CORTEX XXX (XXXX) XXX

Available online at www.sciencedirect.com

ScienceDirect

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

Research Report

Does automatic human face categorization depend on head orientation?

Q9

Q8 *Charles C.-F. Or* [a,b,*], *Talia L. Retter* [b,c] *and Bruno Rossion* [b,d,e]

[a] *Division of Psychology, School of Social Sciences, Nanyang Technological University, Singapore*
[b] *Psychological Sciences Research Institute & Institute of Neuroscience, University of Louvain, Belgium*
[c] *Department of Behavioural and Cognitive Sciences, Institute of Cognitive Science & Assessment, University of Luxembourg Luxembourg*
[d] *Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France*
[e] *CHRU-Nancy, Service de Neurologie, F-54000 Nancy, France*

ARTICLE INFO

Article history:
Received 25 February 2020
Reviewed 01 July 2020
Revised 11 November 2020
Accepted 19 March 2021
Action editor Holger Wiese
Published online xxx

Keywords:
Face categorization
Head orientation
EEG
Frequency-tagging

ABSTRACT

Whether human categorization of visual stimuli as faces is optimal for full-front views, best revealing diagnostic features but lacking depth cues, remains largely unknown. To address this question, we presented 16 human observers with unsegmented natural images of different living and non-living objects at a fast rate ($F = 12$ Hz), with natural face images appearing at $F/9 = 1.33$ Hz. Faces posing all full-front or at ¾ side view angles appeared in separate sequences. Robust frequency-tagged 1.33 Hz (and harmonic) occipito-temporal electroencephalographic (EEG) responses reflecting face-selective neural activity did not differ in overall amplitude between full-front and ¾ side views. Despite this, alternating between full-front and ¾ side views within a sequence led to significant responses at specific harmonics of .67 Hz ($F/18$), objectively isolating view-dependent face-selective responses over occipito-temporal regions. Critically, a time-domain analysis showed that these view-dependent face-selective responses reflected only an earlier response to full-front than ¾ side views by 8–13 ms. Overall, these findings indicate that the face-selective neural representation is as robust for ¾ side faces as for full-front faces in the human brain, but full-front views provide a slightly earlier processing-time advantage as compared to rotated face views.

1. Introduction

Human categorization of faces among other objects in natural scenes is rapid, accurate, and automatic (Crouzet, Kirchner, & Thorpe, 2010; Retter, Jiang, Webster, & Rossion, 2020). The ability for generic face categorization (i.e., assigning variable visual stimuli to the specific category of “faces”) is crucial for social interaction, and constrains the information available for other forms of face processing, such as categorization of

---

* *Corresponding author*. Division of Psychology, School of Social Sciences, Nanyang Technological University, 48 Nanyang Avenue 639818, Singapore.
E-mail address: charlesor@ntu.edu.sg (C.C.-F. Or).

https://doi.org/10.1016/j.cortex.2021.03.030
0010-9452/© 2021 Elsevier Ltd. All rights reserved.

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

emotional expression and identity. However, the visual appearance of faces can vary greatly due to variations in natural presentation. One major factor arises from changes in the subject's head orientation in depth, defining the view angle of the face. Faces posed at the ¾ side views (i.e., mid-profile views, or intermediate views), for example, lose the visual symmetry of full-front faces, and parts of the faces' features may be occluded. Thus, robust generic face categorization is contingent on the ability to generalize across rotations of the face in depth.

Given the sophisticated nature of the problem, extremely little is known about the role of head orientation variations in human generic face categorization. This is surprising especially when the effect of head orientation on human face *identity* recognition has been more extensively studied (e.g., Baddeley & Woodhead, 1983; Benton, Jennings, & Chatting, 2006; Bruce, 1982; Bruce, Valentine, & Baddeley, 1987; Bruce & Young, 1986; Hill, Schyns, & Akamatsu, 1997; Jeffery, Rhodes, & Busey, 2007, 2006; Liu & Chaudhuri, 1998, 2002; Logie, Baddeley, & Woodhead, 1987; O'Toole, Edelman & Bü;lthoff, 1998; Valentin, Abdi, & Edelman, 1997; Van der Linde & Watson, 2010). The results pointing to independent coding of different head orientations have been extensively incorporated in artificial face recognition systems that can handle such transformations (e.g., Schneiderman & Kanade, 2000; Chen & Lien, 2009; for review, see; Ravidas, Ansari, & Kukreja, 2014). Perhaps the only relevant studies on head orientation's effect on human generic face categorization come from work by Bindemann and colleagues, who recorded response times for categorizing stimuli as faces at full-front, ¾ side, or full-profile views when they were embedded in various visual scenes or a white background (Bindemann & Lewis, 2013; Burton & Bindemann, 2009). The authors could not find any differences in response time across the three head orientations, except that categorizing full-profile faces took longer only when these faces were put in a rectangular picture frame positioned in the periphery of large visual scenes. These results suggested that head orientation could only influence generic face categorization in extreme conditions, where faces appeared in the visual periphery and posed at extreme angles (e.g., full-profile views).

However, a lack of effect of head orientation at intermediate angles (¾ side views) during face categorization appears to be inconsistent with head orientation's role in face identity recognition. Categorizing stimuli as faces is so fast that a highly sensitive design may be needed to reveal potential differences across head orientations. Most importantly, direct access to the visual categorization process, i.e., without an explicit task involving attention, decision and motor components, could be required to examine whether the (adult human) face categorization process is affected by different head orientations.

To answer these questions, we used a fast periodic visual stimulation (FPVS) approach coupled with scalp electroencephalogram (EEG), which provides an implicit (i.e., task-free), Q1 sensitive signature of automatic generic face categorization (Rossion, Torfs, Jacques, & Liu-Shuang, 2015). By successively presenting, at a fixed rate, faces embedded in a fast temporal sequence of non-face objects (Fig. 1), a frequency-tagged electrophysiological response at the periodic face inputs reflects the system's capacity to discriminate faces from other non-face objects and generalize this discrimination over a wide range of face stimuli despite visual differences due to lighting, identity, expression, etc. (e.g., Rossion et al., 2015; Retter & Rossion, 2016a; Or, Retter, & Rossion, 2019; Retter et al., 2020). Here, by manipulating the periodicity of presentations of distinct head orientations, we used this highly sensitive FPVS-EEG approach to define the contribution of head orientation in human face categorization. We compared full-front with ¾ side views, as ¾ side views disrupt symmetry and partly mask key features but introduce depth information not available in full-front faces, such as the angles of the forehead and the nose, while information about both sides of the face is still partly available (Baddeley & Woodhead, 1983; Van der Linde & Watson, 2010, Fig. 1). This comparison is also relevant to the controversial claim that ¾ side faces enhance face identity recognition performance compared to full-front faces, resulting in extensive face-recognition studies comparing the two head orientations (e.g., Baddeley & Woodhead, 1983; Bruce et al., 1987; Liu & Chaudhuri, 1998, 2002; Logie et al., 1987; O'Toole, Edelman, & Bülthoff, 1998; Valentin et al., 1997; Van der Linde & Watson, 2010). Resolving the differences between head orientations remains a key issue in the design of artificial face recognition systems (e.g., Farfade, Saberian, & Li, 2015; Yang, Luo, Loy, & Tang, 2018). Q2

Here, we designed an experiment to probe head orientation response differences in two ways. First, we investigated whether there were amplitude differences in face-categorization responses to full-front *vs.* ¾ side views. Only full-front faces (Condition 1) or only ¾ side faces (Condition 2) were embedded at a periodic frequency (1 face every 9 stimuli, 12 Hz/9 = 1.33 Hz) in a fast (12 Hz) periodic sequence of natural images of objects (Fig. 1), such that any periodic response selective to generic face categorization was quantified at 1.33 Hz and harmonics (2.67 Hz, etc.) in each type of the sequences. If full-front faces hold an overall advantage in terms of amplitude of generic face categorization response, the face-selective neural response at 1.33 Hz would be specifically larger in this condition (Condition 1) than when ¾ side faces were presented (Condition 2).

Second, we investigated whether there were differences in the amplitude and/or timing of face-categorization responses to full-front vs. ¾ side views. We designed an original condition (Condition 3) in which the two head orientations were alternately presented throughout the periodic face presentations (i.e., ¾, full-front, ¾, full-front, and so on) such that the view-selective stimulation frequency (1 view every 18 stimuli, 12 Hz/18 = .67 Hz) effectively halved the face stimulation frequency (1.33 Hz). We reasoned that if there were different (i.e., selective) or at least partially different categorization processes for ¾ side and full-front views, then we would observe an EEG response exactly at this halved frequency of .67 Hz, reflecting dissociable neural responses specific to full-front and ¾ side views. By alternating views within a sequence, the response amplitude at .67 Hz and its specific harmonics is sensitive to phase as well as amplitude asymmetries (e.g., see Retter & Rossion, 2016b; Norcia et al., 2015). On the contrary, if detection of full-front and ¾ side Q3 views relied upon common processes tuned to a generic face category, then the system would be blind to differences in

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

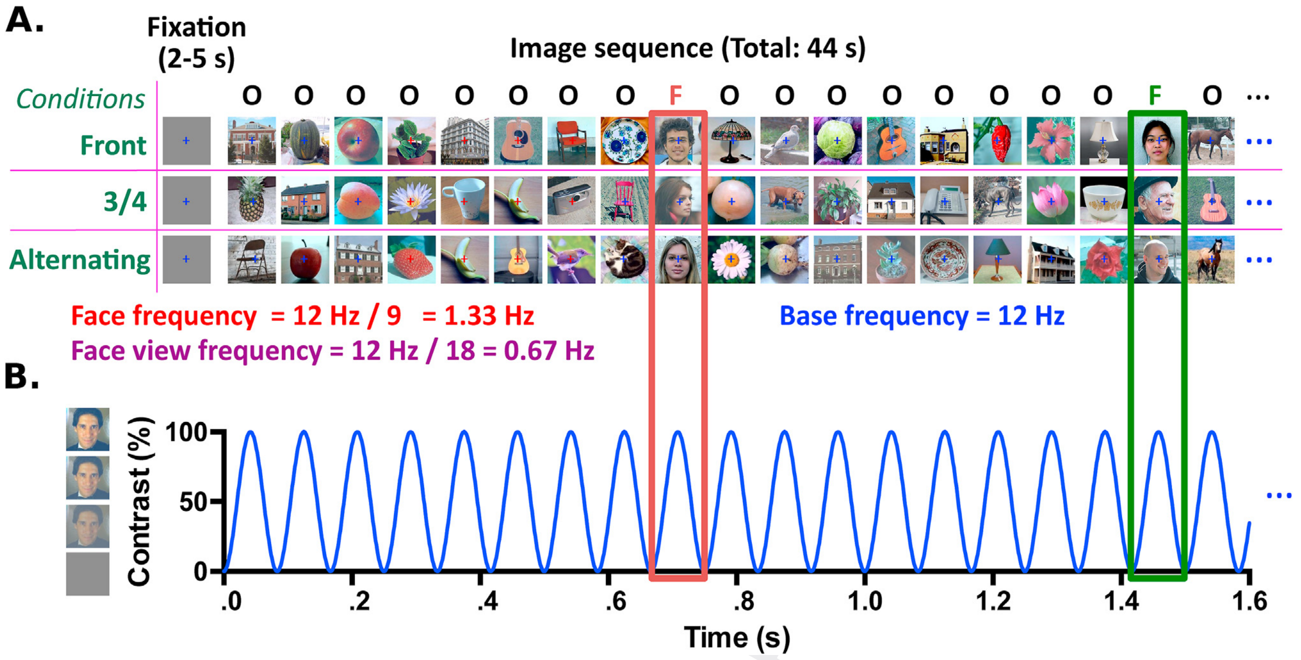

Fig. 1 – Experimental procedure. Following central fixation in each condition, the image sequence of 528 images (only 19 illustrated here) was presented at base frequency, $f_B = 12$ Hz (12 images/s), thus each image was presented for 83.3 ms (1000 ms/12) with sinusoidal contrast modulation from 0% to 100% contrast (blue curve in the contrast-time function; left of function shows example stimuli at 0%, 36%, 65%, and 100% contrasts, bottom to top). Faces (F) were presented periodically following every eight non-face objects (O), thus at a face-selective frequency, $f_F = 12$ Hz/9 = 1.33 Hz. In the Front condition, only full-front faces were presented; in the ¾ condition, only ¾ side faces (consistently facing one side within a sequence; here, right-facing is shown). In the Alternating condition, head orientation alternated between full-front (red box) and ¾ side (green box) views, thus each view presented at $f_V = 1.33$ Hz/2 = .67 Hz. In actual experiments, each sequence included fade-in and fade-out periods (2 sec each) not illustrated here. The task was to press a key when the fixation cross changed colour (blue to red for 300 ms at random times, which did not coincide with the onsets and offsets of images).

head orientation. In this latter case, a response only to the face-selective frequency of 1.33 Hz, but not to the view-selective (and also face-selective) frequency of .67 Hz, would be observed.

Throughout, detailed investigations into the scalp topography probe for spatial differences in the responses to head orientation are reported. Additionally, a complementary time-domain analysis of the face categorization responses under the different viewing conditions will provide more specific information about the source of any observed differences. In sum, this FPVS-EEG study using full-front and ¾ side views investigated whether, to what extent, and along which time course, head orientation affects face categorization in the human brain.

2. Materials and methods

2.1. Participants

The experiment was completed by 16 participants (10 females, mean age = 24.2 ± 4.0, range: 20–36 years) who all had normal, or corrected-to-normal, visual acuity. They were all right-handed according to an adapted Edinburgh Handedness Inventory measurement (Oldfield, 1971). None reported any history of psychiatric or neurological disorders. They were naïve to the purpose of study, and were not aware that faces were presented at a fixed rate of 1 out of 9 stimuli. The sample size was based on previous studies with this paradigm in adult humans, reporting highly significant effects with 11 (Jacques, Retter, & Rossion, 2016) to 20 (Or et al., 2019) participants (M = 14.5). While 16 participants may be considered as modest sample size (although within the range for human EEG studies), the FPVS approach is characterized by a very high signal-to-noise ratio (SNR) (Regan, 1989; Norcia et al., 2015). This high SNR is partly due to the large number of discrimination trials (i.e., face-selective neural responses) recorded in a short amount of time for every individual participant (here, 232 face trials across 4 stimulation sequences for each condition; see Procedure). All participants provided written informed consent prior to the experiment and received honoraria for their participation, as approved by the Biomedical Ethical Committee of the University of Louvain and the 2013 WMA Declaration of Helsinki.

2.2. Stimulus display

The stimuli were generated by a Dell XPS Desktop computer installed with Psychtoolbox 3.0.8 in MATLAB R2009a for Windows (MathWorks, Natick, MA, USA) using previously

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

validated scripts (e.g., Rossion & Boremanse, 2011), passed to a GeForce GTX 560 Ti graphics card, and displayed on a linearly gamma-corrected BenQ XL2420T monitor at a refresh rate of 120 Hz, with a screen resolution of 1920 × 1080 pixels and a colour depth of 24 bits/pixel placed at a viewing distance of 80 cm (pixel size: .0194°) in a dimly lit and sound-attenuated room. The mean luminance after gamma correction was 75.0 cd/m^2.

2.3. Stimuli

Colour photographs of 30 faces posed at full-front views, 30 faces posed at ¾ side views, and 247 non-face objects (animals, plants, man-made objects, houses, etc.) were obtained from ImageNet (http://image-net.org) and elsewhere on the Internet. All objects and faces, varying in size and lighting condition, were embedded in their original, natural scenes without segmentation from a wide range of backgrounds. All face images were taken from different persons. The mirror image of each face posed at ¾ side view was generated in order to create 30 pairs of left-facing and right-facing face images (see Procedure). The stimuli were each cropped to locate the face or non-face object at the centre of a square image and resized to subtend 5.05 ° × 5.05 ° of visual angle. The average luminance of each stimulus was equalized to the screen's mean luminance (75.0 cd/m^2). Note that variations in local colour, luminance and contrast associated with the appearance of the faces and objects in the images remained after this stimulus normalization.

2.4. Procedure

A schematic illustration of the experimental design is shown in Fig. 1. The stimulation sequence was presented through sinusoidal contrast modulation of successive images at a rapid rate of 12.0 Hz (*base frequency*, f_B), the fastest rate at which the largest face-selective response is observed in this paradigm (Retter et al., 2020). Each 83.3-ms (1000 ms/12.0, 10 frames/stimulus) stimulation cycle started with a uniform grey background from which an image appeared as its contrast increased in a sinusoidal fashion from 0% to 100% (full contrast) at 41.7 ms and then decreased at the same rate. The periodic sequence was comprised of eight objects (O) followed by a face (F), all randomly selected from their corresponding categories. Faces were thus presented at a frequency of 1/9 $f_B = 12.0$ Hz/9 = 1.33 Hz such that this *face-selective frequency* (f_F) and its harmonics ($2f_F$ = 2.67 Hz, $3f_F$ = 4.00 Hz, and so on) represented an index of both (1) discrimination between faces and objects, and (2) generalization across different faces (Rossion et al., 2015). Stimuli were repeated randomly (but not consecutively, max. 3 times) within a stimulus sequence.

The experiment was composed of three conditions, differing by the head orientation of the faces presented (Fig. 1). In Condition 1, the faces were all posed at full-front views throughout the stimulation sequence (*Front condition*). In Condition 2, the faces were all posed at ¾ side views (*¾ condition*). In Condition 3, the faces alternated between full-front and ¾ side views within a stimulation sequence (*Alternating condition*). Thus, each view was effectively presented at a frequency of ½ f_F = 1.33 Hz/2 = .67 Hz, such that this *face view-selective frequency* (f_V) and its odd-number harmonics (those not overlapping with the 1.33 Hz harmonics, i.e., $3f_V$ = 2.00 Hz, $5f_V$ = 3.33 Hz, $7f_V$ = 4.67 Hz, and so on) uniquely represented an index of (1) discrimination between objects and specific face views, and (2) discrimination between full-front and ¾ side views. (Note that there was no systematic change at .67 Hz for Conditions 1 and 2.) Each participant performed 12 stimulation sequences (four sequences per condition), each of which contained images in an independently randomized order. For the ¾ and Alternating conditions, the ¾ side views were all left facing in two sequences per condition, and right facing in the other two. The order of presentation of the 12 stimulation sequences was randomized for each participant.

Each stimulation sequence started with a fixation cross (in blue, .31 ° × 0.31 °) centred on a uniform grey background for 2–5 sec (duration randomly determined across sequences) in order to facilitate stable fixation of the participant. Subsequently, the stimulation sequence (528 successive images including 58 faces and 470 non-face objects) was centrally presented for 44 sec, including a 2-sec fade-in period at the beginning of the sequence and a 2-sec fade-out period at the end (with uninterrupted central display of the fixation cross superimposed on the images). As in previous studies (e.g., Or et al., 2019; Retter & Rossion, 2016a; Rossion et al., 2015), the contrast modulation depth of the periodic stimulation gradually increased from 0% to 100% during the fade-in period, and after 40 sec, reduced in the opposite direction from 100% to 0% during the fade-out period (keeping the sinusoidal contrast modulation). The introduction of these fading periods was intended to minimize blinks and abrupt eye movements due to an otherwise sudden appearance or disappearance of the flickering stimuli. Responses during the fading periods were not used in the data analyses (see EEG analyses).

During EEG recording, the participant was instructed to maintain central fixation throughout the image presentations while continuously monitoring the flickering stimuli. Their task was to detect brief colour changes of the fixation cross (blue to red for 300 ms), as in typical previous studies (e.g., Rossion et al., 2015). Such colour changes occurred 8 times randomly throughout each sequence, and were not correlated with the onsets and offsets of images. This task was orthogonal to the manipulation of interest in the study. The accuracy (percentage that the observer correctly pressed the key within 1500 ms after the onset of the colour change) and response times for accurate key presses were analysed to ensure that participants maintained a constant level of attention throughout the experiment.

2.5. EEG acquisition

The electroencephalogram (EEG) was acquired using a 128-channel Biosemi Active 2 system (BioSemi, Amsterdam, The Netherlands), with electrodes including standard 10–20 system locations as well as additional intermediate positions (http://www.biosemi.com/headcap.htm, relabelled to more conventional labels of the 10–5 system, see Supplementary Fig. S1 in Rossion et al., 2015). The EEG was sampled at 512 Hz. Offset of each electrode was reduced to a small ±20 mV

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

range by injecting saline gel on the scalp through the electrode with a plastic syringe. Eye movements were monitored by four additional electrodes placed at the outer canthi of the two eyes, and above and below the right orbit. During the experiment, triggers were sent via parallel port from the stimulation computer to the EEG recording computer at the onset and offset of each stimulation sequence, and at the minima (0% contrast) of all 12.0-Hz stimulation cycles (Fig. 1). Recordings were manually initiated by the experimenter when participants showed artefact-free EEG signals.

2.6. *EEG analyses*

2.6.1. *Preprocessing*

All EEG data were analysed using Letswave 5 (http://nocions.webnode.com/letswave) running on MATLAB R2014b. The signals were first detrended by subtracting the best-fit line (using the least-squares method) from the data, and then passed to a fourth-order low-pass Butterworth filter (Butterworth, 1930) with a cutoff frequency of 120 Hz. The data were then passed to a fast Fourier transform (FFT) multi-notch filter (width = .5 Hz) to remove electrical noise at 50 Hz (frequency of the alternating current) and its second harmonic (100 Hz). Subsequently, the filtered signals were segmented into 48-sec segments, keeping 2 sec each before and after a sequence (i.e., −2 sec through 46 sec). The DC component in each data segment was separately identified and then removed. Noisy and artefact-ridden channels (less than 5% of 128 channels) containing deflections larger than 100 μV in multiple presentation sequences were rebuilt using linear interpolations from immediately-adjacent noise-free channels. Finally, all channels (except the oculars) were re-referenced to a common average.

2.6.2. *Frequency-domain analysis*

The preprocessed data segment of each sequence was cropped again to keep only signals from exactly 2 sec after stimulus onset (the end of the fade-in period) to 41.0 sec after stimulus onset. The end time (41.0 sec) was chosen such that it was the longest possible time point, before the start of stimulus fade-out (at 42 sec), for capturing an integer number of .67-Hz cycles (i.e., .67 Hz × 39.0 sec = 26 cycles, which contains $N = 19{,}971$ time bins). The integer number of cycles was used to avoid spectral leakage of the frequencies of interest, i.e., .67 Hz (face view-selective frequency, f_V) and its harmonics (Note that all harmonics of 1.33 Hz (face-selective frequency, f_F) and 12.0 Hz (base frequency, f_B) are subsets of harmonics of .67 Hz). The sequences were then averaged separately for each condition and for each observer. An FFT was applied to the sequence-averaged data segments, and a positive x-axis-valued amplitude spectrum (normalized by $N/2$, in μV) was extracted in the frequency domain (ranging from 0 to 256 Hz) for each channel. Each spectrum had a high frequency resolution (i.e., distance between two adjacent frequency components) of .0256 Hz, which is the inverse of the segment duration (39.0 sec). This ensured unambiguous identification of the frequencies of interest (.67 Hz and harmonics).

To consider the variations of noise across the amplitude spectrum, a baseline subtraction was applied to each frequency component by subtracting the average amplitude of 20 surrounding frequency components (10 on each side, excluding the immediately adjacent bins and the local minimum and maximum bins; see e.g., Dzhelyova & Rossion, 2014; Mouraux et al., 2011) from the amplitude of the frequency component of interest. In addition, the SNR was also calculated by considering the same 20 surrounding frequency components (e.g., Or et al., 2019; Rossion, Alonso Pireto, Boremanse, Kuefner, & Van Belle, 2012). For group analysis, individual baseline-subtracted amplitude (and SNR) spectra were averaged across observers for each condition, resulting in the grand-averaged spectra.

2.6.2.1. Selecting the range of significant harmonic responses (z-score analyses). To analyse the responses at the base frequency, face-selective frequency, and face view-selective frequency (and their harmonics), we first determined a continuous range of significant harmonic responses for each frequency to include in the analysis. Individual amplitude spectra were first averaged across observers, and then across the 128 EEG channels for each condition. A z-score was calculated for each frequency component of this averaged spectrum by using the mean amplitude and SD of 20 surrounding frequency components (10 on each side, excluding the immediately adjacent bins; see Rossion et al., 2012) from the amplitude of the frequency component of interest. The harmonics to be included in the analysis would range from the fundamental frequency through a cut-off frequency determined by the last significant harmonic that yielded a z-score larger than 2.33 among the three conditions (i.e., beyond the 99.0 percentile of the signal-to-noise distribution; Retter & Rossion, 2016a). The significant harmonic responses (see Results) were summed, separately for each frequency type, in order to compare their overall amplitudes (and scalp topographies) across conditions (Retter & Rossion, 2016a).

2.6.2.2. Statistical comparisons across conditions. In order to analyse the whole-brain responses, individual harmonic-summed, baseline-subtracted amplitudes were further averaged over all 128 EEG channels, and the resulting data (16 observers × 3 conditions) were subjected to a one-way repeated-measures ANOVA. We also defined regions-of-interest (ROIs) over clusters of channels that scored the largest amplitudes, and which were concordant with previous studies (Jacques et al., 2016; Or et al., 2019; Quek & Rossion, 2017; Retter & Rossion, 2016a; Rossion et al., 2015). The baseline-subtracted amplitudes were then averaged over the channels within each ROI and subjected to repeated-measures ANOVA for comparisons across conditions and ROIs.

2.6.3. *Time-domain analysis*

The periodic responses were additionally examined in the time domain (e.g., Jacques et al., 2016; Or et al., 2019; Retter & Rossion, 2016a; Rossion et al., 2015). The preprocessed data segments were each passed to a fourth-order bandpass

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

Butterworth filter with a bandwidth of .1–30 Hz. The filtered data segment was further cropped to keep only signals from stimulus onset (0 sec) to 41.9 sec after. The end time (41.9 sec) was chosen such that it was the nearest time point to the start of stimulus fade-out (at 42 sec) for capturing an integer number of 12.0-Hz cycles (i.e., 12.0 Hz × 41.9 sec = 503 cycles, which contains $N = 21,465$ time bins). An FFT multi-notch filter (width = .5 Hz) was subsequently applied to the cropped signals to selectively remove 12.0 Hz and its first three harmonics, corresponding to the contribution of the base stimulation to the time-domain waveforms. The filtered signals were then cropped into smaller epochs of 1417 ms (17 × 83.3-ms base stimulation cycles), each including responses to a sequence of eight object stimuli, one face stimulus, and another eight object stimuli (OOOOOOOOFOOOOOOOO). Thus, each epoch contains responses for exactly one face stimulus. The cropping began at 2.25 sec after stimulus onset, which was the earliest time point possible after the 2-sec fade-in period. It should be noted that the first eight object stimuli of each epoch correspond to the last eight object stimuli of its immediately preceding epoch. After averaging all epochs per observer for each condition, the data were baseline-corrected by subtracting the mean response amplitude across 167 ms (corresponding to presentations of two object stimuli) preceding presentation of the face stimulus in the epoch sequence. For each condition, the baseline-corrected responses for all 16 participants were subjected to a two-tailed *t*-test at each time point. We defined a face-selective component by a time window that scored significant ($p < .01$) non-zero responses for 12 or more consecutive time points (i.e., ≥21.5 ms; see, e.g., Jacques et al., 2016; Laganaro, 2014). Similar statistical procedures were performed on the within-subjects differences of individual baseline-corrected responses between any two conditions in order to compare the waveforms across conditions. The baseline-corrected responses were subsequently averaged across observers per condition, so as the three sets of within-subjects differences.

Importantly, for the *Alternating condition*, the even-number epochs (before averaging) contain responses only to faces at the full-front views, and the odd-number epochs represent ¾ side view responses. Thus, the epoch-averaging procedures described in the last paragraph were also applied separately to the odd- and even-number epochs in order to compare the time-domain responses towards the full-front and ¾ side views for the *Alternating condition*. This odd/even separation was also applied to the other two conditions (*Front* and *¾* conditions) such that the time-domain responses were compared across the three conditions based on the same numbers of epochs. The face-selective components for the odd and even epochs were separately defined using the aforementioned statistical procedures.

3. Results

3.1. Behavioural data

Observers' percentages for accurate key presses over 32 colour changes (8 changes × 4 sequences per condition) were close to ceiling in all conditions (*Front*: 95.7% ± 7.2%; *¾*: 96.6% ± 4.3%; *Alternating*: 96.0% ± 6.1%; mean ± 1 *SD*). These percentages were not significantly different across conditions, $F(1, 15) = .30$, $p = .59$. The corresponding response times were rapid in all conditions (all shorter than 500 ms: *Front*: 464.2 ± 64.7 ms; *¾*: 458.0 ± 47.0 ms; *Alternating*: 451.4 ± 39.8 ms; mean ± 1 *SD*). No significant differences were found across conditions for response times, $F(1, 15) = .64$, $p = .44$. These results supported that observers maintained a constant level of attention across sequences and conditions.

3.2. EEG data

3.2.1. Frequency-domain analysis

Here, we report responses represented by three frequency rates, i.e., face-selective responses (1.33 Hz and harmonics), face view-selective responses (.67 Hz and odd harmonics), and base-rate responses common to all stimuli (12.0 Hz and harmonics). Q4

3.2.1.1. FACE-SELECTIVE RESPONSES (1.33 Hz)

3.2.1.1.1. AVERAGED ACROSS ALL CHANNELS. For each condition, we first averaged the SNRs across all 128 channels and 16 observers, resulting in the mean frequency spectra shown in Fig. 2a. Robust responses were observed at the face-selective frequency (1.33 Hz) and its harmonics for all three conditions (blue lines), representing the brain's discrimination of faces from other objects (i.e., face-selective responses as in Rossion et al., 2015; Retter & Rossion, 2016a; etc.). As the highest significant harmonic (i.e., z-score > 2.33; see Methods) across the three conditions was 20.0 Hz (14th harmonic), the baseline-subtracted amplitudes were then summed, separately for each observer and condition, across these significant harmonics (i.e., over the range of 1.33 Hz, 2.67 Hz, 4.00 Hz, and so on until 20.0 Hz, excluding 12.0 Hz which coincides with the base frequency; for individual participant spectra, see Supplementary Figure S1) for subsequent analyses. A one-way repeated measures ANOVA on these sums of harmonic amplitudes, averaged over all 128 channels, showed no significant differences across conditions, $F(1, 15) = .46$, $p = .51$. Thus, at an overall level, variation in head orientation did not modulate the face-selective response (Fig. 3a: Chanavg).

3.2.1.1.2. OCCIPITO-TEMPORAL REGIONS. Next, we examined the spatial distributions of such face-selective responses across the scalp, and whether these distributions varied across the three conditions. Fig. 2b shows back-of-the-head scalp topographies illustrating the harmonic-sums of baseline-subtracted amplitudes averaged over observers. The face-selective responses (middle column) were the largest in the occipito-temporal regions. In particular, spatially adjacent channels PO10, P10, and PO8 in the right hemisphere scored the three largest responses (in descending order) consistently across the three conditions, in line with previous studies (Jacques et al., 2016; Or et al., 2019; Quek & Rossion, 2017; Retter & Rossion, 2016a; Rossion et al., 2015); thus they were defined as the right occipito-temporal (rOT) ROI (Fig. 3c). For each condition, individual harmonic-summed baseline-subtracted amplitudes were averaged across these three channels for response quantification and statistical analysis. To assess lateralization, individual amplitudes were also averaged

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

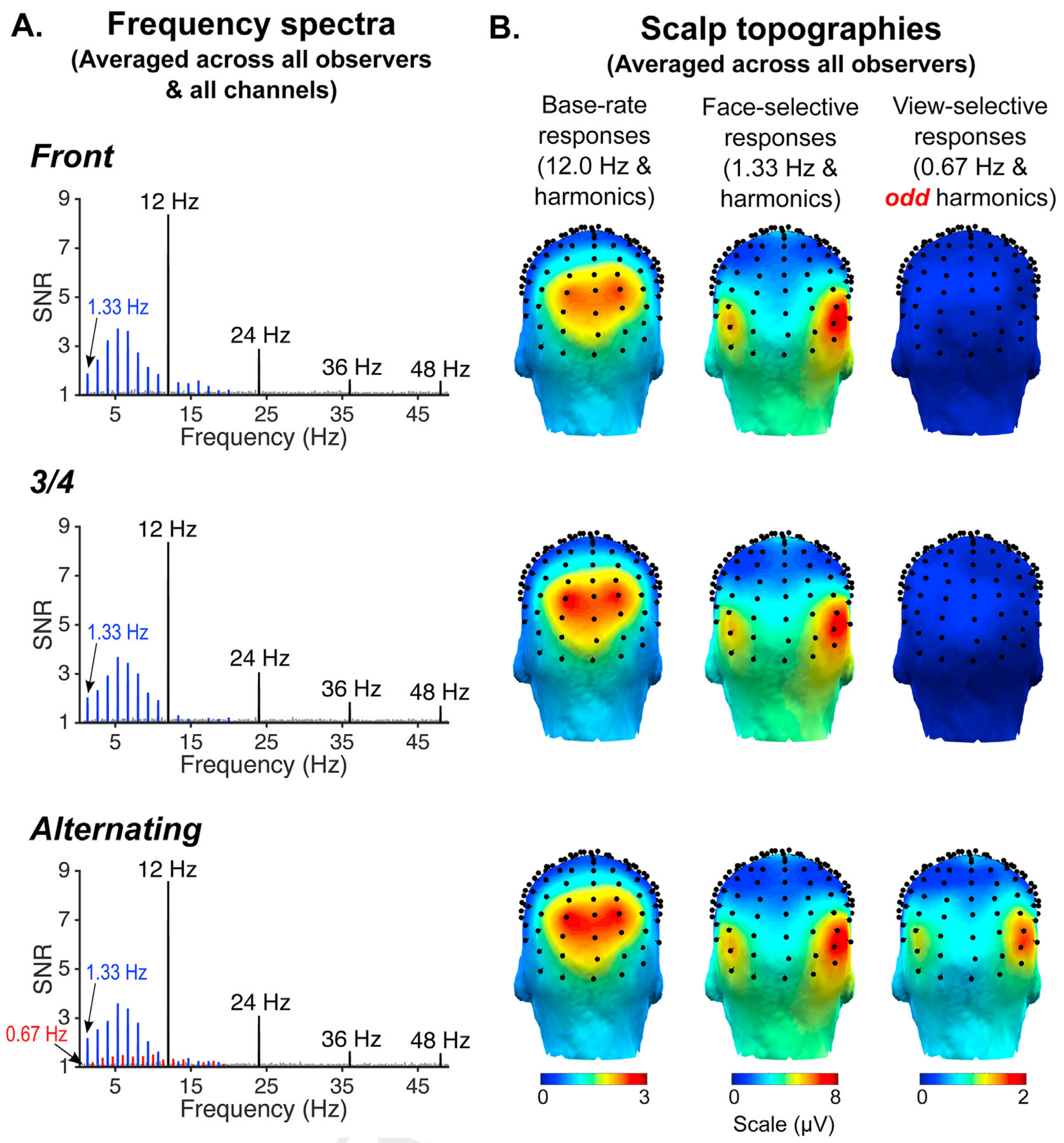

Fig. 2 – Frequency-domain responses (grand-averaged). (a) For each of the three conditions, the frequency spectrum plots the signal-to-noise ratio (SNR) averaged over all observers and all 128 channels as a function of frequency. The black lines denote the base-rate responses (12 Hz and harmonics). Smaller face-selective responses (1.33 Hz and harmonics, in blue) and face view-selective responses (.67 Hz and odd harmonics only for the *Alternating condition*, in red) are also highlighted. The presence of significant view-selective responses showed that face-selective responses differed at least partially between full-front and ¾ faces. (b) Each scalp topography (back of the head) shows the sums of baseline-subtracted amplitudes across significant harmonics of a stimulation frequency (base, face, or view) averaged across all observers. The face-selective and view-selective responses peaked over occipito-temporal regions. Note the different scales across response types.

across contralateral channels over the left occipito-temporal ROI (lOT, encompassing PO9, P9, and PO7; Fig. 3c).

The two-way repeated measures ANOVA (independent variables: ROI, condition) on these individual baseline-subtracted amplitudes (Fig. 3a: lOT & rOT) showed no significant differences across the three conditions, $F(1, 15) = .57$, $p = .46$, between the ROIs, $F(1, 15) = 2.25$, $p = .15$, and their interaction, $F(1, 15) = .001$, $p = .98$.

3.2.1.1.3. Single-channel analysis. In order to further examine potential effects of head orientation on the responses, we conducted a single-channel analysis using a percentile bootstrap approach (sampling subjects with replacement; see Jacques et al., 2016) to statistically compare the two most contrasted conditions (i.e., *Front* condition: all full-front faces *vs.* ¾ condition: all ¾ faces; see Supplementary materials for the less relevant comparisons with the *Alternating* condition) channel-by-channel (10,000 bootstrap samples per condition; significance level = .01). A few left prefrontal and lateral temporal channels showed significantly larger responses to full-front faces than to ¾ side faces

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

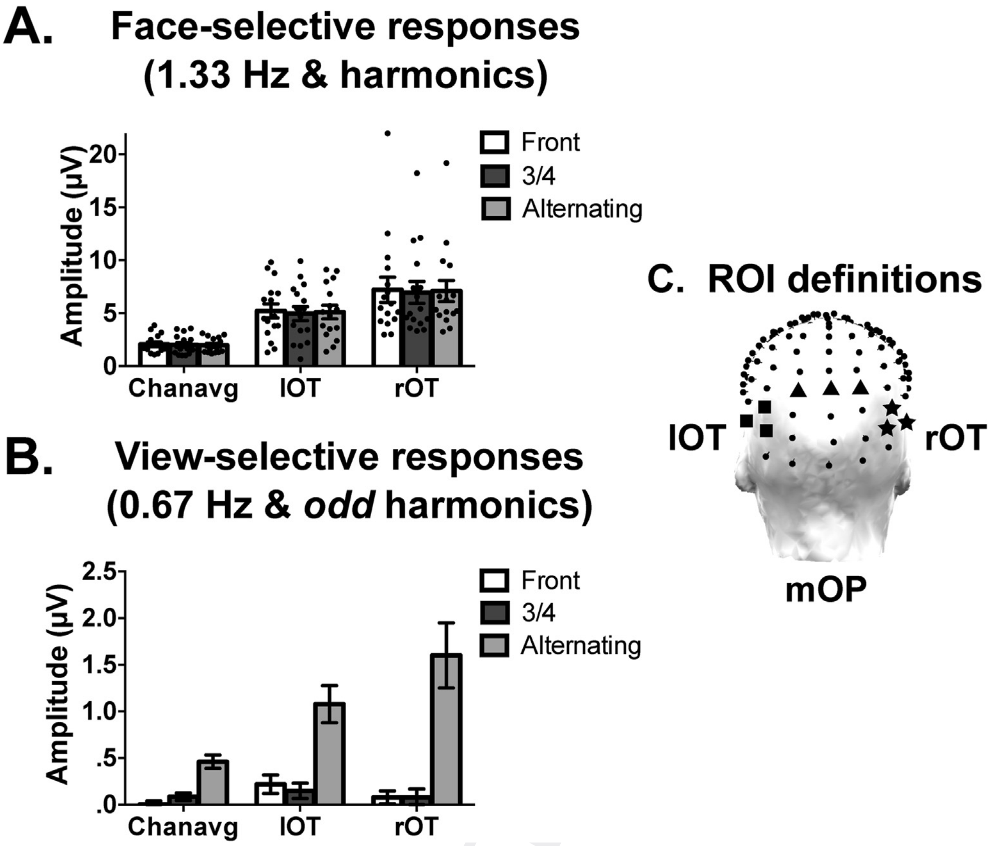

Fig. 3 – Frequency-domain responses in terms of the grand-averaged sums of baseline-subtracted amplitudes over significant (a) face-selective harmonics (1.33 Hz, 2.67 Hz, etc., except 12 Hz), or (b) view-selective (also face-selective) harmonics (.67 Hz, 2.00 Hz, 3.33 Hz, etc.). Each response was averaged over all 128 channels (Chanavg), left occipito-temporal (lOT), or right occipito-temporal (rOT) channels separately. Each bar represents the mean over all 16 observers (error bar = ± 1 *SEM*). For (a), the dots represent responses for individual observers (see also Supplementary Figure S1). The face-selective responses were similar across all view conditions, but significant view-selective responses were observed only in the *Alternating* condition. (c) Corresponding channel locations that define the lOT (■), rOT (★), and medial occipito-parietal (mOP, ▲) ROIs.

(Fpz: .37 μV, $p = .0046$; Fp1: .41 μV, $p = .0061$; AF7: .49 μV, $p = .0001$; F7: .43 μV, $p = .0019$; FT7: .48 μV, $p = .001$; C5: .32 μV, $p < .0001$; T7h: .39 μV, $p = .0011$; other channels: $ps > .0135$; Front − ¾ for all differences). It should be noted, however, that these channels were associated with the smallest face-selective responses (.80–1.39 μV across the two conditions; rank 80th or below among 128 channels in both conditions). Also, none of these channels were over the occipito-temporal ROIs. Together, these results suggest that the effect of head orientation on the magnitude of face-selective responses is negligible (Fig. 3a).

3.2.1.1.4. Inter-participant correlations and intra-participant comparisons between views. We also assessed potential differences of head orientation at the individual participant level. First, we found that individual amplitudes were highly correlated across the three conditions (rOT: $rs = .95–.99$, lOT: $rs = .89–.97$; $ps < .0001$ for all pairwise Pearson's correlation coefficients; see Fig. 4 for individual scalp topographies). Hence, the lack of significant difference at the group level between face-selective responses for full-front and ¾ faces cannot be attributed to any meaningful positive (e.g., Front > ¾) *vs.* negative (Front < ¾) differences in separate subsets of individuals potentially cancelling out each other in the overall analysis.

Then, for each condition and channel separately, the individual baseline-subtracted amplitude spectrum was cropped into segments each consisting of a 1.33 Hz harmonic and 20 surrounding bins (10 on each side; for indicating the amount of variation within individual data). These 21-bin segments were summed across 1.33 Hz harmonics (1.33 Hz–20.0 Hz, except 12 Hz). Supplementary Figure S1 shows such amplitude segments averaged over the rOT region, showing the largest overall responses in the *Front* and *¾* conditions. For each participant and channel separately, the harmonic-summed amplitude segments between *Front* condition and *¾* condition were compared using paired-samples *t*-tests (significance level = .01, two-tailed). Among all 16 participants and 128 channels, none of the comparisons showed any significant differences (Even for significance level = .05, only three channels within one single participant showed significant

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

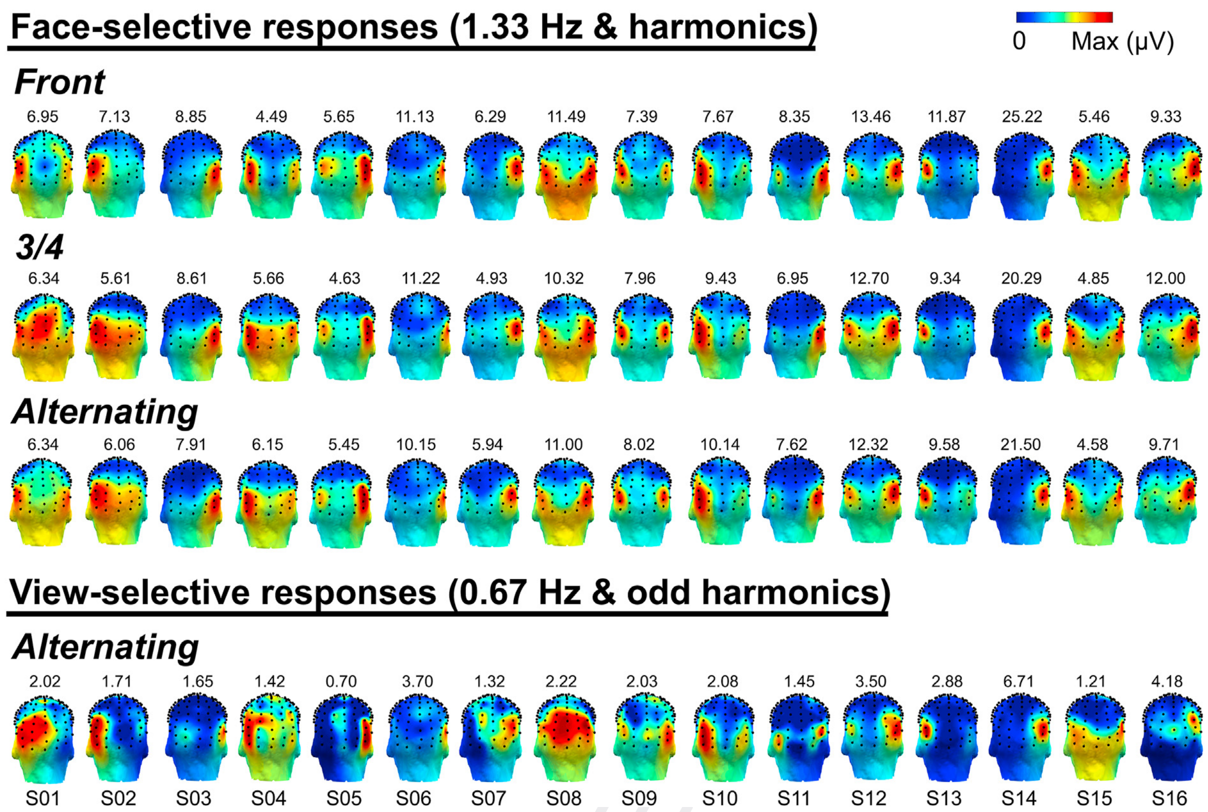

Fig. 4 – Individual frequency-domain scalp topographies for all 16 observers (S01–S16). A back-of-the-head topography shows the sums of baseline-subtracted amplitudes across significant face-selective harmonics (1.33–20.0 Hz, except 12 Hz) or view-selective harmonics (.67–19.33 Hz) for each observer. The colour scale (minimum amplitude always 0, maximum amplitude above each topography) is adapted for each topography in order to highlight the response variations over the scalp across participants.

differences, *ps* = .04–.047; other *ps* = .079–1.00). These results further demonstrated an absence of head orientation's effect on the magnitude of face-selective responses in the frequency domain.

3.2.1.2. Face view-selective responses (.67 Hz). The lack of significant amplitude difference and the high correlation between frequency-tagged face-selective 1.33 Hz (and harmonic) occipito-temporal EEG responses for full-front and ¾ side views point to equally strong neural representations of these two views in the human brain, and no overall amplitude advantage for full-front faces in generic face categorization, supported by analyses at both the group level and the individual level. However, there might still be differences in timing between these two views in terms of face-selective responses. This is where the alternation between full-front and ¾ side views within a sequence (Condition 3) is critical: if significant responses are found at specific harmonics of .67 Hz (*F*/18), they objectively isolate view-dependent face-selective amplitude and/or timing response differences over occipito-temporal regions.

3.2.1.2.1. Averaged across all channels. In line with this hypothesis, we found significant responses (z-score > 2.33) for the *Alternating* condition over a range of view-selective harmonics (3.33 Hz, 4.67 Hz, and so on until 18.00 Hz, except .67 Hz and 15.33 Hz) in the channel-averaged frequency spectrum (red lines in Fig. 2a). The amplitudes of view-selective responses were much smaller than face-selective responses (blue lines) common to both face views. No significant responses were found (z-score < 2.33) at any view-selective .67 Hz harmonics in the *Front* and the ¾ conditions, confirming that the view-selective responses were unique to alternations of the two head orientations.

To summarize the data, we summed view-dependent face-selective responses over all view-selective harmonics up to 20.0 Hz (odd-number harmonics only: .67 Hz, 2.00 Hz, 3.33 Hz, 4.67 Hz, and so on until 19.33 Hz; even-number harmonics were excluded as they coincide with the harmonics of common face-selective responses) such that this frequency range was made consistent with the one defined for the common face-selective responses (even though three harmonics: .67 Hz, 15.33 Hz, and 19.33 Hz did not show significant view-selective responses). A one-way repeated measures ANOVA on these sums of amplitudes, averaged over all 128 channels, showed a significant main effect of *Condition*, $F(1, 15) = 28.6$, $p < .001$. Post hoc pairwise comparisons showed that responses for the *Alternating* condition were significantly larger than both *Front* and ¾ conditions ($p < .001$ in both comparisons), but no significant difference between *Front* and ¾ conditions ($p = .07$). These results again confirm that the view-

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

selective harmonics reflect periodic responses to alternating face views (Fig. 3b).

3.2.1.2.2. Occipito-temporal regions. Fig. 2b (rightmost column) shows the scalp topographies of observer-averaged, view-selective harmonic-summed, baseline-subtracted amplitudes for the face view-selective responses. Such responses peaked at PO10, P10, and PO8 (in terms of baseline-subtracted amplitudes in descending order), identical to what we observed from the face-selective responses (but here, only in the *Alternating* condition). Thus, we defined the occipito-temporal ROIs by the same bilateral channels (lOT: PO9, P9, PO7, rOT: PO10, P10, PO8; Fig. 3c) as in the face-selective response analysis. Then, we compared individual view-selective responses (view-selective harmonic-summed baseline-subtracted amplitudes) averaged separately for the two ROIs and the three conditions by use of a two-way repeated measures ANOVA.

The main effect of *Condition* was significant, $F(1, 15) = 37.2$, $p < .001$, since responses were present only in the *Alternating* condition. The main effect of *ROI* was not significant, $F(1, 15) = .36$, $p = .56$, nor was there a significant interaction effect, $F(1, 15) = 2.39$, $p = .14$. Post hoc pairwise comparisons indicated that the mean view-selective responses for the *Alternating* condition were significantly larger than the other two conditions ($p < .001$ for both *Alternating vs. Front* and *Alternating vs. ¾* comparisons; no significant differences for *Front vs. ¾*, $p = .70$), again confirming that the face view-selective responses were unique to alternating face views within a sequence (Fig. 3b). Despite a right hemispheric advantage of .52 μV on the grand-averaged view-selective responses in the Alternating condition (Fig. 3b), no significant differences were found between the lOT and rOT data, $t(15) = 1.24$, $p = .24$ (two-tailed *t*-test) (see Fig. 4 for individual scalp topographies). These results, and in fact even just the presence of significant face view-selective responses, indicate that face-selective responses differ at least partially between full-front and ¾ faces.

Since the approach used here allows straightforward quantifications of the EEG responses of interest, we computed the proportion of the view-selective response to the face-selective response for each ROI: 22.5% ± 2.9% for all 128-channel averages, 20.4% ± 3.0% over lOT, 20.5% ± 1.6% over rOT (mean ± SE). Thus, the amplitude of the view-selective responses was only about 20–23% of that of the face-selective responses.

3.2.1.3. Base-rate responses (12.0 Hz). Responses to base-rate stimulations merely represent the brain's sensitivity to images, regardless of whether they are faces or non-face objects. Analysis on the channel-averaged spectra (Fig. 2) revealed that the first four harmonics (12.0 Hz, 24.0 Hz, 36.0 Hz, and 48.0 Hz; black lines) were significant (i.e., z-score > 2.33) in all conditions. Thus, we summed the baseline-subtracted amplitudes across these four harmonics separately for each observer and condition. A one-way repeated measures ANOVA on these individual sums of amplitudes, averaged over all channels, showed no significant differences across conditions, $F(1, 15) = .21$, $p = .65$. Fig. 2 shows the scalp topographies of the harmonic-sum averaged across observers for each condition. Channels POO5, POO6, and POOz over the medial occipito-parietal region (mOP; Fig. 3c) scored the largest responses (in terms of harmonic-summed baseline-subtracted amplitudes) consistently for each condition. A one-way repeated measures ANOVA on individual sums of baseline-subtracted amplitudes, averaged over the mOP channels, again revealed no significant differences across the three conditions, $F(1, 15) = .84$, $p = .37$. These results indicate that the base-rate responses did not differ across view conditions.

3.2.2. *Time-domain analysis*

The significant view-dependent face-selective response at .67 Hz and its specific harmonics in the *Alternating* condition demonstrate clearly that face-selective populations of neurons respond (at least partially) differently to full-front and ¾ side views. Since there were no spatial and amplitude differences between full-front and ¾ side views at 1.33 Hz and its harmonics, it is likely that the response at .67 Hz and harmonics is driven by differences in timing sensitive to the two views. In order to directly test this assumption, we conducted a time-domain analysis, in which the (spatio-)temporal patterns of activity to full-front and ¾ side views would be revealed.

Fig. 5 shows the time-domain responses, in terms of baseline-corrected amplitudes averaged across all epochs and observers (see Methods), separately for all 128 channels and the three conditions. The left column plots the original, unfiltered waveforms, where the consistent periodic fluctuations corresponded to responses to periodic base-rate stimulations at 12 Hz. The right column plots the differential waveforms after selectively filtering out such base-rate responses (see Methods; further details in Retter & Rossion, 2016a). These differential waveforms, time-locked to the periodic face stimuli (face stimulus onset at 0 sec in Fig. 5), indicated face-selective responses for all view conditions (as in Rossion et al., 2015; Jacques et al., 2016; Retter & Rossion, 2016a). Qualitatively, at least three distinctive components underlying the face-selective responses are observed over time: P1-face, N1-face, and P2/P3-face (Fig. 5). These components' timings are broadly consistent with previous results (e.g., Or et al., 2019; Retter & Rossion, 2016a; Rossion et al., 2015).

Fig. 6 plots the scalp topographies for the filtered waveforms evolving over time (in steps of 9.77 ms; see also the video in Supplementary Materials). Significant temporal delays were observed for components in the ¾ condition as compared to those in the *Front* condition (see scalp topographies for *Front* − ¾ conditions: These scalp topographies plot only significant differences, $p < .01$, estimated using the single-channel percentile bootstrap approach, Jacques et al., 2016), with timings of the *Alternating* condition in between. In addition to temporal differences, the early P1-face component (Fig. 6a) had a significant ($p < .01$) scalp topography difference between the peak responses in the *Front* and the ¾ conditions. In particular, responses peaked over the right occipito-temporal area for the *Front* condition but over the medial occipital region for the ¾ condition, with the *Alternating* condition in-between. As time progressed, however, the later N1-face and P2/P3-face components, with peaks of much higher amplitudes (Fig. 6b and c, respectively), became more

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

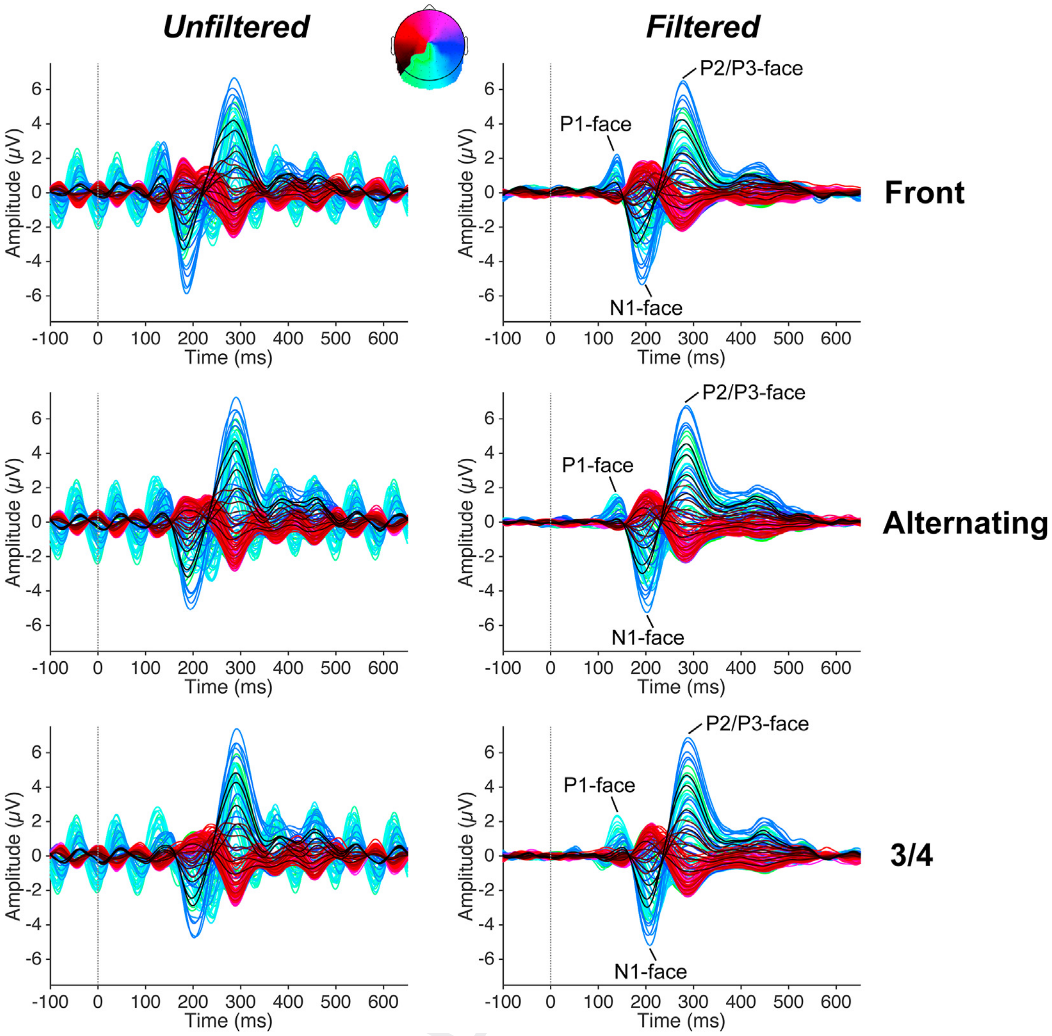

Fig. 5 – Time-domain waveforms (baseline-corrected amplitudes) for all 128 channels for the three conditions. The left column shows unfiltered data, i.e., including periodic responses at the base frequency (12 Hz). The right column plots corresponding data notch-filtered to remove 12 Hz and harmonics (see Methods), showing the differential, face-selective responses. Time 0 ms indicates the onset of sinusoidal face stimulation. The 2D head map (viewed from top of the head) represents the colour codes of the channels.

consistent in scalp topography across conditions; both showed the largest responses over the occipito-temporal regions.

Supplementary video related to this article can be found at https://doi.org/10.1016/j.cortex.2021.03.030

In order to quantify the different temporal latencies of the P1-face components across conditions, we defined a right-and-medial occipito-temporal ROI (rmOT; Fig. 7) encompassing 17 channels that consistently ranked among the top responses in the three conditions (within 100–170 ms) but in a different order (e.g., PO10, P10, PO12, PO8 as top four channels for the *Front* condition, but Oiz, Oz, Iz, and POOz accordingly for the ¾ condition). Fig. 7a (left plot) shows the (base-rate filtered) waveforms averaged over rmOT channels for the three conditions. The three face-selective components (coloured horizontal lines near the bottom of the plot) were defined by significant non-zero responses ($p < .01$) over 12 consecutive time points (i.e., 21.5 ms) on the average waveforms. The significant response times for the P1-face were 109–148 ms (*Front*), 113–154 ms (*Alternating*), and 122–163 ms (¾) after stimulus onset. Thus, there was a 13-ms delay ($p < .05$) for all ¾ side face presentations compared to all full-front views. The N1-face components were significant from 167 to 228 ms (*Front*), 173–236 ms (*Alternating*), and 179–240 ms (¾) accordingly, showing a 12-ms delay for ¾ views ($p < .01$). For P2/P3-face, significant responses in the *Front* condition were present in three intervals: 247–351 ms, 374–484 ms, and 511–536 ms. For the *Alternating* and the ¾ conditions, the significant time ranges were continuous: 251–552 ms and 255–552 ms, respectively. The temporal delay of the onset of P2/P3-face between the ¾ side and full-front views was 8 ms ($p < .01$). Minor, but significant, responses were recorded from 618 to 644 ms only in the ¾ condition. To evaluate the statistical significance of the

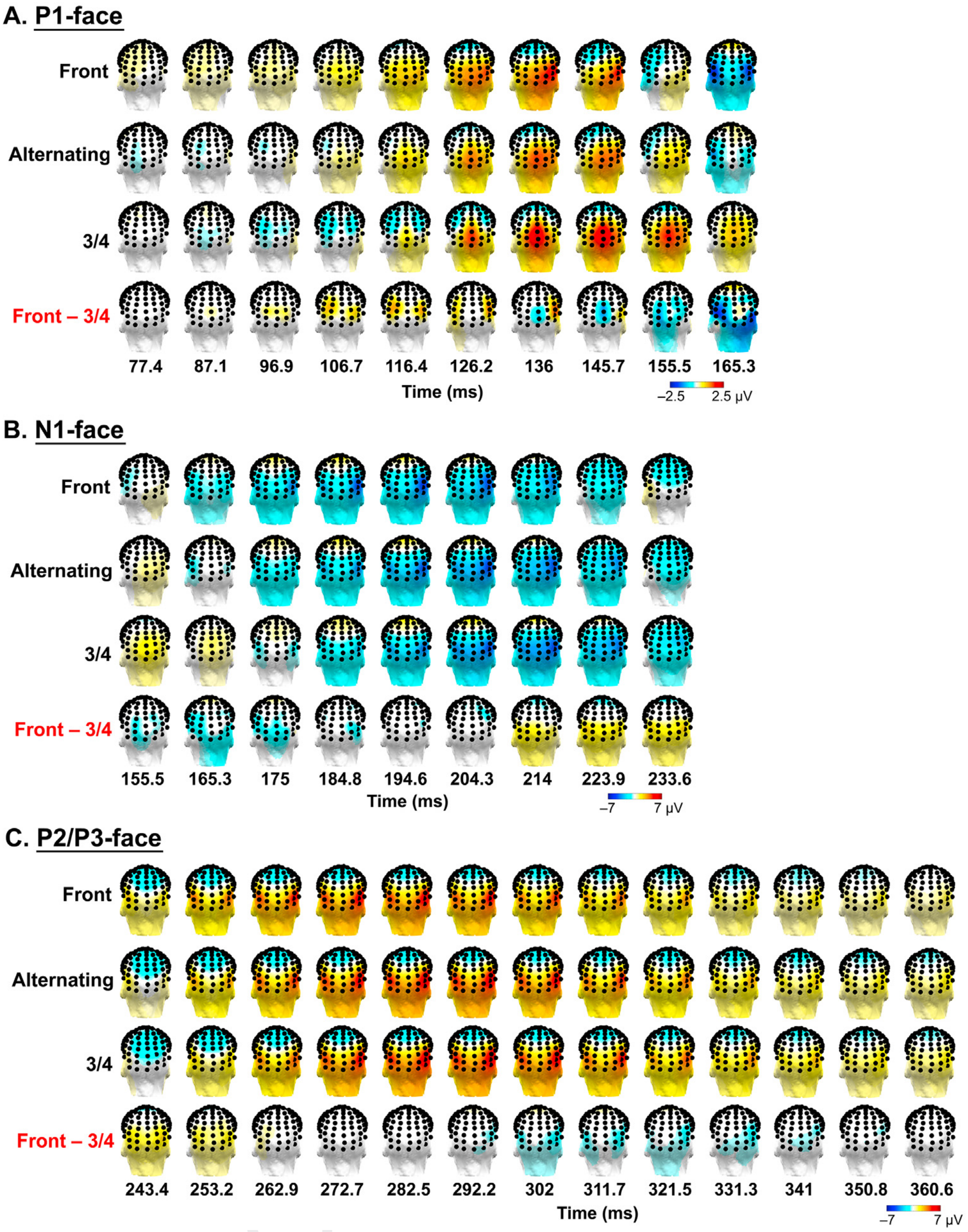

Fig. 6 – Time-domain scalp topographies (back of the head) for the filtered waveforms (Fig. 5, right column) around the three components, (a) early positive P1-face, (b) negative N1-face, and (c) late positive P2/P3-face, for the three conditions (Front, Alternating, 3/4) over time in steps of 9.77 ms. The differences between *Front* and *¾* conditions (Front – 3/4) are drawn in the bottom rows (only significant differences, $p < .01$, are shown; i.e., white areas represent no significant differences). Note that the scales are different for (a) compared to (b) and (c), as P1-face has smaller amplitudes. See also the video in Supplementary Materials.

temporal delays, we subtracted individual rmOT-averaged response waveforms between *Front* and *¾* conditions (significant differences over 21.5 ms in the bottommost lines in Fig. 7a; black: $ps < .01$, grey: $.01 \leq ps < .05$, two-tailed paired-samples t tests). The time windows of the significant differences ($p < .05$) were 95–120 ms, 152–191 ms, 210–261 ms, 290–365 ms, 433–493 ms, and 616–650 ms, which were consistent with the temporal delays of the face-selective components. Waveforms for most individual observers (Supplementary Figure S2) also showed the three components with the temporal shift across views. Aside from temporal delays, there were no significant differences of peak amplitudes between *Front* and *¾* conditions in any of the three components over the rmOT ($ps = .10–.89$; two-tailed paired-samples t tests).

We specifically analysed the time-domain responses for channel PO10 (Fig. 7, right column) over the right occipito-temporal area, as this channel scored the largest responses in both N1-face and P2/P3-face, as well as in the frequency domain (both face-selective and view-selective responses). Similar to analysis over rmOT (which PO10 is part of), three

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

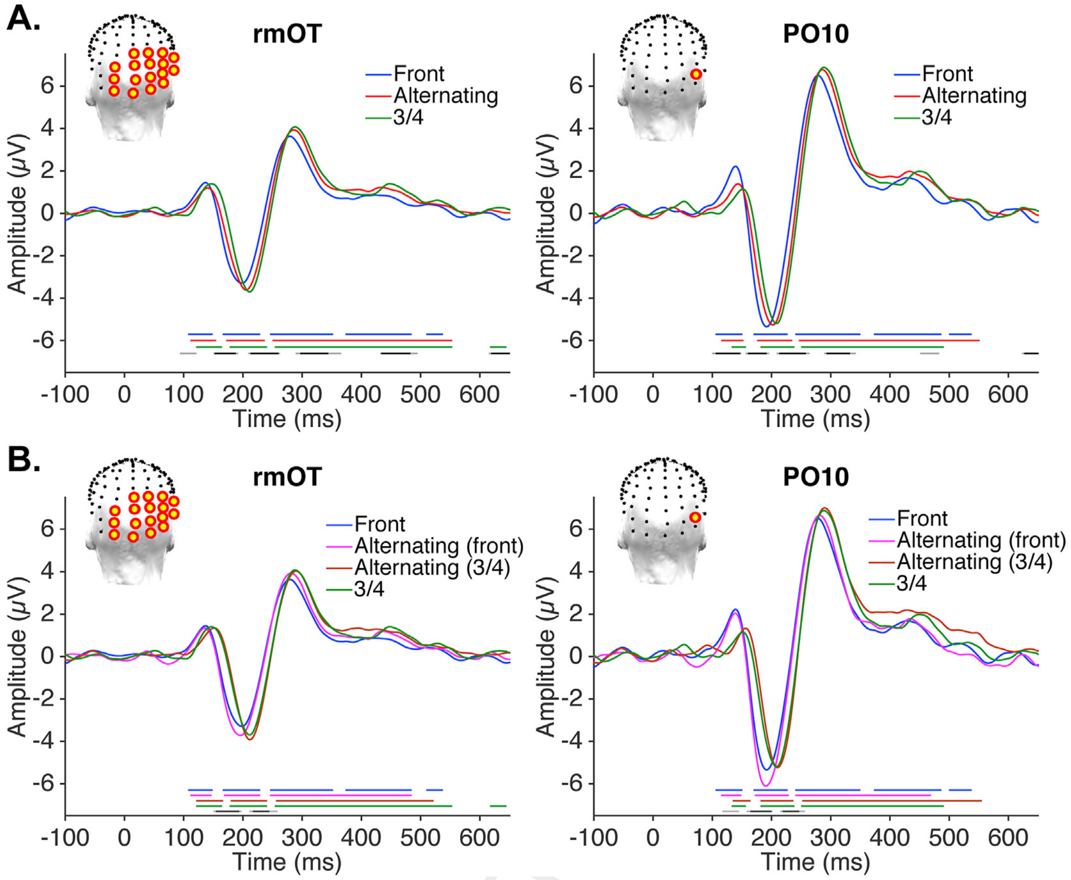

Fig. 7 – Time-domain responses (averaged over observers) over the rmOT (left column) and channel PO10 (right column) for (a) the three conditions, and (b) after splitting responses in the *Alternating* condition into those only for full-front views (magenta) and only for ¾ side views (brown). Horizontal lines below the waveforms (with matching colours) represent significant responses for 12 consecutive time points (i.e., *ps* < .01 over a range of 21.5 ms, two-tailed one-sample *t*-test against zero). The bottommost horizontal lines represent significant (*ps* < .01 (black), .01 ≤ *ps* < .05 (grey) over a range of 21.5 ms) pairwise differences (a) between *Front* and ¾ conditions (blue *vs.* green), or (b) between full-front-only and ¾-only waveforms in the *Alternating* condition (magenta *vs.* brown).

significant (*ps* < .01 over 21.5 ms) components underlying the face-selective responses were defined, and verified among most individual waveforms (Supplementary Figure S3). The P1-face components were present from 107 to 150 ms (Front), 116–152 ms (Alternating), and 134–156 ms (¾) after stimulus onset. The temporal delay for ¾ views was 27 ms (*p* < .01), longer than 13 ms over rmOT. The longer delay over PO10 was partly due to a spatiotemporal shift of the P1-face peak from the right occipito-temporal region for the *Front* condition to the medial occipital area in the ¾ condition. Indeed, over PO10, the P1-face peak amplitude was significantly higher for the *Front* condition than for the ¾ condition (Mean difference = 1.06 μV, *t*(15) = 3.83, *p* = .0016, two-tailed paired-samples *t* test). The N1-face components were present from 171 to 226 ms (Front), 177–234 ms (Alternating), and 183–238 ms (¾) accordingly, showing a 12-ms delay (*p* < .01) for ¾ side views that was identical to the delay over rmOT. For P2/P3-face, significant responses in the *Front* condition occurred over three intervals: 241–349 ms, 374–486 ms, and 501–536 ms. For the *Alternating* and the ¾ conditions, the significant time ranges were continuous: 247–550 ms and 251–490 ms, respectively. The temporal delay of the onset of P2/P3-face was 10 ms (*p* < .01). For these two later components, no significant differences of peak amplitudes were found between *Front* and ¾ conditions (*ps* > .19).

In summary, it appears that the time-domain waveform averaged across all epochs for the *Alternating* condition (Fig. 7a) lies between those for *Front* and ¾ conditions, in terms of the spatiotemporal dynamics of the components. We wondered whether this is a result of averaging epochs representing full-front and ¾ side face presentations. Thus, for the *Alternating* condition, we averaged epochs for the two face views separately (Fig. 7b: magenta *vs.* brown lines), permitting a within-condition comparison that was not possible in the frequency domain (since full-front and ¾ side views were presented at the same, .67-Hz frequency). After splitting the data in the *Alternating* condition (see Methods), significant pairwise differences (*ps* < .05 over 21.5 ms; see bottommost lines in Fig. 7b) between the full-front-only and the ¾-only waveforms over both rmOT and PO10 demonstrated once again the significant temporal

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

shift across face views using *only* the *Alternating* condition (except that, with half of the trials compared, the differences in P1-face onset over rmOT did not reach significance, $ps = .07–.23$). Pairwise comparisons of the areas under the two waveforms over rmOT did not show any significant difference, $t(15) = .24$, $p = .81$ (two-tailed), supporting that the differences between the two waveforms were only in the form of a temporal shift (similarly over PO10: $t(15) = .81$, $p = .43$). In addition, the full-front-only waveform in the *Alternating* condition (magenta) coincides ($ps > .05$ over 21.5 ms) with the waveform in the *Front* condition (blue), and the ¾-only waveform in the *Alternating* condition (brown) overlaps ($ps > .05$ over 21.5 ms) with the waveform in the *¾* condition (green) over both rmOT and PO10, demonstrating the consistency of responses to the two views regardless of condition. We also verified that splitting the epochs in the same manners (odd/even) for the *Front* and *¾* conditions (see Methods) led to the same temporal-shift conclusion. Taken together, the temporal-domain results suggested that full-front views led to earlier face-selective responses than ¾ side views.

4. Discussion

The current study investigated whether head orientation (full-front *vs.* ¾ side face views) modulated human generic face categorization in rapidly presented natural images. We pursued this goal by examining the responses to fast periodic visual stimulations (FPVS) of either or both of the two head orientations against presentations of non-face object images, while recording EEG responses from the human brain.

The aggregate face-selective response amplitudes in the frequency domain (1.33 Hz and harmonics), which peaked over the bilateral occipito-temporal regions, did not differ between the two face views. Although a null effect should always be interpreted with caution, the FPVS paradigm used here is very sensitive and there was no hint of a difference between *Front* and *¾* conditions (with consistent head orientations throughout the stimulation sequences), with nearly overlapping distributions of EEG response amplitudes across the 16 individuals tested (Figs. 2a and 4). Individual EEG spectra, computed over a large number of trials per participant (from 232 face presentations per condition), also highlight a lack of overall quantitative difference in response amplitude between the two conditions (Supplementary Figure S1). This result indicates that face-selective neural representation—forming the basis for generic face categorization—is as robust for ¾ side faces as for full-front faces in the human brain. This observation is interesting because ¾ side views disrupt symmetry and partly mask internal features (e.g., part of an eye, part of the mouth) that are thought to play a key role in human generic face categorization (e.g., Omer, Sapir, Hatuka, & Yovel, 2019; Paras & Webster, 2013). However, ¾ side views also introduce depth information not available in full-front faces, such as the angles of the forehead and the nose, while information about both sides of the face is still partly available (Baddeley & Woodhead, 1983; Van der Linde & Watson, 2010). The effect of these different cues may thus counteract each other, so that the overall face-selective neural response, supporting generic face categorization, does not differ significantly between full-front and ¾ side faces.

Despite this lack of significant difference when comparing the face-selective response amplitudes in the first two conditions of the present experiment (and a near perfect correlation of amplitude values across individuals), significant responses at .67 Hz and harmonics associated specifically to alternating presentations of full-front and ¾ side views were found in the third, *Alternating*, condition. This result indicates that the face-selective system *is* tuned to head orientation, i.e., providing objective and quantitative evidence that head orientation influences automatic human face categorization. Again, in this condition, responses at .67 Hz and specific harmonics can be potentially accounted for by differences in amplitude and/or timing between the alternating face views (Norcia et al., 2015; Retter & Rossion, 2016b); since no amplitude differences were found previously, we reasoned that this effect was likely produced by timing differences in the responses to the two views.

We therefore examined the differential waveforms selective to face stimulations in the time domain, which enabled separate, directly comparable spatiotemporal analyses on responses selective to each view angle. The results revealed a systematic temporal advantage for full-front faces, for which the average face-selective components occurred 8–13 ms earlier than those for ¾ side views. In addition, we found an early spatial dissociation between responses to the two face views, where peak responses to full-front views were more right lateralized, while those to ¾ side views were more medial, over occipital regions. The absence of peak amplitude differences over the rmOT between views at any deflection was in agreement with the absence of amplitude differences between 1.33 Hz responses in *Front* and *¾* conditions. Hence, encoding of face views occurs as early as the earliest steps of generic face categorization via spatiotemporal dissociation of the head orientation signals.

The earlier onset of response to full-front faces highlights one of the advantages of using (a spatiotemporal analysis of) FPVS-EEG data for understanding how head orientation influences face categorization. While Bindemann and colleagues (Bindemann & Lewis, 2013; Burton & Bindemann, 2009) found no response-time difference in head orientation (full-front *vs.* ¾ side views) during face categorization, this could be due to specific issues with experimental design. For example, face stimuli in those studies were always placed within a picture frame, being quite different from normal appearances of faces in a natural visual environment. More generally, response-time measures encompass the contributions of many generic attentional, decisional and motor-related processes. When the difference in processing time is significant but small, an explicit behavioural task associated with a large variability across individuals may not be sensitive enough to capture it. In contrast, the FPVS approach isolates automatic face-selective neural responses to the rapid presentation of successive images, with the additional advantage of allowing for a detailed analysis of the dynamic cortical variations of the response.

In regards to view-selective scalp topography, in the frequency-domain analysis, face view-selective response amplitudes peaked over the same occipito-temporal cortical

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

regions as face-selective response amplitudes (i.e., common response amplitudes to both face views). Also, face-selective response amplitudes to full-front and ¾ side views showed overall the same scalp topographies, apparently offering no evidence for a spatial dissociation at the level of the scalp for face categorization responses to different head orientations. In the time-domain analysis, however, early spatial dissociation between face views became evident, despite that such scalp topography differences were relinquished later. The discrepancy between frequency-domain and time-domain results can be understood in that the frequency-domain face-selective response was a generalized response across the response waveform in the time domain, while the relatively smaller amplitude of the early, spatially dissociated P1-face component was overshadowed by later, larger components that did not exhibit a spatial dissociation.

The current results showed that sensitivity to head orientation, at least for a comparison between full-front and ¾ side views, was mainly exhibited in the form of a temporal difference in responses to different view angles over high-level, face-selective occipito-temporal regions. Several fMRI studies have examined view selectivity in face-selective regions, either through repetition suppression/adaptation paradigms (Fang, Murray, & He, 2007; Grill-Spector et al., 1999; Xu & Biederman, 2010) or multivariate pattern analysis (Axelrod & Yovel, 2012; Ramírez, Cichy, Allefeld, & Haynes, 2014). In this context, Axelrod and Yovel (2012) suggested that high-level face-selective regions including the right occipital face area (OFA), fusiform face areas (FFA) and the right superior temporal sulcus (STS) were able to decode a wide range of face views and favoured full-front over side views, suggesting mirror-symmetric face coding as the full-front view nearly achieves mirror symmetry. Interestingly, face views were also decoded in low-level, non-face-selective early visual cortex (though not favouring the mirror-symmetric full-front view), perhaps a result of enhanced low-level differences between head orientations due to the use of highly homogenous sets of segmented stimuli (but see Ramírez, et al., 2014 over effective control of low-level retinotopic effects). It is important to note that these studies involved classification of head orientations alone (i.e., without direct comparisons against non-face objects), thus their results cannot be directly implicated on the current findings resulting from high-level categorization of faces from other objects in natural scenes (i.e., face-selective responses).

This issue of distinct representations to full-front or ¾ side faces has also been addressed with neurophysiological recordings performed in the infero-temporal (IT) cortex of macaque monkey, where neurons responding selectively to faces have been reported in many studies (since Gross, Rocha-Miranda, & Bender, 1972; e.g., Perrett, Rolls, & Caan, 1982; Desimone, Albright, Gross, & Bruce, 1984; Young & Yamane, 1992; Tsao, Freiwald, Tootell, & Livingstone, 2006; Freiwald & Tsao, 2010). Early studies reported that most face-selective neurons in the monkey IT were highly sensitive to head orientation, i.e., with different populations of neurons coding for full-front or ¾ side views of faces (Desimone et al., 1984; Perrett et al., 1984, 1985, 1991, 1998). These monkey studies suggested that faces were first encoded selectively (i.e., categorized) as view-dependent representations, which could then pool their outputs to smaller populations of neurons higher in the hierarchy to generate view-invariant representations (Perrett & Oram, 1998; see also Booth & Rolls, 1998). More recent studies measuring face-selective responses in fMRI-defined regions of the monkey IT (since Tsao et al., 2006) have been taken as supporting this view, even suggesting a progressive evolution from view-specific face-selective neurons in posterior regions to mirror-symmetric representations in intermediate regions and almost full view invariance in the most anterior IT region (Freiwald & Tsao, 2010). These monkey data form the basis of neurofunctional face recognition models (e.g., Duchaine & Yovel, 2015; Haxby, Hoffman, & Gobbini, 2000; Tovée & Cohen-Tovée, 1993), suggesting a view-sensitive module feeding to a view-independent module in face processing.

However, to our knowledge, none of these studies in the monkey brain have shown that face-selective single neuron responses to ¾ side faces were delayed relative to full-front faces. Rather, it has been suggested that delays observed at the population level could be due to more cells tuned to the view(s) most frequently experienced, so that activity amongst the population of cells selective for the object's appearance would accumulate more slowly when the object is seen in an unusual view, orientation, or size (Perrett, Oram, & Ashbridge, 1998). However, in this case, one has to make an assumption that ¾ side faces are less frequently experienced, which is unknown and unlikely for humans, as it is well documented that ¾ side faces are recognized for their identity by humans better than, or at least as accurately as, full-front faces (Bruce et al., 1987; Hill et al., 1997; Troje & Bülthoff, 1996, etc.). Even if such an assumption would hold (implying more cells tuned to full-front than ¾ side faces) in the human brain, both a delay *and* a weaker response to ¾ side faces would be predicted at the population level, which was not the case in our EEG findings (only a delay was found, but comparable response amplitudes to the two views).

The discrepancy between our human study and previous studies in the monkey brain might be accounted for by differences across the two species. Besides differences between the type of brain signals recorded (postsynaptic population activity for EEG *vs.* action potentials of single units in the majority of monkey studies), we note that there are considerable differences between humans' and monkeys' face processing abilities and cortical face networks, such that the monkey brain can hardly be considered as an adequate model of the human brain for this function. In particular, compared to humans, macaque monkeys are poor at face identity recognition (Griffin, 2020; Parr, Heintz, & Pradhan, 2008; Rossion & Taubert, 2019) and lack a ventral cortical face-selective network (Rossion & Taubert, 2019). This and other sources of evidence (e.g., Heywood & Cowey, 1992) suggest that monkeys' cortical face network in the STS may be particularly sensitive to the coding of head (and gaze) orientation, rather than for identifying faces. Future studies with higher spatial resolution methods such as intracerebral recordings in face-selective regions with a similar paradigm as used here (e.g., Jonas et al., 2016) may be able to provide better information regarding the putative spatial dissociation of face-selective neural responses to different head orientations in the human brain.

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

The processing time advantage for full-front as compared to ¾ side views may potentially be attributed to the symmetry and the visibility of internal features (e.g., eyes, mouth) in full-front faces, which play a key role in human generic face categorization (e.g., Omer et al., 2019; Paras & Webster, 2013). However, a potential confounding factor in our design is that head orientation is associated with different perceived gaze directions. While the full-front faces we used generally had a direct gaze (i.e., the face appeared looking at the observer), the ¾ faces were randomized between direct gaze and averted gaze (i.e., looking away from the observer). Although this scenario may be putatively similar to our natural experience, potential effects of gaze on speed of categorization cannot be ignored, as we are faster to detect a person looking at us than one ignoring us (gaze cueing effect, see for review, Frischen, Bayliss, & Tipper, 2007).

In particular, recent work in our laboratories (Retter et al., 2020, data not shown; Or, Goh, & Lee, under review) suggested that, in behavioural rapid face categorization experiments, manual response to direct gaze was faster than to averted gaze. However, Or, Goh, & Lee (under review) found that responses to ¾ faces with either gaze were both slower than full-front faces with direct gaze, consistent with our current EEG responses in the time domain. Although it would be difficult for the current design to disentangle head orientation and perceived gaze direction, we set out to reanalyse the time-domain data by separating the ¾ face epochs showing direct gaze (14 out of 30 such faces) from those showing averted gaze (16 out of 30 faces) in the ¾ condition (this was not performed on full-front faces as they all had direct gaze). The results suggested that perceived gaze direction did not result in a time shift in the waveforms; averted gaze resulted in significantly higher amplitude of response than direct gaze in the P2/P3-face component (211–250 ms) over one mediofrontal channel of AFF4h only (Supplementary Figure S4). There was a reversed tendency over the right occipito-temporal region (e.g., PO10), where direct gaze led to higher response amplitudes than averted gaze, but the differences were not significant. This ad hoc analysis suggested a potential influence of gaze on EEG responses. However, it should be noted that gaze was not deliberately studied here, where direct gaze and averted gaze were presented randomly (rather than periodically) over time. Thus, further studies in a more controlled setting (e.g., separate periodic sequences of ¾ faces with direct gaze or averted gaze, as well as full-front faces with either gaze) would be required to address this issue.

In summary, we found no evidence for differences in overall face categorization response amplitudes between full-front and ¾ side views of faces, suggesting that face-selective neural representation is as robust for these two views despite their differences in terms of diagnostic cues provided. This does not imply that all head orientations would generate the same magnitude of face-selective response, which could even be substantially reduced for full-profile faces and back head views. Despite the absence of amplitude difference, face-selective responses emerged earlier to full-front than ¾ side views by 8–13 ms. Overall, these findings indicate that face-selective neural representation is as robust for ¾ side faces as for full-front faces in the human brain, although full-front views provide an early processing-time advantage as compared to rotated face views.

Credit author statement

Charles C.-F. Or: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Visualization, Funding acquisition.

Talia L. Retter: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Review & Editing, Funding acquisition.

Bruno Rossion: Conceptualization, Methodology, Software, Validation, Resources, Data Curation, Writing - Review & Editing, Supervision, Project administration, Funding acquisition.

Author statement

No part of the study procedures and analyses was pre-registered prior to the research being conducted. We report how we determined our sample size, all data exclusions, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. The program code and data are publicly available at https://www.doi.org/10.17605/OSF.IO/YZHCP. Legal copyright restrictions prevent us from making the stimuli publicly available or sharing them on request with any individual outside the author team.

Open Practices

The study in this article earned an Open Data badge for transparent practices. Data for this study can be found at https://www.doi.org/10.17605/OSF.IO/YZHCP.

Acknowledgments

This work was supported by Singapore MOE AcRF Grant 2019-T1-001-060, NTU HASS Start-Up Grant, and F.R.S.-FNRS postdoctoral fellowship (FC 2773) to C. O., AcRF Grant 2018-T1-001-069 to C. O. & B. R., F.R.S.-FNRS doctoral grant to T. L. R. (FC 7159), and ERC grant to B. R. (facessvep 284025). We thank Denise Lim for generating the video.

Q5

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cortex.2021.03.030.

Uncited References

Q7

De Heering and Rossion, 2015, Or et al., .

REFERENCES

Axelrod, V., & Yovel, G. (2012). Hierarchical processing of face viewpoint in human visual cortex. *Journal of Neuroscience, 32*(7), 2442–2452. https://doi.org/10.1523/jneurosci.4770-11.2012

Baddeley, A., & Woodhead, M. (1983). Improving face recognition ability. In S. M. A. Lloyd-Bostock, & B. R. Clifford (Eds.), *Evaluating witness evidence: Recent psychological research and new perspectives* (pp. 125–136). Chichester, England: John Wiley & Sons.

Benton, C. P., Jennings, S. J., & Chatting, D. J. (2006). Viewpoint dependence in adaptation to facial identity. *Vision Research, 46*, 3313–3325. https://doi.org/10.1016/j.visres.2006.06.002

Bindemann, M., & Lewis, M. B. (2013). Face detection differs from categorization: Evidence from visual search in natural scenes. *Psychonomic Bulletin & Review, 20*(1), 1140–1145. https://doi.org/10.3758/s13423-013-0445-9

Booth, M. C. A., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cerebral Cortex, 8*(5), 510–523. https://doi.org/10.1093/cercor/8.6.510

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73*(1), 105–116. https://doi.org/10.1111/j.2044-8295.1982.tb01795.x

Bruce, V., Valentine, T., & Baddeley, A. (1987). The basis of the 3/4 view advantage in face recognition. *Applied Cognitive Psychology, 1*, 109–120. https://doi.org/10.1002/acp.2350010204

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305–327. https://doi.org/10.1111/j.2044-8295.1986.tb02199.x

Burton, A. M., & Bindemann, M. (2009). The role of view in human face detection. *Vision Research, 49*, 2026–2036. https://doi.org/10.1016/j.visres.2009.05.012

Butterworth, S. (1930). On the theory of filter amplifiers. *Experimental Wireless & the Wireless Engineer, 7*(85), 536–541.

Chen, J.-C., & Lien, J.-J. J. (2009). A view-based statistical system for multi0view face detection and pose estimation. *Image and Vision Computing, 27*, 1252–1271. https://doi.org/10.1016/j.imavis.2008.11.004

Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*, 10(4), 1–17. https://doi.org/10.1167/10.4.16, 16.

De Heering, A., & Rossion, B. (2015). Rapid categorization of natural face images in the infant right hemisphere. *eLife*, 4, 1–14. https://doi.org/10.7554/elife.06564. e06564.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience*, 4(8), 2051–2062.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, 1, 393–416. https://doi.org/10.1146/annurev-vision-082114-035518

Dzhelyova, M., & Rossion, B. (2014). The effect of parametric stimulus size variation on individual face discrimination indexed by fast periodic visual stimulation. *BMC Neuroscience*, 15(87), 1–12. https://doi.org/10.1186/1471-2202-15-87

Fang, F., Murray, S. O., & He, S. (2007). Duration-dependent fMRI adaptation and distributed viewer-centered face representation in human visual cortex. *Cerebral Cortex, 17*, 1402–1411. https://doi.org/10.1093/cercor/bhl053

Farfade, S. S., Saberian, M., & Li, L.-J. (2015). Multi-view face detection using deep convolutional neural networks. *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, 643–650. https://doi.org/10.1145/2671188.2749408

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science, 330*(6005), 845–851. https://doi.org/10.1126/science.1194908

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin, 133*(4), 694–724. https://doi.org/10.1037/0033-2909.133.4.694

Griffin, J. W. (2020). Quantifying the face inversion effect in nonhuman primates: A phylogenetic meta-analysis. *Animal Cognition, 23*, 237–249. https://doi.org/10.1007/s10071-019-01340-8

Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*, 187–203. https://doi.org/10.1016/s0896-6273(00)80832-6

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology, 35*(1), 96–111. https://doi.org/10.1152/jn.1972.35.1.96

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223–233. https://doi.org/10.1016/s1364-6613(00)01482-0

Heywood, C. A., & Cowey, A. (1992). The role of the 'face-cell' area in the discrimination and recognition of faces by monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 335(1273), 31–38. https://doi.org/10.1098/rstb.1992.0004

Hill, H., Schyns, P. G., & Akamatsu, S. (1997). Information and viewpoint dependence in face recognition. *Cognition, 62*, 201–222. https://doi.org/10.1016/s0010-0277(96)00785-8

Jacques, C., Retter, T. L., & Rossion, B. (2016). 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. *Neuroimage, 137*, 21–33. https://doi.org/10.1016/j.neuroimage.2016.04.045

Jeffery, L., Rhodes, G., & Busey, T. (2006). View-specific coding of face shape. *Psychological Science, 17*(6), 501–505. https://doi.org/10.1111/j.1467-9280.2006.01735.x

Jeffery, L., Rhodes, G., & Busey, T. (2007). Broadly tuned, view-specific coding of face shape: Opposing figural aftereffects can be induced in different views. *Vision Research, 47*, 3070–3077. https://doi.org/10.1016/j.visres.2007.08.018

Jonas, J., Jacques, C., Liu-Shuang, J., Brissart, H., Colnat-Coulbois, S., Maillard, L., et al. (2016). A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proceedings of the National Academy of Sciences, 113*(28), E4088–E4097. https://doi.org/10.1073/pnas.1522033113

Laganaro, M. (2014). ERP topographic analyses from concept to articulation in word production studies. *Frontiers in Psychology*, 5(493), 1–10. https://doi.org/10.3389/fpsyg.2014.00493

Liu, C. H., & Chaudhuri, A. (1998). Are there qualitative differences between face processing in photographic positive and negative? *Perception, 27*, 1107–1122. https://doi.org/10.1068/p271107

Liu, C. H., & Chaudhuri, A. (2002). Reassessing the 3/4 view effect in face recognition. *Cognition, 83*, 31–48. https://doi.org/10.1016/s0010-0277(01)00164-0

Logie, R. H., Baddeley, A. D., & Woodhead, M. M. (1987). Face recognition, pose and ecological validity. *Applied Cognitive Psychology*, 1(1), 53–69. https://doi.org/10.1002/acp.2350010108

Mouraux, A., Iannetti, G. D., Colon, E., Nozaradan, S., Legrain, V., & Plaghki, L. (2011). Nociceptive steady-state evoked potentials elicited by rapid periodic thermal stimulation of cutaneous nociceptors. *Journal of Neuroscience, 31*(16), 6079–6087. https://doi.org/10.1523/jneurosci.3977-10.2011

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Omer, Y., Sapir, R., Hatuka, Y., & Yovel, G. (2019). What is a face? Critical features for face detection. *Perception, 48*(5), 437–446. https://doi.org/10.1177/0301006619838734

Or, C. C.-F., Goh, B. K., & Lee, A. L. F. (under review). The roles of gaze and head orientation in face categorization during rapid serial visual presentation. Q6

Or, C. C.-F., Retter, T. L., & Rossion, B. (2019). The contribution of color information to rapid face categorization in natural scenes. *Journal of Vision, 19*(5), 1–20. https://doi.org/10.1167/19.5.20, 20.

O'Toole, A. J., Edelman, S., & Bülthoff, H. H. (1998). Stimulus-specific effects in face recognition over changes in viewpoint. *Vision Research, 38*, 2351–2363. https://doi.org/10.1016/S0042-6989(98)00042-X

Paras, C. L., & Webster, M. A. (2013). Stimulus requirements for face perception: An analysis based on "totem poles". *Frontiers in Psychology, 4*(18), 1–16. https://doi.org/10.3389/fpsyg.2013.00018

Parr, L. A., Heintz, M., & Pradhan, G. (2008). Rhesus monkeys (*Macaca mulatta*) lack expertise in face processing. *Journal of Comparative Psychology, 122*(4), 390–402. https://doi.org/10.1037/0735-7036.122.4.390

Perrett, D. I., & Oram, M. W. (1998). Visual recognition based on temporal cortex cells: Viewer-centred processing of pattern configuration. *Zeitschrift für Naturforschung, 53*, 518–541.

Perrett, D. I., Oram, M. W., & Ashbridge, E. (1998). Evidence accumulation in cell populations responsive to faces: An account of generalisation of recognition without mental transformations. *Cognition, 67*, 111–145.

Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., et al. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research, 86*, 159–173.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47*, 329–342. https://doi.org/10.1007/bf00239352

Perrett, D. I., Smith, P. A., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., et al. (1984). Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology, 3*(4), 197–208.

Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., et al. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society B: Biological Sciences, 223*, 293–317.

Quek, G. L., & Rossion, B. (2017). Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability. *Neuropsychologia, 104*, 182–200. https://doi.org/10.1016/j.neuropsychologia.2017.08.010

Ramírez, F. M., Cichy, R. M., Allefeld, C., & Haynes, J.-D. (2014). The neural code for face orientation in the human fusiform face area. *Journal of Neuroscience, 34*(36), 12155–12167. https://doi.org/10.1523/jneurosci.3156-13.2014

Ravidas, S., Ansari, A., & Kukreja, J. (2014). Multi-view face detection: A comprehensive survey. *International Journal of Computer Science and Mobile Computing, 3*(3), 1193–1203.

Retter, T. L., Jiang, F., Webster, M. A., & Rossion, B. (2020). All-or-none face categorization in the human brain. *Neuroimage, 213*(116685), 1–16. https://doi.org/10.1016/j.neuroimage.2020.116685

Retter, T. L., & Rossion, B. (2016a). Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia, 91*, 9–28. https://doi.org/10.1016/j.neuropsychologia.2016.07.028

Retter, T. L., & Rossion, B. (2016b). Visual adaptation provides objective electrophysiological evidence of facial identity discrimination. *Cortex; a Journal Devoted To the Study of the Nervous System and Behavior, 80*, 35–50. https://doi.org/10.1016/j.cortex.2015.11.025

Rossion, B., Alonso Pireto, E., Boremanse, A., Kuefner, D., & Van Belle, G. (2012). A steady-state visual evoked potential approach to individual face perception: Effect of inversion, contrast-reversal and temporal dynamics. *Neuroimage, 63*(3), 1585–1600. https://doi.org/10.1016/j.neuroimage.2012.08.033

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. *Journal of Vision, 11*(2), 1–21. https://doi.org/10.1167/11.2.16, 16.

Rossion, B., & Taubert, J. (2019). What can we learn about human individual face recognition from experimental studies in monkeys? *Vision Research, 157*, 142–158. https://doi.org/10.1016/j.visres.2018.03.012

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. *Journal of Vision, 15*(1), 1–18. https://doi.org/10.1167/15.1.18, 18.

Schneiderman, H., & Kanade, T. (2000). A statistical method for 3D object detection applied to faces and cars. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2000), 1*, 746–751.

Tovée, M. J., & Cohen-Tovée, E. M. (1993). The neural substrates of face processing models: A review. *Cognitive Neuropsychology, 10*(6), 505–528. https://doi.org/10.1080/02643299308253471

Troje, N. F., & Bülthoff, H. H. (1996). Face recognition under varying poses: The role of texture and shape. *Vision Research, 36*(12), 1761–1771. https://doi.org/10.1016/0042-6989(95)00230-8

Tsao, D. Y., Freiwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311*(5761), 670–674. https://doi.org/10.1126/science.1119983

Valentin, D., Abdi, H., & Edelman, B. (1997). What represents a face: A computational approach for the integration of physiological and psychological data. *Perception, 26*(10), 1271–1288. https://doi.org/10.1068/p261271

Van der Linde, I., & Watson, T. (2010). A combinatorial study of pose effects in unfamiliar face recognition. *Vision Research, 50*, 522–533. https://doi.org/10.1016/j.visres.2009.12.012

Xu, X., & Biederman, I. (2010). Loci of the release from fMRI adaptation for changes in facial expression, identity and viewpoint. *Journal of Vision, 10*(14), 1–13. https://doi.org/10.1167/10.14.36, 36.

Yang, S., Luo, P., Loy, C. C., & Tang, X. (2018). Faceness-Net: Face detection through deep facial part responses. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40*(8), 1845–1859. https://doi.org/10.1109/TPAMI.2017.2738644

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science, 256*(5061), 1327–1331. https://doi.org/10.1126/science.1598577

Please cite this article as: Or, C. C.-F et al., Does automatic human face categorization depend on head orientation?, Cortex, https://doi.org/10.1016/j.cortex.2021.03.030