://dx.doi.org/10.1371/journal.pone.0071408>.
- Minnebusch, D.A., Suchan, B., Daum, I., 2009. Losing your head: behavioral and electrophysiological effects of body inversion. *J. Cogn. Neurosci.* 21, 865–874. <http://dx.doi.org/10.1162/jocn.2009.21074>.
- Nakamura, K., Kawashima, R., Sato, N., Nakamura, A., Sugiura, M., Kato, T., Hatano, K., Ito, K., Fukuda, H., Schormann, T., Zilles, K., 2000. Functional delineation of the human occipito-temporal areas related to face and scene processing. A PET study. *Brain* 123 (Pt 9), 1903–1912.
- Nasr, S., Echavarria, C.E., Tootell, R.B.H., 2014. Thinking outside the box: rectilinear shapes selectively activate scene-selective cortex. *J. Neurosci.* 34, 6721–6735. <http://dx.doi.org/10.1523/JNEUROSCI.4802-13.2014>.
- Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottareau, B.R., Rossion, B., 2015. The steady-state evoked potential in vision research: A review. *J. Vis.* 6 (6), 4. <http://dx.doi.org/10.1167/15.6.4>.
- Orlov, T., Makin, T.R., Zohary, E., 2010. Topographic representation of the human body in the occipitotemporal cortex. *Neuron* 68, 586–600. <http://dx.doi.org/10.1016/j.neuron.2010.09.032>.

- Peelen, M.V., Downing, P.E., 2005. Selectivity for the human body in the fusiform gyrus. *J. Neurophysiol.* 93, 603–608. <http://dx.doi.org/10.1152/jn.00513.2004>.
- Potter, M.C., 2012. Recognition and memory for briefly presented scenes. *Front. Psychol.* 3, 1–9. <http://dx.doi.org/10.3389/fpsyg.2012.00032>.
- Potter, M.C., Levy, E.L., 1969. Recognition memory for a rapid sequence of pictures. *J. Exp. Psychol.* 81, 10–15.
- Puce, A., Perrett, D., 2003. Electrophysiology and brain imaging of biological motion. *Philos. Trans. R. Soc. B Biol. Sci.* 358, 435–445. <http://dx.doi.org/10.1098/rstb.2002.1221>.
- Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.* 18, 2188–2199.
- Regan, D., 1989. Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine. Elsevier, New York.
- Rossion, B., Jacques, C., 2008. Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage* 39, 1959–1979. <http://dx.doi.org/10.1016/j.neuroimage.2007.10.011>.
- Rossion, B., Caharel, S., 2011. ERP evidence for the speed of face categorization in the human brain: disentangling the contribution of low-level visual cues from face perception. *Vis. Res.* 51, 1297–1311. <http://dx.doi.org/10.1016/j.visres.2011.04.003>.
- Rossion, B., Jacques, C., 2011. In: Kappenman, E.S., Luck, S.J. (Eds.), *The N170: understanding the time-course of face perception in the human brain. The Oxford Handbook of ERP Components*, New York, pp. 115–142.
- Rossion, B., Gauthier, I., Tarr, M.J., Despland, P., Bruyer, R., Linotte, S., Crommelinck, M., 2000. The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *Neuroreport* 11, 69–74.
- Rossion, B., Hanseeuw, B., Dricot, L., 2012a. Defining face perception areas in the human brain: a large-scale factorial fMRI face localizer analysis. *Brain Cogn.* 79, 138–157. <http://dx.doi.org/10.1016/j.bandc.2012.01.001>.
- Rossion, B., Prieto, E.A., Boremanse, A., Kuefner, D., Van Belle, G., 2012b. A steady-state visual evoked potential approach to individual face perception: effect of inversion, contrast-reversal and temporal dynamics. *NeuroImage* 63, 1585–1600. <http://dx.doi.org/10.1016/j.neuroimage.2012.08.033>.
- Rossion, B., 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 (1), 1–18. <http://dx.doi.org/10.1167/15.1.18> (18).
- Rousselet, G.A., Macé, M.J.-M., Fabre-Thorpe, M., 2003. Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J. Vis.* 3, 440–455. <http://dx.doi.org/10.1167/3.6.5>.
- Rousselet, G.A., Husk, J.S., Bennett, P.J., Sekuler, A.B., 2008. Time course and robustness of ERP object and face differences. *J. Vis.* 8 (12), 1–18. <http://dx.doi.org/10.1167/8.12.3> (3).
- Schneider, B.L., DeLong, J.E., Busey, T.A., 2007. Added noise affects the neural correlates of upright and inverted faces differently. *J. Vis.* 7 (4), 1–24. <http://dx.doi.org/10.1167/7.4.4> (4).
- Schwarzlose, R.F., Baker, C.I., Kanwisher, N., 2005. Separate face and body selectivity on the fusiform gyrus. *J. Neurosci.* 25, 11055–11059. <http://dx.doi.org/10.1523/JNEUROSCI.2621-05.2005>.
- Sergent, J., Ohta, S., Macdonald, B., 1992. Functional neuroanatomy of face and object processing — a positron emission tomography study. *Brain* 115, 15–36.
- Srinivasan, R., Russell, D.P., Edelman, G.M., Tono, G., 1999. Increased synchronization of neuromagnetic responses during conscious perception. *J. Neurosci.* 19, 5435–5448.
- Tanskanen, T., Nasanen, R., Montez, T., Paallysaho, J., Hari, R., 2005. Face recognition and cortical responses show similar sensitivity to noise spatial frequency. *Cereb. Cortex* 15, 526–534.
- Taylor, J.C., Roberts, M.V., Downing, P.E., Thierry, G., 2010. Functional characterisation of the extrastriate body area based on the N1 ERP component. *Brain Cogn.* 73, 153–159. <http://dx.doi.org/10.1016/j.bandc.2010.04.001>.
- Thierry, G., Pegna, A.J., Dodds, C., Roberts, M., Basan, S., Downing, P., 2006. An event-related potential component sensitive to images of the human body. *NeuroImage* 32, 871–879. <http://dx.doi.org/10.1016/j.neuroimage.2006.03.060>.
- Thompson, J.C., Clarke, M., Stewart, T., Puce, A., 2005. Configural processing of biological motion in human superior temporal sulcus. *J. Neurosci.* 25, 9059–9066. <http://dx.doi.org/10.1523/JNEUROSCI.2129-05.2005>.
- Tsao, D.Y., Moeller, S., Freiwald, W.A., 2008. Comparing face patch systems in macaques and humans. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19514–19519. <http://dx.doi.org/10.1073/pnas.0809662105>.
- Weiner, K.S., Grill-Spector, K., 2010. Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage* 52, 1559–1573. <http://dx.doi.org/10.1016/j.neuroimage.2010.04.262>.
- Weiner, K.S., Grill-Spector, K., 2011. Not one extrastriate body area: using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *NeuroImage* 56, 2183–2199. <http://dx.doi.org/10.1016/j.neuroimage.2011.03.041>.
- 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. <http://dx.doi.org/10.1016/j.neuroimage.2015.03.010>.