

**Selective attention to faces in a rapid visual stream:
Hemispheric differences in enhancement and suppression of
category-selective neural activity**

Genevieve Quek*, Dan Nemrodov*, Bruno Rossion, Joan Liu-Shuang

Corresponding author:

Genevieve Quek

Psychological Sciences Research Institute and Institute of Neuroscience

University of Louvain

10 Place du Cardinal Mercier

Louvain-la-Neuve, 1348, Belgium

Email: genevieve.quek@uclouvain.be

Abstract

In daily life, efficient perceptual categorisation of faces occurs in dynamic and highly complex visual environments. Yet the role of selective attention in guiding face categorisation has predominantly been studied under sparse and static viewing conditions, with little focus on disentangling the impact of attentional enhancement and suppression. Here we show that attentional enhancement and suppression exert a differential impact on face categorisation supported by the left and right hemispheres. We recorded 128-channel EEG while participants viewed a 6Hz stream of object images (e.g., buildings, animals, objects...) with a face image embedded as every 5th image (i.e., OOOOFOOOOFOOOOF...). We isolated face-selective activity by measuring the response at the face presentation frequency (i.e., $6\text{Hz}/5 = 1.2\text{Hz}$) under three conditions: *Attend Faces*, in which participants monitored the sequence for instances of female faces; *Attend Objects*, in which they responded to instances of guitars, and *Baseline*, in which they performed an orthogonal task on the central fixation cross. During the orthogonal task, face-specific activity was predominantly centred over the right occipito-temporal region. Actively attending to faces enhanced face-selective activity much more evidently in the left hemisphere than in the right; whereas attending to objects suppressed the face-selective response in both hemispheres to a comparable extent. Additionally, the time-courses of attentional enhancement and suppression did not overlap. These results suggest the left and right hemispheres support face-selective processing in distinct ways – where the right hemisphere is mandatorily engaged by faces, the left hemisphere is more flexibly recruited to serve current tasks demands.

Introduction

Throughout our waking moments, the visual system is constantly bombarded by dynamically changing sensory input from the environment. Remarkably, however, perceptual categorisation within this overwhelming datastream happens rapidly and accurately. This is particularly true in the case of human faces – which, as objects of high ecological relevance, enjoy a privileged status in our visual system (Jonas & Rossion, 2016; Sergent, Ohta, & Macdonald, 1992). Effective social interaction depends critically on our ability to discriminate faces from a wide range of other perceptual categories (e.g., animals, plants, bodies, etc.), a complex and meaningful categorisation that the brain achieves between 100-200ms (Crouzet, Kirchner, & Thorpe, 2010; Crouzet & Thorpe, 2011; Rousselet, Mace, & Fabre-Thorpe, 2003). Although indisputably efficient, however, perceptual categorisation is far from capacity-free (Broadbent, 1958; Schneider & Shiffrin, 1977). Rather, visual information processing is limited such that multiple objects present in a scene must compete for neural representation (Kastner & Ungerleider, 2000). Selective attention allows us to cope with this visual competition, prioritising processing of information relevant to our current behavioural goals. Yet despite the critical role that selective attention plays in guiding behaviour in dynamic and complex environments, investigations of selective attention and face processing have often employed sparse and static viewing conditions. Typically, perceptual discriminations in these studies are binary (e.g., faces vs. houses), and stimuli are often spatially and/or temporally isolated (Baldauf & Desimone, 2014; Downing, Liu, & Kanwisher, 2001; Eimer, 2000a; Engell & McCarthy, 2010; Haxby et al., 1994; Holmes, Vuilleumier, & Eimer, 2003; Lueschow et al., 2004; Pessoa, McKenna, Gutierrez, & Ungerleider, 2002; Vuilleumier, Armony, Driver, & Dolan, 2001; Williams, McGlone, Abbott, & Mattingley, 2005; Wojciulik, Kanwisher, & Driver, 1998; Yi, Kelley, Marois, & Chun, 2006). In this way, existing studies do not impose the twin constraints that characterise effective generic face categorisation in the real world – namely, speed and high categorical

diversity, and as such, are limited in what they can tell us about how selective attention modulates face categorisation in natural vision.

Not only do extant studies of faces/attention utilise a simplified form of face categorisation, they also operationalise selective attention at a relatively coarse level. That is, although selectively attending to a specific visual feature is known to both *enhance* the neural response to the attended feature, and *suppress* the response to unattended features (Cohen & Maunsell, 2011; Ho, Brown, Abuyo, Ku, & Serences, 2012; Kastner & Ungerleider, 2000; Martinez-Trujillo & Treue, 2004; Treue & Trujillo, 1999), few face processing studies have attempted to disentangle these two attentional components. Instead, the classic approach has been to contrast face-processing under *maximally* and *minimally* attended conditions, by presenting two stimulus types concurrently (e.g., superimposed face and house images) and having participants selectively attend to one category at a time (Baldauf & Desimone, 2014; Downing et al., 2001; Eimer, 2000a; Engell & McCarthy, 2010; Furey et al., 2006; Haxby et al., 1994; Holmes et al., 2003; Lueschow et al., 2004; Pessoa et al., 2002; Vuilleumier et al., 2001; Williams et al., 2005; Wojciulik et al., 1998; Yi et al., 2006). Since attending to one category under these conditions necessarily involves actively ignoring the other, this approach can provide no insight into how the enhancement and suppression aspects of attention contribute to the overall effect. To separately characterise these two attentional components, the maximally and minimally attended conditions must be contrasted with a third condition, in which neither category is actively attended to (or ignored), i.e., an attentional baseline. Yet where such contrasts are commonplace in attention studies using low-level stimuli (Martinez-Trujillo & Treue, 2004; Treue & Trujillo, 1999), comparatively few face studies have taken this approach (but see Chadick & Gazzaley, 2011; Gazzaley et al., 2008; Gazzaley, Cooney, Rissman, & D'Esposito, 2005; Zanto, Hennigan, Östberg, Clapp, & Gazzaley, 2010).

That most studies of faces index selective attention at a global level has had a direct impact on the nature of the conclusions drawn about the role of this important cognitive mechanism in face perception. One such concerns the similarity in attentional modulation between the left and right face processing networks. It is now well-established that although normal observers process face information via a *bilateral* network of specialised regions in occipito-temporal cortex, this response to faces is stronger in the right hemisphere than in the left (Frassle et al., 2016; Jonas & Rossion, 2016; Kanwisher, McDermott, & Chun, 1997; Rossion, Hanseeuw, & Dricot, 2012; Sergent et al., 1992), a pattern of lateralisation which emerges very early in life (de Heering & Rossion, 2015). One possible consequence of the right hemisphere's *specialisation* for face perception could be a differential benefit of attentional allocation between the left and right face networks. Specifically, we might predict that processing supported by the less-efficient left hemisphere should be more sensitive to attentional enhancement than that supported by face-dominant right hemisphere¹. Studies that operationalise attention at a global level are unlikely to detect such a nuanced pattern of differential attentional sensitivity, and indeed existing investigations have by and large reported identical (or uncomparing) effects of selective attention in the two hemispheres (Baldauf & Desimone, 2014; Eimer, 2000a; Engell & McCarthy, 2010; Furey et al., 2006; Haxby et al., 1994; Holmes, Kiss, & Eimer, 2006; Holmes et al., 2003; Lueschow et al., 2004; Müsch et al., 2014; O'Craven, Downing, & Kanwisher, 1999; Pessoa et al., 2002; Vuilleumier et al., 2001; Williams et al., 2005; Wojciulik et al., 1998; Yi et al., 2006).

In this paper, we clarify these outstanding issues concerning the role of selective attention in face categorisation. Specifically, we sought to separately quantify attentional enhancement and suppression of face categorisation that is subject to the same constraints

¹ This is the case for the vertical asymmetry in face perception, in which face processing in the upper visual field is both faster and more robust to attentional influence than face processing in the lower visual field (Quek & Finkbeiner, 2014, 2016).

that underlie effective perceptual categorisation in the real world – namely, speed and categorical diversity. To this end, we recorded high-density electroencephalography (EEG) while participants viewed a continuous stream of object images taken from many different perceptual categories (e.g., animals, vehicles, man-made objects, trees, etc.). Images appeared at a rapid and strictly periodic rate of exactly 6 Hz (i.e., 6 stimuli/second), allowing just a single fixation per 167ms stimulus. Critically, we embedded a second periodicity in the sequence by inserting highly variable face stimuli as every fifth image, giving a face presentation frequency of $6\text{Hz}/5$, or 1.2 Hz². In line with previous studies using this approach, we expect this stimulation sequence to elicit two specific responses in the EEG spectrum: One at 6 Hz, reflecting visual processing common to both object and face images (referred to here as the *common response*), and one at 1.2 Hz (i.e., $6\text{Hz}/5$), reflecting the differential response to faces as compared to objects (Jacques, Retter, & Rossion, 2016; Retter & Rossion, 2016; Rossion, Torfs, Jacques, & Liu-Shuang, 2015). We refer to this 1.2 Hz response as the face-selective response, as it can *only* arise if the neural response evoked by each briefly presented face consistently differs from that evoked by the many other object categories appearing in the sequence. In this way, the 1.2 Hz signal captures high-level face-selective processing without the need for conditional subtraction (e.g., face activity – object activity). We compared this 1.2 Hz face selective response under three conditions of task-based attention: On ‘Baseline’ trials, participants performed the same orthogonal task used in all previous studies (Retter & Rossion, 2016; Rossion et al., 2015), in which they responded each time the central fixation cross overlaid on the images changed colour. On ‘Attend Faces’ trials, participants covertly monitored the face images in the sequence and responded each

² Note that we are not the first to employ a frequency-tagging approach to the study of selective attention, however this work has focused almost exclusively on low-level stimuli and simple features (Chen, Seth, Gally, & Edelman, 2003; Keil, Moratti, Sabatinelli, Bradley, & Lang, 2005; Morgan, Hansen, & Hillyard, 1996; Müller et al., 2006; Müller & Hübner, 2002; Müller et al., 1998; Norcia, Appelbaum, Ales, Cottureau, & Rossion, 2015; Wang, Clementz, & Keil, 2007).

time they detected a female face. On ‘Attend Objects’ trials, they monitored the object images and responded each time they detected a guitar. The probability of presentation for each target type was constant in every sequence (always five of each target type), such that only the focus of participants’ task-based attention changed on each trial. Critically, from these conditions we calculated indices of attentional enhancement (Attend Faces – Baseline) and attentional suppression (Attend Objects – Baseline) for occipito-temporal regions of interest in both the left and right hemispheres.

To anticipate our results, we found that when observers were engaged in an orthogonal task (i.e., Baseline), natural face images in our sequences activated the right hemisphere much more strongly than the Non-Preferred hemisphere. Actively attending to faces *enhanced* the face-selective response, much more evidently in the left hemisphere than in the right hemisphere. Actively attending to a stimulus category other than faces served to *suppress* the face-selective response in both hemispheres to the same extent. Interestingly, the temporal dynamics of attentional enhancement and suppression differed across the temporal unfolding of the face-selective response. These results suggest that face-selective regions in the two hemispheres may support face categorisation in distinct ways – whereas the right hemisphere may be mandatorily engaged by faces, the left hemisphere appears to be flexibly recruited to serve current tasks demands.

Methods

Participants

Twenty adult participants took part in this study in exchange for monetary compensation. Three were excluded due to technical issues during EEG recording and two were excluded due to low behavioural performance (< 80% accuracy in one of the conditions). The final sample consisted of 15 participants (age = 22 ± 2.63 , 7 females). All were right-handed and had normal or corrected-to-normal vision. None reported any history

of neurological or psychiatric disease. We obtained written informed consent prior to testing in accordance with the guidelines set out by the Biomedical Ethical Committee of the University of Louvain.

Stimuli

The stimulus set consisted of 44 face images in total and 250 images of various non-face objects in total (animals, plants, man-made objects, and houses), all collected from the internet and used in previous studies (de Heering & Rossion, 2015; Retter & Rossion, 2016; Rossion et al., 2015). Each image was converted to greyscale, resized to 200 x 200 pixels, and equalised in terms of mean pixel luminance and RMS contrast in MatLab (MathWorks, USA). Both faces and objects were left embedded in their original naturalistic background and varied in their size, position, viewpoint, and lighting. Target images were female faces (13 individual exemplars) and guitars (15 individual exemplars). The full image set can be downloaded from our website (<http://face-categorisation-lab.webnode.com/resources/natural-face-stimuli/>).

Design

The current design was similar to that reported in previous studies (Rossion et al., 2015). We used PsychToolbox running on MATLAB R2009a (MathWorks, USA) to present stimuli at a periodic rate of exactly 6 Hz (i.e., 6 images per second). Each stimulation cycle lasted 167 ms and began with a uniform grey background from which an image gradually appeared and disappeared as its contrast respectively increased and decreased (0-100-0%). We used a sinusoidal contrast modulation since it can be described with a single parameter (SOA), and gives a smoother, virtually continuous visual stimulation, with only one frame (8.33 ms) per cycle in which the contrast is at 0% (Movie 1) (Liu-Shuang, Norcia, & Rossion, 2014; Rossion et al., 2015). Throughout the sequence, a small black fixation cross overlaid the images. Each 60-s stimulation sequence consisted of randomly selected object images

(without guitars) with a randomly selected male face interleaved as every 5th image. Thus, faces appeared periodically at a frequency of exactly 6 Hz/5 (i.e., 1.2 Hz). Periodic EEG responses at the 1.2 Hz frequency and its harmonics reflect the differential response to faces as compared to objects (i.e., face categorisation), while responses at the base stimulation frequency of 6 Hz and harmonics reflect visual processing common to all stimuli (for an overview of how harmonic responses arise in the FPVS frequency spectrum, see Retter & Rossion, 2016). On any given sequence, participants performed one of three behavioural tasks (conditions). In the *Baseline* condition, participants attended to the central fixation cross and were instructed to press the spacebar whenever it changed colour from black to red (duration = 200 ms). In the *Attend Faces* condition, participants monitored the face images in the stimulation sequence and responded whenever they saw a female face. We reasoned that this task would increase attention to all faces in the sequence, since in order to discriminate between male and female faces, an observer must first identify that a stimulus is a face. Finally, in the *Attend Objects* condition, participants monitored the object images in the stimulation sequence and responded when they saw an image of a guitar. Female face targets randomly replaced male faces, and guitar targets randomly replaced another object image. Targets were distributed throughout the whole sequence (time range between consecutive targets = 6.88 – 15.98 s) to ensure that observers maintained attention for the entirety of the 60-s sequences. Importantly, all target types occurred in every sequence the same number of times (i.e., each contained five fixation cross changes, five female faces, and five guitars), ensuring equal probability of each target type on each trial. In this way, only the participant's active task differed from trial to trial, while the visual stimulation itself was held constant. There were 4 x 60 s trials per condition, making for 12 trials in total (total testing time = 15-20 mins, including breaks). There were several pseudo-random trial orders assigned to each participant in a counterbalanced order.

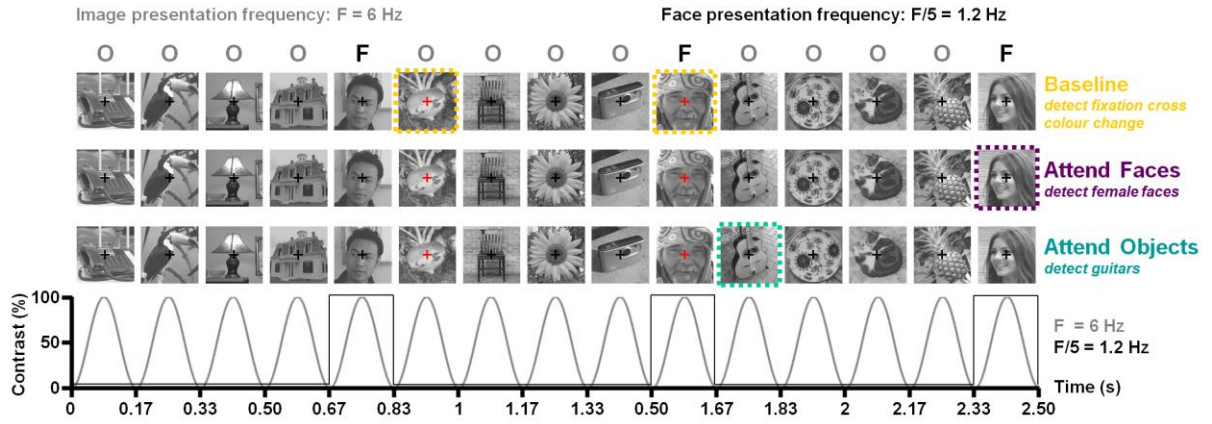


Figure 1. Schematic illustration of the experimental design. Images were presented through sinusoidal contrast modulation at fixed rate of 6 images per second (6 Hz). The 60-s stimulation sequences contained various natural and man-made objects (O), and a face as every 5th stimulus (F). There were three target types contained within each sequence: the central black fixation cross turning red, female faces, and guitars, each of which occurred randomly five times within the sequence. On any given trial, participants performed one of three behavioural tasks: detect fixation cross changes (*Baseline*), detect the appearance of a female face (*Attend Faces*), or detect the appearance of a guitar (*Attend Objects*).

EEG acquisition

The experiment was run in a quiet, low-lit room. Participants sat 80 cm away from an LED monitor (BenQ XL2420T) with a 1920 x 1080 resolution and a 120 Hz refresh rate. A curtain isolated the participant from the experimenter; participant behaviour was monitored with a webcam. Stimuli appeared centrally and subtended 3.93° of visual angle. We used the ActiveTwo Biosemi system (Biosemi, Amsterdam, The Netherlands) to acquire high-density 128-channel EEG at a 512 Hz sampling rate. The magnitude of the offset of all electrodes, referenced to the common mode sense (CMS), was held below 50 μV . Four additional flat-type Active-electrodes recorded vertical and horizontal electro-oculogram (EOG): two above and below the participant's right eye and two lateral to the external canthi.

Behavioural Analysis

We calculated response time (RTs) relative to target onset, and considered responses to be accurate if they occurred between 250 ms to 1500 ms following target onset. Only accurate responses were taken into account for the RT analysis. We also calculated an inverse efficiency score (correct RT/accuracy) to take into account a speed-accuracy trade-off.

EEG Analysis

Pre-processing.

We analysed the EEG data using open source software (Letswave5 <http://www.nocions.org/letswave/>) running in MATLAB R2012b (MathWorks, USA), with similar processing steps and parameters as in previous studies (e.g., Retter & Rossion, 2016; Rossion et al., 2015). We firstly band-pass filtered the EEG data between 0.1 Hz and 100 Hz using a 4th order Butterworth filter and downsampled it to 256 Hz for faster processing. We then segmented the continuous EEG trace relative to the starting trigger of each trial, including an additional 2 seconds before and after each sequence. We removed blink artefacts using Independent Components Analysis (ICA) performed with a square mixing matrix (Jung et al., 2000). For each participant, we removed the single component corresponding to blinks based on the visual inspection of the topography and time-course. We removed additional artefacts by interpolating bad channels with the three neighbouring channels. No more than 5% of channels were interpolated for any given participant (i.e., maximum 6 channels out of 128). We re-referenced the clean data to the average of all scalp channels, and averaged each participant's trials by condition. Electrode labels were changed to closely match a more conventional 10/20 system (see Rossion et al., 2015, Figure S2 for exact re-labeling.).

Frequency-domain analysis.

To avoid spectral leakage, we re-segmented the pre-processed EEG data into epochs containing an integer number of cycles of the face presentation frequency (i.e., 1.2 Hz =

0.0833-s per cycle). We discarded the first and last two seconds of each trial to remove eye-movements and transients related to the abrupt onset and offset of the flickering stimuli. The final cropped epochs were 55.84 s long and contained 67 face presentation cycles. We subjected these to Fast Fourier Transformation (FFT) and extracted the amplitude spectra with a frequency resolution of 0.018 Hz (i.e., $1/55.84$).

To establish the presence of significant periodic EEG responses at the relevant stimulation frequencies in both individual participants and at the group-level, we pooled the amplitude spectra across all scalp channels and calculated Z-scores. The Z-scores at a given frequency were computed as the difference in amplitude between that frequency and the mean of the 20 neighbouring frequency bins, divided by the standard deviation of the 20 neighbouring bins. The 20 neighbouring bins represented a frequency range of 0.36 Hz (0.18 Hz on either side) and excluded the two immediately adjacent frequency bin (Retter & Rossion, 2016; Rossion, Prieto, Boremanse, Kuefner, & Van Belle, 2012; Rossion et al., 2015; Srinivasan, Russell, Edelman, & Tononi, 1999). As per previous studies (Jacques et al., 2016), we considered Z-scores greater than 3.1 ($p < 0.001$, one-tailed, i.e., signal > noise) to be significant.

We quantified the size of periodic EEG responses in two steps. First, we applied a baseline-correction to the raw amplitude spectra. For each frequency bin, we subtracted the mean amplitude of the 20 surrounding frequency bins (again excluding the two immediately adjacent bins) from the amplitude at that frequency (Jacques et al., 2016; Retter & Rossion, 2016). This enabled us to quantify the magnitude of the response at each individual relevant harmonic frequency³. Second, we assessed responses at the global level by summing the baseline-corrected amplitudes across the relevant frequencies (Retter & Rossion, 2016). We

³ Harmonic responses of frequencies of interest arise since the periodic EEG response is not purely sinusoidal, but rather a complex, multi-component waveform (for details, see Norcia et al., 2015; Regan, 1989; Retter & Rossion, 2016).

calculated a *face-selective response* by summing the response at the first 8 harmonics of the face presentation frequency, excluding the 5th harmonic which was actually the 6 Hz base stimulation frequency (i.e., 1.2 Hz, 2.4 Hz, 3.6 Hz, 4.8 Hz, 7.2 Hz, 8.4 Hz, 9.6 Hz, & 10.8 Hz). We selected these harmonics of 1.2 Hz as they were the most consistently present in all participants. We also calculated a *common response* by summing the response amplitudes across the first 3 harmonics of the base stimulation frequency (i.e., 6 Hz, 12 Hz, and 18 Hz) (note that responses at harmonics above the 20 Hz frequency range were largely decreased).

Following inspection of scalp topographies, and on the basis of several of our previous studies that have used the exact same stimulation parameters (Jacques et al., 2016; Rossion et al., 2015), we defined one occipito-temporal region-of-interest (ROI) for each hemisphere: right channels were P8, P10, PO8, PO10, PO12, and left channels were P7, P9, PO7, PO9, PO11. Given that studies using this paradigm often identify a few observers whose face-selective response is left lateralised (roughly 2-3 participants out of 16, e.g., Retter & Rossion, 2016), we took account of individual lateralisation patterns by identifying each participant's Preferred and Non-Preferred hemisphere, where preferred refers to the hemisphere with the strongest face-selective response. We then computed indices of attentional enhancement (Attend Faces – Baseline) and attentional suppression (Baseline – Attend Objects) within each ROI. We carried out statistical analyses using repeated measures ANOVAs with Greenhouse-Geisser corrections applied to degrees of freedom whenever the assumption of sphericity was violated. We used pairwise *t*-tests for post-hoc comparisons and unless specified otherwise, all *p*-values were two-tailed. We used a Bonferroni correction to control for multiple comparisons where necessary.

Time-domain analysis.

We also inspected periodic EEG responses in the time-domain (Jacques et al., 2016; Rossion et al., 2015). Here we low-pass filtered the re-referenced data with a 30 Hz cut-off

(4th order zero-phase Butterworth filter) and cropped each sequence to be an integer number of cycles of the face presentation frequency (0 – 58 s, 14722 bins = 69 face presentation cycles). To remove aspects of the data common to both face *and* object processing, we applied a multi notch-filter (width = 0.1 Hz) that encompassed five harmonics of the base stimulation frequency (i.e., 6 Hz – 36 Hz). We then averaged sequences by condition, and starting from 3 s after the start of the sequence, segmented smaller epochs containing 5 stimulation cycles (\approx 832 ms), corresponding to four objects and one face presentation (*OFOOO*). After averaging across these resulting smaller epochs, we performed a baseline-correction relative to the first object cycle (-167 – 0 ms). As a final step, we calculated indices of attentional enhancement (Attend Faces – Baseline) and attentional suppression (Baseline – Attend Objects) within *i*) the ROIs from the frequency-domain analysis (Standard ROIs), and *ii*) larger ROIs including more dorsal and medial occipital channels selected based on visual inspection of time-domain components (referred to here as Wide ROIs). We identified statistically significant timepoints of enhancement and suppression by asking whether the 99% confidence interval at each time-point excluded zero. Only clusters containing minimum five consecutively significant time-points (\sim 19 ms) between 100 – 600 ms post-stimulus were considered.

Phase analysis.

We inspected the phase of the individual harmonics of the face-selective response in Matlab using the Circular Statistics Toolbox (Berens, 2009). For each participant, we averaged the complex values output of the FFT within each ROI, and then grand-averaged across participants to plot the mean amplitude and phase of the relevant face-selective response harmonics. Since the first three harmonics were the strongest, and because phase estimation depends on signal strength, we focused on this subset of frequencies (i.e., 1.2 Hz, 2.4 Hz, & 3.6 Hz) to examine phase shifts across attentional conditions. For each harmonic,

we computed the difference in phase between the Attend Faces vs. Baseline conditions (attentional enhancement) and the Baseline vs. Attend Objects conditions (attentional suppression) for each participant. These scores were then averaged with the *circ_mean* function and converted to degrees. We used the same function to calculate the grand-average phase-shift across harmonics.

Results

Behavioural performance

Behavioural performance is summarised in Figure 2. Repeated measures ANOVAs with *Condition* (Baseline, Attend Faces, Attend Objects) as a within-subject factor showed a significant difference between conditions in terms of RT ($F(1.45, 20.28) = 8.61, p < 0.004, \text{partial } \eta^2 = 0.38$) and accuracy ($F(1.44, 20.18) = 5.37, p < 0.02, \text{partial } \eta^2 = 0.28$). In both cases, there was no significant difference between the Attend Faces and Attend Objects conditions (RT: $t(14) = 0.33, p = 1$; accuracy: $t(14) = -0.34, p = 1$). Although accuracy in the Baseline condition was at ceiling relative to the Attend Faces and Attend Objects conditions (Baseline vs. Attend Faces: $t(14) = 2.94, p < 0.03$; Baseline vs. Attend Objects: $t(14) = 3.95, p < 0.003$), RT was also significantly higher in this condition compared to the other two (Baseline vs. Attend Faces: $t(14) = 3.86, p < 0.006$; Baseline vs. Attend Objects: $t(14) = 2.93, p < 0.03$). Hence, these differences were driven by a speed-accuracy trade-off in the Baseline condition, as demonstrated by the lack of difference between conditions when using inverse efficiency scores ($F(2, 28) = 0.37, p = 0.69, \text{partial } \eta^2 = 0.03$). As such, there was no evidence that participants' behavioural performance varied meaningfully across conditions.

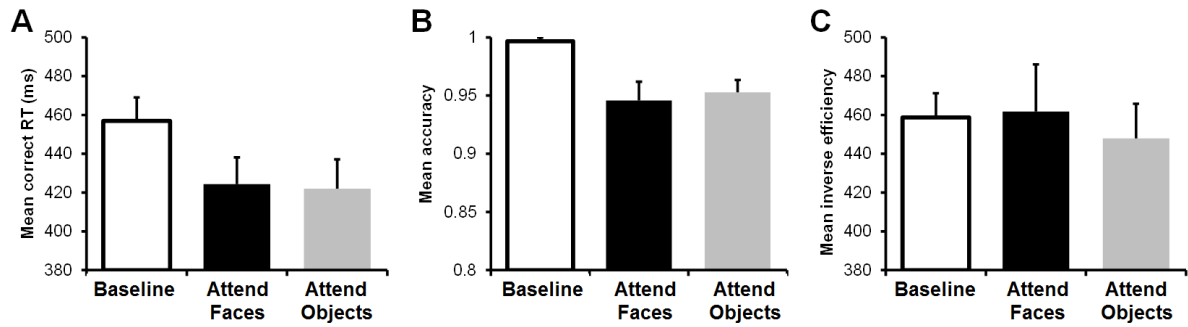


Figure 2. Behavioural results by condition. **A.** Mean RT for correct trials. **B.** Mean accuracy. **C.** Mean inverse efficiency (correct RT/accuracy). Error bars represent the standard error of the mean.

Periodic EEG Responses: Frequency-Domain

Response significance.

There were large peaks at the frequencies of base stimulation (6 Hz) and face presentation (1.2 Hz), as well as at the harmonics (i.e., integer multiples) of these frequencies (see Figure 3A). To avoid task- or channel-related biases, we averaged across all conditions and channels before determining the range of frequencies to consider for quantification. At both the group and individual participant level, we observed significant responses at multiple harmonics of the base stimulation frequency and the face presentation frequency. General visual responses were mostly distributed over the first 3 harmonics (6 Hz, 12 Hz, and 18 Hz) and face categorisation responses were most consistent across participants within the range of the first 9 harmonics (1.2 Hz to 10.8 Hz, excluding 6 Hz). Further analyses therefore concentrated on these frequencies.

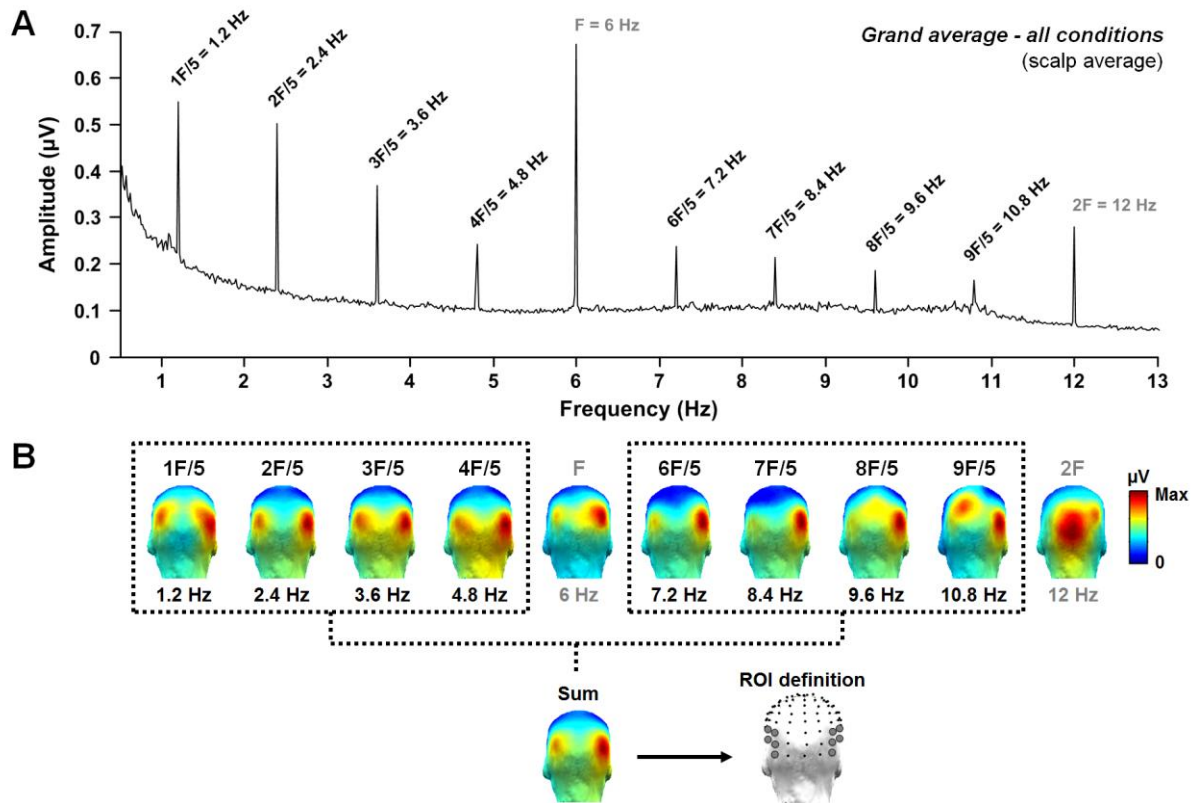


Figure 3. Group level periodic EEG responses. **A.** The amplitude spectrum averaged across all conditions and channels, with labelling over peaks at frequencies of interest. There were large significant peaks at the face presentation frequency (1.2 Hz) and its harmonics (2.4 Hz, 3.6 Hz...), as well as at the base stimulation frequency (6 Hz) and its second harmonic (12 Hz; third harmonic not shown here). **B.** Topographical distribution of responses at each relevant frequency. The colour-scale is adjusted to the maximal value of individual frequencies. Responses at the face presentation frequency harmonics show a consistent bilateral occipito-temporal topography. Dotted boxes indicated the 8 frequencies whose responses were summed to give the *face-selective response* (Retter & Rossion, 2016) used for all subsequent analyses. Below is the topography of this overall face-selective response, used to define bilateral occipito-temporal ROIs.

Face-selective response.

Having identified the relevant frequency range, we quantified the overall face categorisation response in each condition by summing responses at the first 9 harmonics (excluding the 5th harmonic, 6 Hz, which is confounded with the base stimulation frequency). The magnitude of this face-selective response fluctuated across conditions. First, we analysed

the data minimising spatial bias by considering the scalp-averaged response (128 channels). Grouped this way, the data showed a clear and significant effect of *Condition* ($F(1.16, 16.25) = 30.73, p < 0.001$), as face-selective responses were increased in the Attend Faces condition relative to the Baseline condition, $t(14) = 4.22, p < 0.003$ (i.e., enhancement), and decreased in the Attend Objects condition relative to the Baseline condition, $t(14) = 6.01, p < 0.0001$ (i.e., suppression). As expected, responses in the Attend Faces condition were also significantly larger than the Attend Objects condition, $t(14) = 6.75, p < 0.0001$.

Next, we used a region-of-interest (ROI) approach by focusing on channels where the face-selective response was maximal. Collapsing across conditions, the face-selective response exhibited a stable bilateral occipito-temporal topography with a clear right hemisphere preference (Figure 3B), with the exception of two participants (S07 and S15) who showed a left-lateralised face-selective response (Figure 4; Table 1). As such, we created two regions-of-interest (ROIs) composed of right (P8, P10, PO8, PO10, PO12) and left (P7, P9, PO7, PO9, PO11) occipito-temporal channels. On average, the response in the right ROI was 21% stronger than the response in the left ROI (30% stronger when the two left-lateralised participants were excluded). When considered as a function of condition, visual inspection of the group level topographies revealed that participants' task modulated the relative contribution of the two hemispheres to the face-selective response. The right hemisphere dominance was most evident in the Baseline condition, both at the group-level and for the majority of the individual participants (Figure 4; Table 1). This pattern is in line with previous studies that have used the same orthogonal task in this paradigm (Rossion et al., 2015; Jacques, Retter, & Rossion, 2016; Retter & Rossion, 2016; see Jonas et al., 2016 for intracerebral recording evidence), as well as the general right hemispheric dominance of face processing (Kanwisher et al., 1997; Sergent et al., 1992). In contrast, actively attending to faces generated a more bilateral response profile (Figure 4).

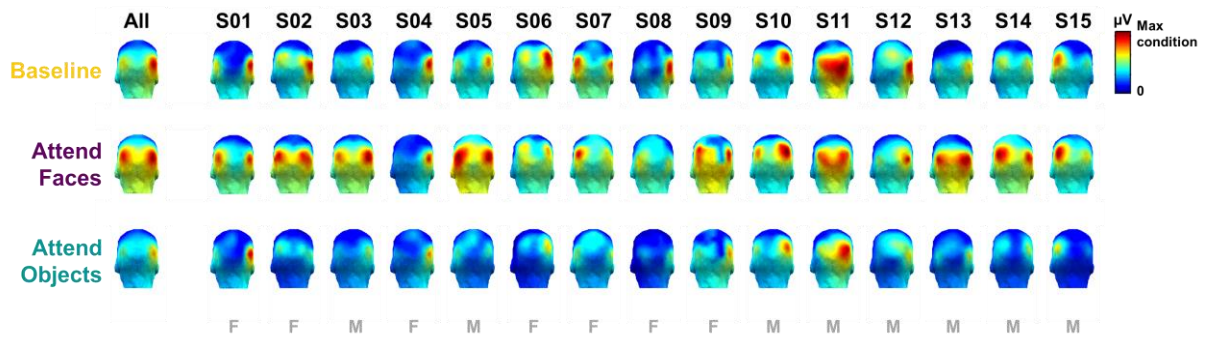


Figure 4. Topographies of the face-selective response across conditions at the group-level (far left) and in individual participants. Colour scales are adjusted according to the maximal value across conditions within each individual participant; participant gender (M/F) is indicated below each column. The face-selective response in the Baseline condition was right lateralised for 13/15 participants. Two participants (S07 & S15) were left-lateralised. Regardless of the pattern of lateralisation, attending to faces recruited the contralateral hemisphere (e.g., compare S01 vs. S15, see also Table 1).

Table 1

Face-selective response lateralisation index (right ROI – left ROI) in μV .

Participant	Average all conditions	Baseline	Attend Faces	Attend Objects
1	1.74	1.92	0.18	3.11
2	0.58	2.12	-0.64	0.27
3	1.14	1.33	0.98	1.10
4	4.90	5.58	4.97	4.14
5	0.26	1.09	-0.14	-0.18
6	1.25	1.68	0.74	1.31
7	-0.53	-0.64	-0.79	-0.15
8	1.15	2.57	-0.18	1.05
9	0.65	0.83	0.26	0.87
10	1.29	1.27	0.87	1.74
11	1.06	0.86	0.29	2.04
12	1.82	2.52	1.27	1.68
13	0.96	1.34	0.55	1.00
14	0.53	0.56	0.30	0.72
15	-1.32	-1.63	-1.00	-1.33

We extracted the face-selective response in each ROI to statistically test the differential effects of attention within each hemisphere (Figure 5A). Because of inter-individual differences in response lateralisation, we defined the ROI with the strongest face-selective response (averaging across all conditions) as the Preferred hemisphere (the right hemisphere for 13/15 participants) and the contralateral ROI as the Non-Preferred hemisphere. A repeated measures ANOVA with *Condition* (Baseline, Attend Faces, Attend Objects) and *Hemisphere* (Preferred, Non-Preferred) as within-subject factors showed significant main effects of *Condition* ($F(2,28) = 24.36, p < 0.001, \text{partial } \eta^2 = 0.63$) and *ROI* ($F(1,14) = 20.39, p < 0.001, \text{partial } \eta^2 = 0.59$), which were qualified by a significant

Condition x *Hemisphere* interaction ($F(2,28) = 9.21, p < 0.001, \text{partial } \eta^2 = 0.40$)⁴. This interaction was due to the presence of differential attentional modulation in the two hemispheres. In the Preferred hemisphere, responses in the Attend Faces and Baseline conditions did not differ ($t(14) = -0.12, p = 1$), but the response in both these conditions was larger than that in the Attend Objects condition (Attend Objects vs. Baseline: $t(14) = 4.61, p < 0.002$; Attend Objects vs. Attend Faces: $t(14) = 3.87, p < 0.01$). In the Non-Preferred hemisphere, the Baseline ($t(14) = 6.70, p < 0.001$) and Attend Faces conditions ($t(14) = 7.15, p < 0.001$) also both evoked a larger response than the Attend Objects condition. However, here the response in the Attend Faces condition was larger than the response in the Baseline condition ($t(14) = -4.07, p < 0.007$). In other words, when attention was selectively allocated to faces, responses increased relative to baseline over the Non-Preferred hemisphere, but not over the Preferred hemisphere.

To further test the differential effect of attention in each hemisphere, we computed indices of attentional modulation (Figure 5B). Attentional enhancement (Attend Faces - Baseline) was stronger for the Non-Preferred hemisphere than the Preferred hemisphere ($t(14) = -4.46, p < 0.001$). In the Non-Preferred hemisphere, face-selective responses increased by around 40% on average when participants explicitly attended to faces within the stimulation sequence. On the other hand, attentional suppression (Attend Objects - Baseline) did not differ across the two hemispheres ($t(14) = -1.58, p = 0.146$). On average, the face-

⁴ Note that the same analysis comparing the left and right hemispheres directly yielded identical results: A repeated measures ANOVA with *Condition* (Baseline, Attend Faces, Attend Objects) and *Hemisphere* (right, left) as within-subject factors showed significant main effects of *Condition* ($F(2,28) = 24.36, p < 0.001, \text{partial } \eta^2 = 0.63$) and *Hemisphere* ($F(1,14) = 8.82, p < 0.01, \text{partial } \eta^2 = 0.39$), as well as a significant *Condition* x *Hemisphere* interaction ($F(2,28) = 7.83, p < 0.002, \text{partial } \eta^2 = 0.36$). Within the left hemisphere, planned pairwise comparisons indicate that all conditions differed significantly (p-values range = 0.001 – 0.005), with the largest responses occurring in the Attend Faces condition. However, in the right hemisphere, response amplitudes were similar in the Baseline and Attend Faces conditions ($t(14) = 0.25, p = 1$), and both were larger than in the Attend Objects condition (p-values range = 0.001 – 0.006).

selective response in both the left and right ROI decreased by around 25% when attention was directed towards a category other than faces (i.e., towards objects).

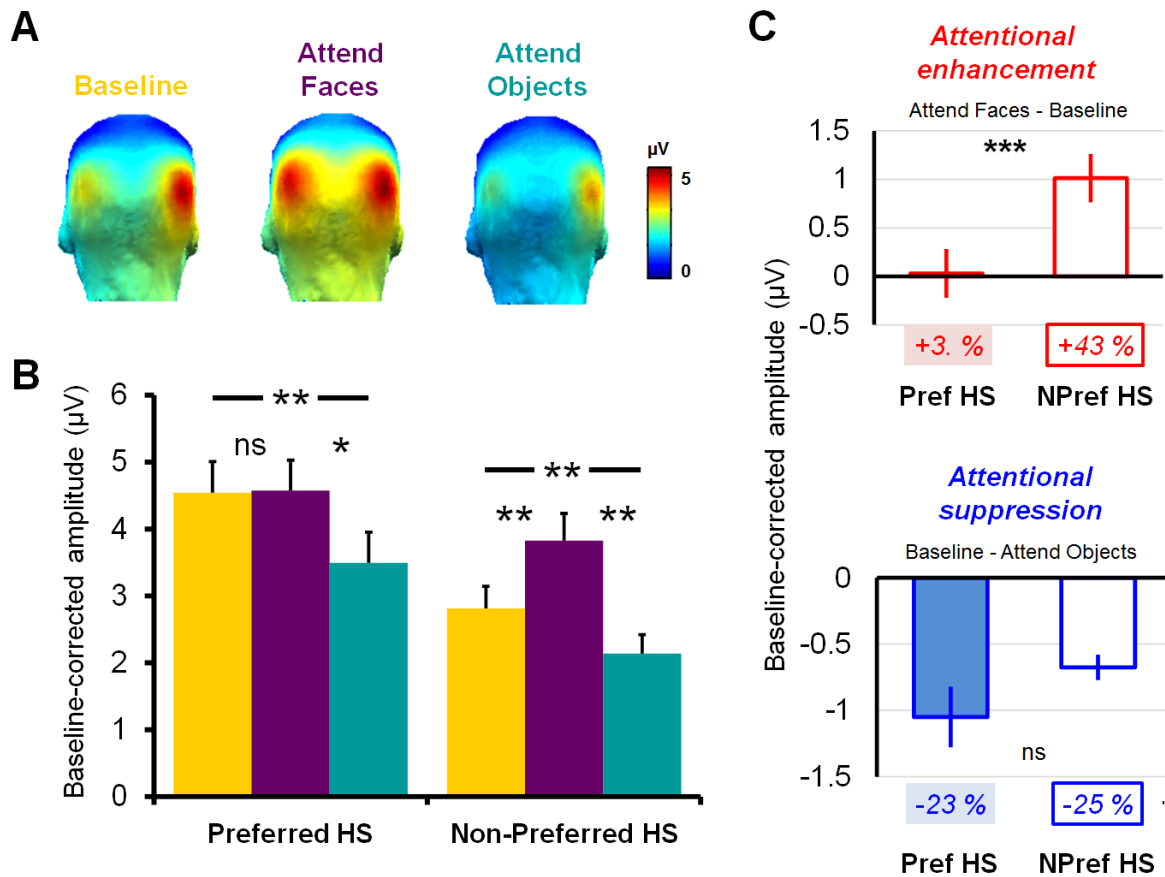


Figure 5. **A.** Group-level topographies corresponding to the face-selective response in the three conditions. **B.** Mean face-selective response amplitude as a function of Hemisphere and Condition. **C.** Indices of attentional enhancement and suppression for the Preferred and Non-Preferred hemispheres. Values are the amplitude difference expressed as percentage response increase (top) and decrease (bottom) relative to Baseline. Error bars represent standard error of the mean. ns = non-significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Common response.

Responses at the base stimulation frequency and its harmonics represent visual processing common to all images, both objects and faces. In this way, this general visual response provides a baseline measure of how well stimuli were perceived across different task conditions (Figure 6). The spatial topography of the base stimulation frequency

harmonics was variable, as in previous studies (Jacques et al., 2016; Rossion et al., 2015). Since we were primarily interested in determining whether the common response differed across attentional task conditions, we considered responses summed over the first three harmonics of the base stimulation rate (6 Hz, 12 Hz, & 18 Hz). We first tested for potential common response differences in the same occipito-temporal ROIs as those used in the analysis of the face-selective response. A repeated measures ANOVA with *Condition* (Baseline, Attend Faces, Attend Objects) and *Hemisphere* (Preferred, Non-Preferred) as within-subject factors did not yield any significant main effects of Condition ($F(1.37, 19.16) = 1.68, p = 0.21, \text{partial } \eta^2 = 0.11$) or of Hemisphere ($F(1,14) = 4.1, p = 0.06, \text{partial } \eta^2 = 0.23$), nor were there any significant Condition x Hemisphere interactions ($F(1.31, 18.33) = 0.76, p = 0.43, \text{partial } \eta^2 = 0.05$). However, given that in contrast to the face-selective response, the common response was located over more (right) dorsal and medial occipital channels, we also ran a second analysis within a medial occipital ROI defined specifically for the common response (PPO6, PO8, PO10, POO6, O2, POI2, Oz, Oiz). This additional analysis also gave no evidence that common response amplitudes were modulated by attentional task ($F(1.37, 19.19) = 3.0, p = 0.088, \text{partial } \eta^2 = 0.18$). Hence, stimuli appeared to be equally well perceived in all conditions, i.e., regardless of the attentional task.

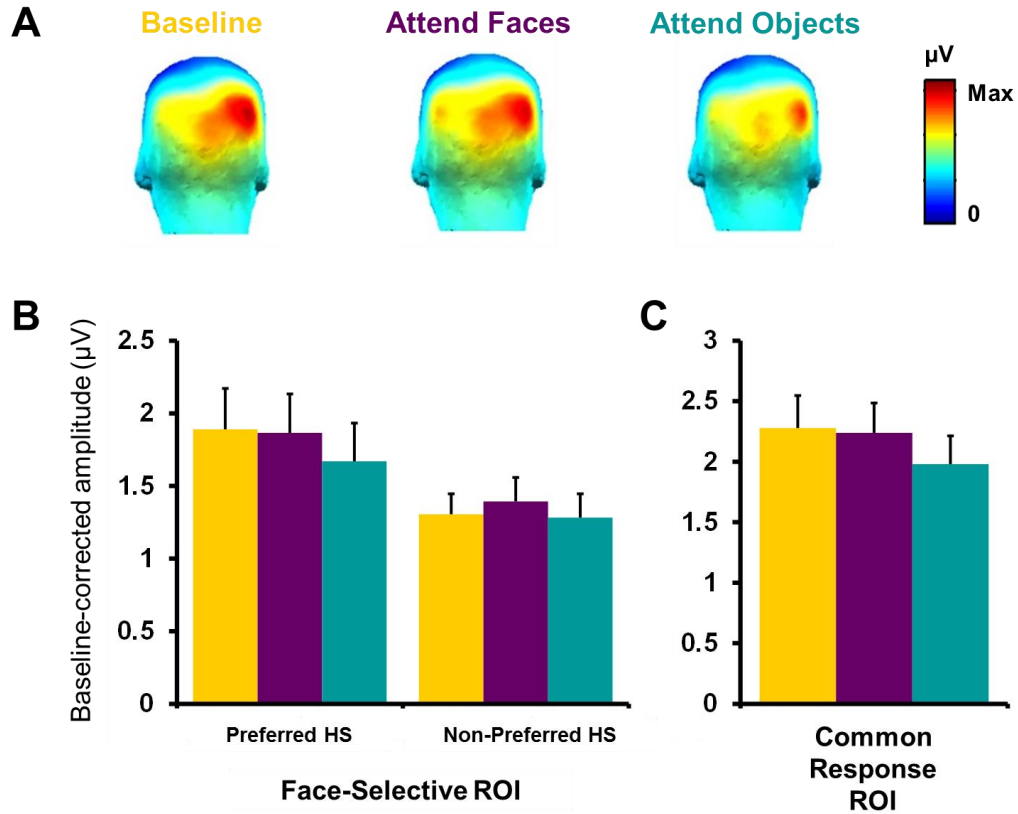


Figure 6. Common response analysis. A. Group-level topographies of the common response across conditions; colours scaled according to the maximal value of each condition. B. Mean amplitudes within Preferred and Non-Preferred Hemisphere ROIs. C. Mean amplitudes within a medial occipital common response ROI (see text). Error bars represent standard error of the mean. No common response differences between conditions were found.

Periodic EEG Responses: Time-Domain

Next, we sought to examine the temporal dynamics of attentional modulation by analysing the face-selective response in the time-domain. We removed the base stimulation frequency (6 Hz and harmonics) from these data using notch-filtering, such that the resulting waveforms isolate the *differential* responses to faces relative to objects (Retter & Rossion, 2016). Importantly, the latency of all observed components should be interpreted taking into account the sinusoidal contrast modulation. The epochs here were cropped relative to face onset, corresponding to the start of the sine cycle at 0% contrast, where the face is invisible.

Taking 30% contrast as a reference point for when the faces became visible to observers, the “true” latencies of the current time-domain components are shifted ~30 ms earlier in time (i.e., a sinusoidal stimulation at 6 Hz with a 120 Hz screen refresh rate reaches 30% contrast around ~30 ms post stimulus onset; see Retter & Rossion, 2016 for comparison between sinewave and squarewave stimulation).

The waveform of the face-selective response contained multiple components (Figure 7) similar to those we have described in previous studies (Rossion et al., 2015; Jacques, Retter, & Rossion, 2016; Retter & Rossion, 2016). A first positivity (“P1-face”, peaking at ~170 ms) arose over medial and lateral occipital channels, followed by a large negativity (“N1-face”, ~245 ms) with a bilateral distribution. This negativity was prolonged when observers explicitly attended to faces (second peak at ~337 ms). Finally, a large positivity (“P2-face”, ~427 ms) was apparent over bilateral occipito-temporal channels, but more ventral to the first negativity.

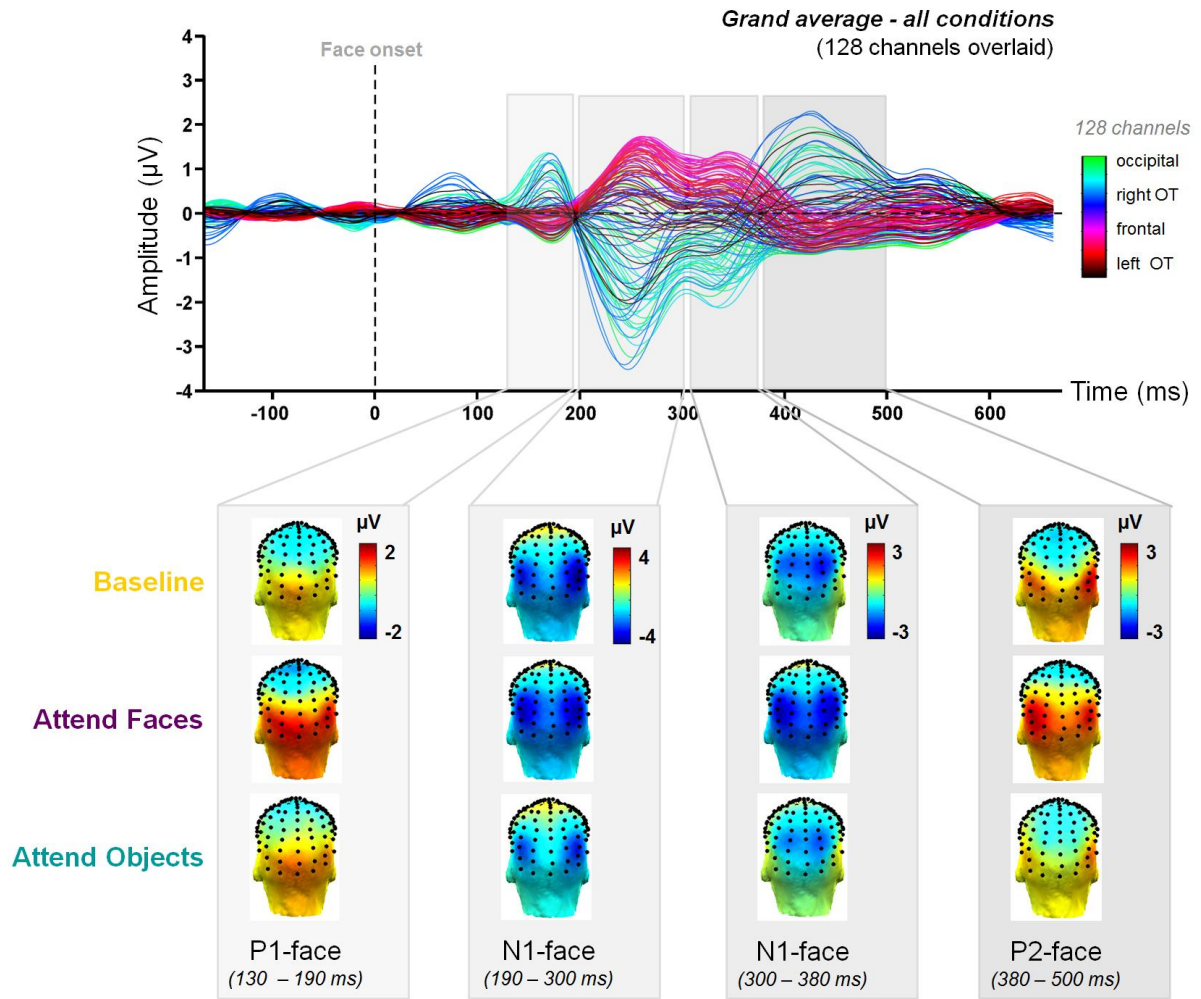


Figure 7. Mean time-domain waveform of the face-selective response, averaged across all conditions and with all 128 channels overlaid (colour legend at right). Below are the topographies of the peak amplitude (defined within the time-windows between parentheses) of the four main time-domain components for each condition. Note that the time of face onset here refers to the start of the sinewave cycle at 0% contrast, where faces are not yet visible. Hence, the “true” timing of the face-selective components should be shifted earlier by ~30 ms, corresponding to the time for faces to be shown at a visible 30% contrast.

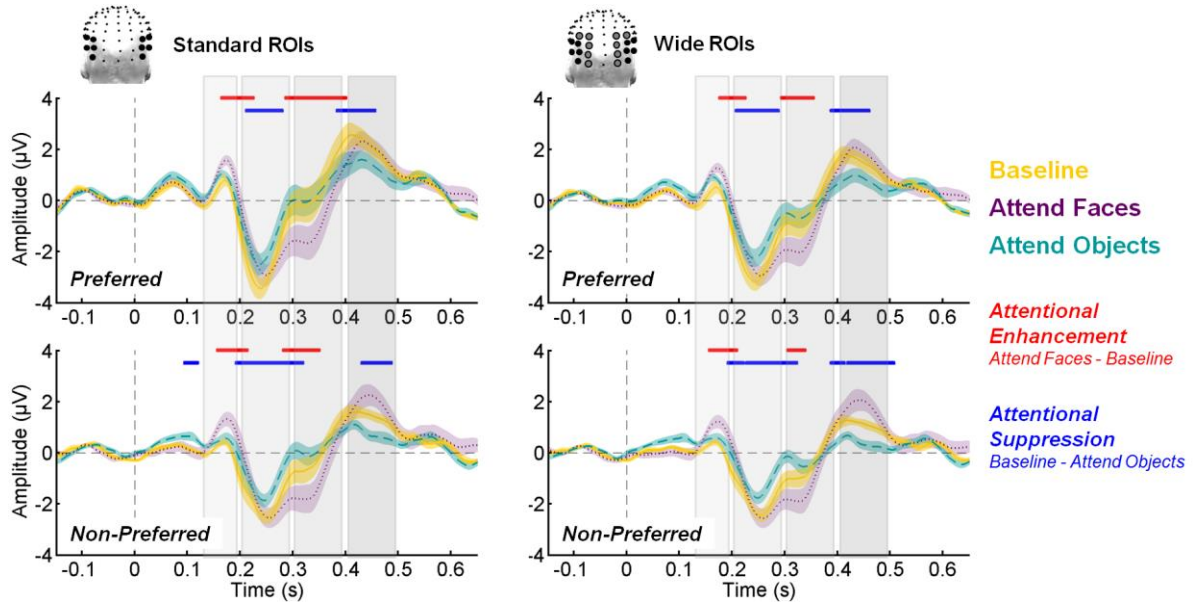


Figure 8. Mean face-selective responses shown as a function of Condition, Hemisphere, and ROI type. The shaded regions represent standard error of the mean. Grey shaded boxes delineate approximate time-windows of the components shown in Figure 7. The Standard ROIs refer to the bilateral ROIs defined in the frequency-domain analysis while the Wide ROIs contain additional channels over dorsal and medial occipital regions, based on the topographies of the negative time-domain components. Coloured lines above each plot represent the time-points during which we observed significant attentional modulation (red = enhancement; blue = suppression; see Methods).

To compute indices of attentional enhancement and suppression, we calculated conditional difference-waves (i.e., Attend Faces – Baseline; Baseline – Attend Objects) and their 99% confidence intervals. Time-points at which these intervals exclude zero during at least five consecutive bins suggest meaningful attentional effects (as indicated by coloured lines above the plots in Figure 8). We observed consistent differences between the time-course of attentional enhancement and suppression in both types of ROI analyses. Attentional enhancement was present during an early time-window (150-200 ms), corresponding to the “P1-Face” component and between 300-400 ms during the prolongation of the “N1-Face” component. By contrast, attentional suppression occurred between 200-300 ms during the “N1-Face” component and later at 400-500 ms during the “P2-Face” time-window. Hence,

these two opposing attentional modulations unfolded over largely non-overlapping time-windows in both the Preferred and Non-Preferred hemispheres.

One aspect of these analyses that may be surprising at first is the apparent discrepancy between the frequency-domain and time-domain analyses in terms of attentional enhancement in the Preferred hemisphere. Specifically, where enhancement in this hemisphere *was* evident in the time-domain difference waveforms (i.e., Attend Faces – Baseline, see Figure 8), there was no evidence of enhancement in the frequency-domain quantification (see Figure 5B). This divergence is underpinned by the nature of the response quantification in each analysis. According to the Fourier theory, any signal in time can be reconstructed with a sum of sinusoids at different frequencies, amplitudes, and phases. For our frequency-domain analysis, responses were quantified as the sum of amplitudes, discarding the phase information in the process. In contrast, phase information is preserved in the time-domain analysis. Two response waveforms can be described by the same frequencies at the same amplitudes, but their respective phases will determine the *shape* of their waveforms and therefore any potential differences in temporal dynamics (Figure 9). Hence, as illustrated in Figure 10, while the amplitude of each harmonic of the face presentation frequency was similar between the Baseline and Attend Faces conditions in the Preferred Hemisphere, there was a large ($\pm 20^\circ$) phase-shift on the first three harmonics between these two conditions (see Table 2). The observed amplitude enhancement of the time-domain response in the Attend Faces condition is likely driven by this phase-shift. Conversely, in the Non-Preferred hemisphere, there was both a phase shift *and* an amplitude difference between the Baseline and Attend Faces conditions, making it easier to detect a difference regardless of whether phase information is taken into account (i.e., in both the frequency- and time-domains). Interestingly, we note that attending to a category other than faces appears to reduce the face-selective response amplitude, but does not modify its phase

(i.e., the shape of response waveforms are comparable). This again suggests two different processes underlying attending towards faces and attending away from faces. Overall, it appears that explicitly attending to faces *does* impact face-selective processing in the Preferred hemisphere, albeit at a much more subtle level that is evident only when the temporal dynamics of attentional modulation are taken into account.

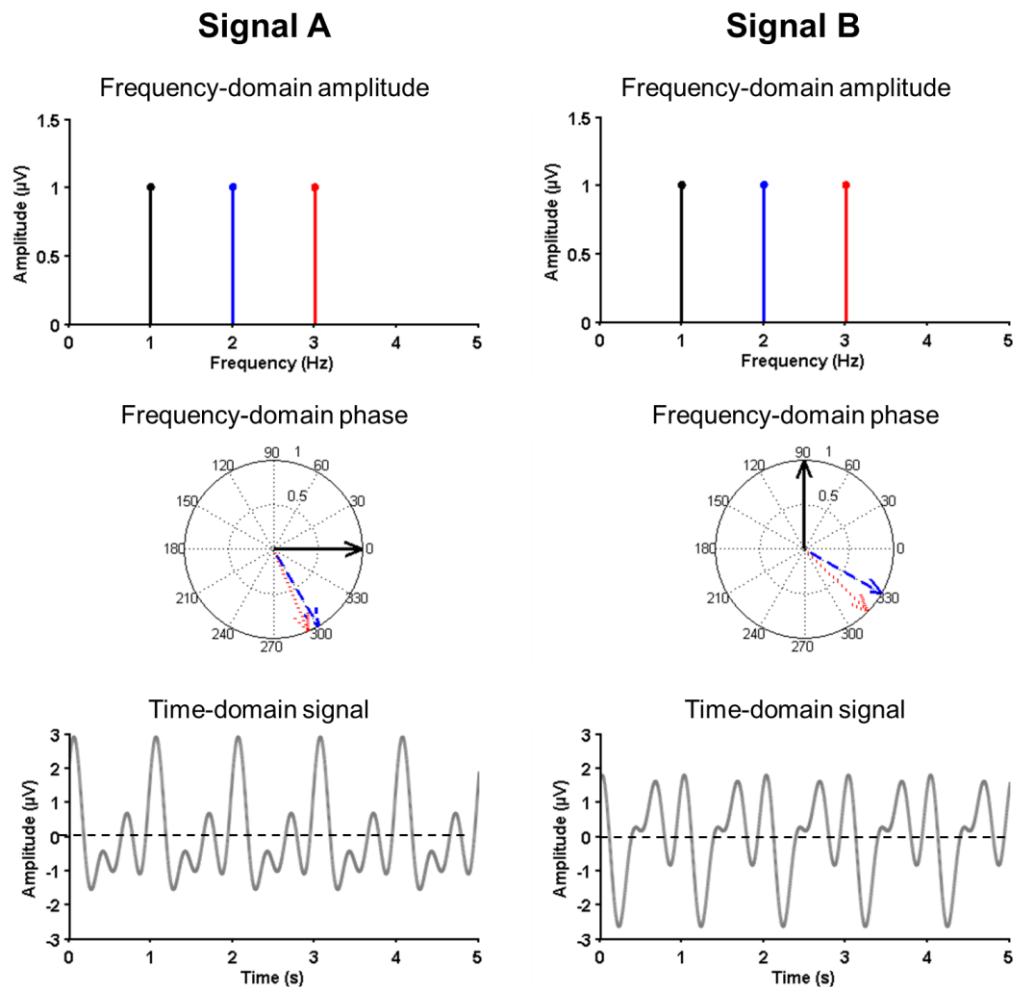


Figure 9. Illustration of the relationship between the time-domain signal and its frequency-domain amplitude and phase components. Signals A and B are both composed of three sinusoids at 1 Hz, 2 Hz, and 3 Hz. The magnitude of these harmonic frequencies are equal, as shown by the identical amplitude spectra (first row). However, the phase of the harmonic frequencies differ between the signals (second row), leading to vastly different shapes of the waveforms (third row). Hence, while the overall magnitudes of signals A and B are the same, their temporal dynamics do differ.

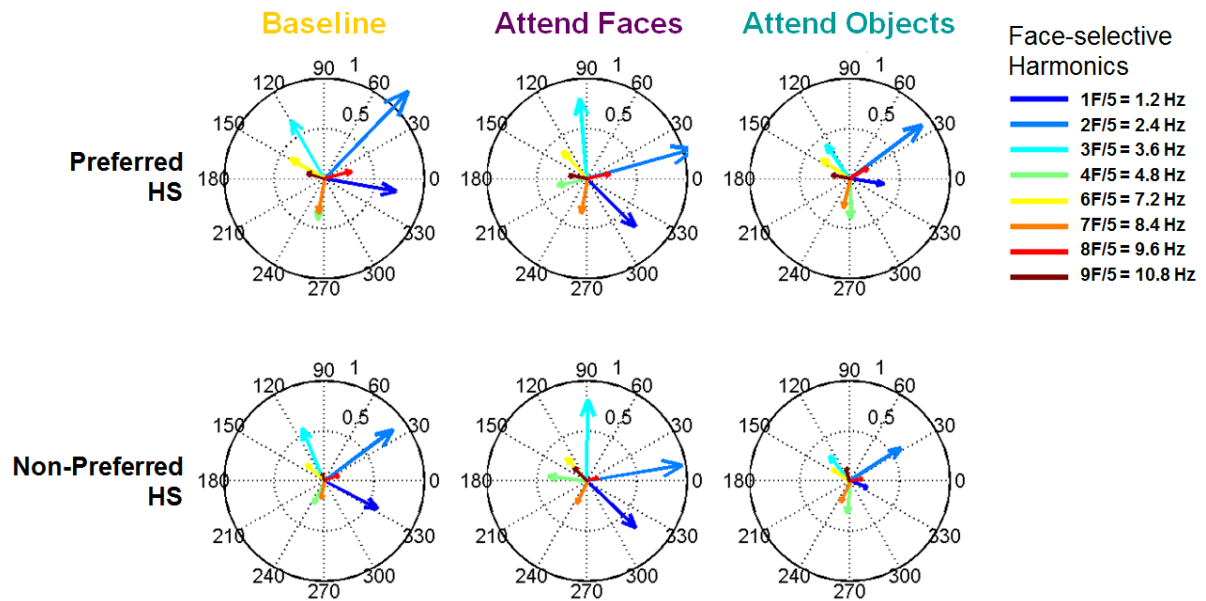


Figure 10. Mean phase values of harmonics of the face presentation frequency used for frequency-domain quantification. Here circumference reflects phase in degrees and arrow length reflects frequency amplitude. On average, a shift in the phase of the first three harmonics was present in the Attend Faces condition relative to Baseline.

Table 2

Mean phase shift in degrees for the three first harmonics of the face-selective response

	Preferred hemisphere		Non-Preferred hemisphere	
	Attentional Enhancement	Attentional Suppression	Attentional Enhancement	Attentional Suppression
1F/5 = 1.2 Hz	-20.44	-3.66	-17.53	-3.29
2F/5 = 2.4 Hz	-28.21	8.90	-22.37	4.37
3F/5 = 3.6 Hz	-19.97	-9.14	-22.55	-13.87
Mean phase shift	-22.87	-1.31	-20.82	-4.26

General Discussion

Efficient perceptual categorisation in daily life occurs in dynamic and highly complex visual environments. Yet the role of selective attention in guiding meaningful categorisation has predominantly been studied under sparse and static viewing conditions. Here we asked how task-based attention modulates face categorisation which is characterised by the same temporal and complexity constraints typical of effective perceptual categorisation in real world vision. We provide the first evidence that attentional enhancement and suppression exert a differential impact on face processing supported by the left and right hemispheres. Relative to an attentional baseline, actively attending to faces enhances the face-selective neural response much more evidently in the left hemisphere than in the right; whereas attending to a stimulus category other than faces suppresses the face-selective response in both hemispheres to an equal extent.

Task-based Attentional Modulation of Face Categorisation

Hemispheric differences in attentional sensitivity

Our results in the frequency-domain indicate that categorisation of highly variable face images in a rapid, dynamic visual stream is efficient and robust, unfolding in under 167 ms over predominantly right occipito-temporal regions. Although a perceptual discrimination response for faces vs. objects was evident under all conditions of task-based attention, this response was nevertheless still sensitive to attentional modulation. Critically, however, this attentional effect differed between the two hemispheres. When faces and objects were *equally* task-irrelevant (i.e., in the orthogonal task condition), the face-selective response was ~20-30% stronger in the right ROI compared to the left ROI. This right hemispheric dominance is consistent with previous studies that have shown task-irrelevant faces in dynamic visual streams preferentially engage the right occipito-temporal region (de Heering & Rossion, 2015; Jacques et al., 2016; Retter & Rossion, 2016; Rossion et al., 2015). Actively attending

to faces in the sequence enhanced the face-selective response much more prominently in the left (Non-Preferred) hemisphere (43% increase) than in the right (Preferred) hemisphere (3%). In contrast, selectively attending to objects (i.e., ignoring faces) produced an attentional suppression effect that was largely comparable in the two hemispheres (~24% reduction). Taken together, these frequency-domain results suggest that selective attention exerts a differential influence on face categorisation processes supported by the left and right face perception networks. These findings are in line with our prediction that face processing supported by the right occipito-temporal region – the functional core of the face perception network (de Heering & Rossion, 2015; Jonas & Rossion, 2016; Michel, Poncet, & Signoret, 1989; Rossion, Prieto, et al., 2012; Sergent et al., 1992; Sergent & Signoret, 1992; Zhen et al., 2015) – should be comparatively robust to attentional enhancement. To the best of our knowledge, the present study provides the first empirical demonstration of a clear hemispheric difference in attentional sensitivity in face categorisation.

Distinct Time-Courses of Attentional Enhancement and Suppression

An interesting and unexpected aspect of our results concerns the differing time-course of attentional enhancement and suppression of face-selective processing. In the time-domain, attentional enhancement was reflected in an increased amplitude for the P1-face (100-200 ms), and in a prolongation of the N1-face (300-400 ms). In contrast, attentional suppression was evident in an amplitude decrease for the N1-face and P2-face (400-500 ms; see Figure 8). To our knowledge, an alternating influence of enhancement and suppression during the unfolding of the face-selective response has not been reported before. Thus far, enhancement and suppression of face-related activity has been documented only in N1 (120-220ms) latency shifts and P1 (50-150ms) amplitude differences (Gazzaley et al., 2008; Gazzaley et al., 2005; Zanto et al., 2010). While this intriguing finding should be interpreted conservatively until it can be replicated, it is nevertheless interesting to consider the

theoretical implications of the pattern we report here. Specifically, if a particular face-selective component is consistently robust to attentional enhancement, does this suggest that this aspect of face processing is automatically engaged to its saturation point? Inversely, if some components cannot be suppressed by the focus of task-based attention, what does this say about the compulsory/automatic nature of the face processing carried by these components?

While at first glance the presence of attentional enhancement in the right (Preferred) hemisphere in the time-domain might appear to conflict with the absence of such an effect in the frequency-domain, results from these two analyses are in fact entirely complementary and reveal different aspects of the same response. Periodic visual presentation of face images at a given frequency elicits a periodic face-selective response in the brain at this same frequency. This face-selective response is a complex waveform comprised of multiple components (see Retter & Rossion, 2016). Frequency-domain analysis allows us to easily identify and quantify the overall magnitude of the face-selective responses and compare them between conditions, with the drawback of not knowing how exactly waveforms in the two conditions differ. Conversely, time-domain analysis provides exactly this information, giving a detailed picture of how responses unfold, but with less power to detect differences that are spread across time. Consequently, it is entirely possible for an effect to be present in one type of analysis and absent in the other (or vice versa). In our case, the combined information from the frequency- and time-domains suggests that attentional enhancement is stronger overall for the left hemisphere, where it both modulates the temporal dynamics of the response as well as its amplitude, whereas it only impacts response shape in the right hemisphere.

On balance, both the frequency- and time-domain results suggest a differential sensitivity to attention in the left and right face perception networks, a finding which represents a critical step forward in our understanding of the role of selective attention in face

categorisation. Moreover, the data here point to the interesting possibility that, at least for face categorisation in the right hemisphere, the *global* attentional effect that has been so frequently reported in the literature is predominantly underpinned by suppression, rather than enhancement.

Advantages of the Current Design

The present design has several important advantages over existing paradigms that enabled us to observe hemispheric differences in attentional sensitivity where others have not. First, and most importantly, our design here employed an attentional baseline condition which directly enabled the separate quantification of attentional enhancement and suppression, an approach that remains surprisingly rare in studies of high-level object perception (Chadick & Gazzaley, 2011; Gazzaley et al., 2008; Gazzaley et al., 2005; Zanto et al., 2010). Had we taken the conventional approach of simply contrasting the face-selective response elicited under maximally and minimally attended conditions (e.g., "Attend Face" vs. "Attend House", Baldauf & Desimone, 2014; Downing et al., 2001; Engell & McCarthy, 2010; Holmes et al., 2003; Lueschow et al., 2004; O'Craven et al., 1999; Sreenivasan, Goldstein, Lustig, Rivas, & Jha, 2009; Vuilleumier et al., 2001; Williams et al., 2005; Wojciulik et al., 1998; Yi et al., 2006), we would have entirely missed this interesting finding. Moreover, the attentional baseline used here was an orthogonal task which actively constrained observers' task-based attention, potentially providing a more stable attentional baseline against which to observe enhancement and suppression effects. This is in contrast to a handful of studies which have used a 'passive-viewing' attentional baseline condition, in which participants are at liberty to, deliberately or otherwise, preferentially direct their attention to one stimulus category over the other (cf. Chadick & Gazzaley, 2011; Gazzaley et al., 2008; Gazzaley et al., 2005; Zanto et al., 2010). Given that faces are thought to capture attentional resources more than other stimulus categories (for a discussion, see Palermo &

Rhodes, 2007), participants in these studies may well have attended preferentially to the faces presented during passive-viewing.

Second, the paradigm here indexes face categorisation imposes the twin constraints that characterise effective perceptual categorisation in the real world, i.e., speed and high categorical diversity (Crouzet et al., 2010; Crouzet & Thorpe, 2011; Rousselet et al., 2003). Our dynamic rapid display (e.g., over 360 images per minute) places considerable processing strain on the visual system, which may have helped to pull face processing “off the ceiling”, making it easier in turn to detect attentional benefits (Lavie, 2005). Importantly, this paradigm also minimises the contribution of low-level differences to perceptual categorisation, enabling us to target attentional modulation of high-level face categorisation processes (Gao, Gentile, & Rossion, 2017; Rossion et al., 2015). By presenting observers with a large number of unsegmented images that vary widely in composition, lighting, viewing angle, etc., the face-selective response necessarily reflects both successful *discrimination* of faces from the many other object types, as well as successful *generalisation* across multiple varied face exemplars. In this way, we are able to objectively quantify attentional modulation of high level face categorisation processes (projected to 1.2 Hz and harmonics) *in isolation* from more general visual processing common to both faces and objects (projected to 6 Hz and harmonics). Attentional modulations may be more readily detected in the context of such a truly face-selective response, rather than activity that is simply face-related (e.g., responses elicited by faces), as is common in studies using standard EEG or fMRI approaches.

Given how the experimental framework of the current study differs from previous studies, it is important to consider whether our main results and conclusions may be explained by solely design-related factors. One concern is that the observed patterns of hemispheric lateralisation are driven by the stimulation method itself, i.e., by presenting a fast

stream of natural images. However, there is strong evidence against the possibility that fast periodic presentation of *any* image type should give rise to a right lateralised response profile. Indeed, not only is the right lateralisation of the periodic responses to faces in the baseline condition consistent with the well-documented specialisation of this hemisphere for face processing (see above), but we also have evidence that such periodic responses, and their lateralisation, inherently represent functionally selective perceptual processing. More precisely, the periodic response to faces at 1.2 Hz can only arise from the detection of face images among object images, and therefore reflects selective visual processing. Other visual category contrasts measured with this approach elicit vastly different response topographies. For instance, presenting words among letter strings (Lochy, Van Belle, & Rossion, 2015), elicits a left-lateralised periodic response at the word frequency that is consistent with the recruitment of the left hemisphere for specialised language processing. In another example, responses to faces, houses, or body-parts among a stream of other objects leads to distinct response topographies despite identical presentation frequencies, with only faces leading to a significant right lateralisation (Jacques et al., 2016). Hence, in this study, we have good reason to assert that the periodic face-selective responses in each condition directly relate to how the specialised face network processes the face stimuli in each case. In other words, the response lateralisations are not spurious but functionally relevant.

A separate concern is whether the hemisphere differences we observed might be an artefact of temporal attention, i.e., participants attending to the periodic *frequency* of stimulation rather than its contents. Regarding this, we have recently demonstrated that temporal expectation does not influence the face-selective response (Quek & Rossion, 2017). However, if we consider that participants were indeed exploiting periodicity to complete the Attend Face and Attend Object tasks, the image presentation rate (6 Hz) would in fact be a more salient temporal cue than the embedded face presentation rate. In this case, both

attentional conditions would be equally aided by periodicity, since participants would use the 6 Hz “beat” to focus attention. Alternatively, had participants been able to attend to the frequency specific to each task (e.g. 1.2 Hz for attending to faces and 6 Hz for attending to objects), there should have been similar attentional enhancement of the left hemisphere for both frequencies. In other words, we should have observed an increase of the face-selective response over the left hemisphere when participants were attending to faces *and* an increase of the common response when participants were attending to objects. However, there was no such response profile on the common response. Given that the ultimate goal here was to investigate task-based attention, we would argue that the current findings remain relevant on how attention differentially modulates face processing across hemispheres, regardless of the exact mechanism by which attention was selectively engaged to the task-relevant stimuli.

A final question is whether directing participants’ attention specifically to face gender could have somehow driven the pronounced attentional enhancement effect over the left hemisphere. Yet such a task-specific effect would seem unlikely given that the existing literature does not suggest face gender itself is processed in the left hemisphere (Sergent & Corballis, 1989; Sergent et al., 1992; Wiese, Schweinberger, & Neumann, 2008). Note that although some reports suggest an interaction between response lateralisation and participant gender (e.g., Lovén, Svärd, Ebner, Herlitz, & Fischer, 2013) our pattern of results was consistent across individual participants regardless of their gender (see Figure 4). Still another possibility might be that the gender task we employ depends on local feature processing (Brown & Perrett, 1993; Dupuis-Roy, Fortin, Fiset, & Gosselin, 2009; Yamaguchi, Hirukawa, & Kanazawa, 2013), and that this local processing drives the left hemisphere recruitment in the Attend Faces condition (Bourne, Vladeanu, & Hole, 2009; Hillger & Koenig, 1991; Parkin & Williamson, 1987). However, several important factors undermine this argument. First and most importantly, our face stimuli were greyscale and

highly variable in their lighting, pose, size, position, and external facial features (e.g., ears, hair, accessories...). In the absence of colour cues or systematic overlap of features, local details would not be efficient diagnostic cues for gender. For example, participants cannot reliably monitor the mouth of each face, since the position of the mouth changes across each face presentation. Moreover, the short image presentation duration ($SOA = 167$ ms) prevented participants making multiple saccades across the faces to inspect individual features. As such, regardless of exactly how participants completed the face gender discrimination task (which they did with high accuracy), it is highly unlikely they relied on local processing to do so. In summary, we believe that the current findings are neither artefacts of the experimental design nor the specific task used, but rather reflect functionally relevant hemispheric differences in face categorisation.

Conclusion & Future Research

Selective attention guides behaviour in dynamic and complex visual environments, yet its role in face categorisation has not yet been examined under conditions that enforce the strong processing constraints which characterise effective perceptual categorisation in the real world. Using an original dynamic visual stimulation approach, we uncovered the hitherto unknown finding that selective attention influences face categorisation in the left and right face perception networks differently. Where the right hemisphere is mandatorily activated by faces and benefits little from the allocation of attention, the left hemisphere appears to be flexibly recruited to serve current task demands. Additionally, we show that attentional enhancement and suppression occur over distinct time-windows during the face-selective response. An outstanding question is whether this pattern of differential attentional sensitivity across hemispheres extends to other high-level object categories – for instance, the opposite pattern (i.e., greater attentional enhancement of right hemisphere responses compared to left) might be predicted for word stimuli, the processing of which is left-lateralised. Similarly, for

a visual category with bilateral responses, such as objects, we might expect the attentional enhancement effect to be equally distributed across hemispheres.

Acknowledgements

The authors thank Talia Retter for assistance with data acquisition, and three anonymous reviewers for their helpful comments on a previous version of this paper. This work was supported by a co-funded initiative by the University of Louvain and the Marie Curie Actions of the European Commission awarded to GLQ [grant no: F211800012], an FRSM-FNRS grant awarded to DN [grant no: 3.4601.12], a European Research Council grant awarded to BR [grant no: facessvep 284025], and an FSR-FNRS postdoctoral grant awarded to JLS [grant no: FC 91608].

References

- Baldauf, D., & Desimone, R. (2014). Neural mechanisms of object-based attention. *Science*, 344(6182), 424-427. doi:10.1126/science.1247003
- Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *Journal of Statistical Software*, 31(10), 1-21.
- Bourne, V. J., Vladeanu, M., & Hole, G. J. (2009). Lateralised repetition priming for featurally and configurally manipulated familiar faces: evidence for differentially lateralised processing mechanisms. *Laterality: Asymmetries of Body, Brain and Cognition*, 14(3), 287-299.
- Broadbent, D. E. (1958). *Perception and communication*. New York, NY: Oxford University Press.
- Brown, E., & Perrett, D. I. (1993). What gives a face its gender? *Perception*, 22(7), 829-840.
- Chadick, J. Z., & Gazzaley, A. (2011). Differential coupling of visual cortex with default or frontal-parietal network based on goals. *Nat Neurosci*, 14(7), 830-832. doi:<http://www.nature.com/neuro/journal/v14/n7/abs/nn.2823.html#supplementary-information>
- Chen, Y., Seth, A. K., Gally, J. A., & Edelman, G. M. (2003). The power of human brain magnetoencephalographic signals can be modulated up or down by changes in an attentive visual task. *Proceedings of the National Academy of Sciences*, 100(6), 3501-3506. doi:10.1073/pnas.0337630100
- Cohen, M. R., & Maunsell, J. H. R. (2011). Using neuronal populations to study the mechanisms underlying spatial and feature attention. *Neuron*, 70(6), 1192-1204. doi:10.1016/j.neuron.2011.04.029
- Crouzet, S. M., Kirchner, H., & Thorpe, S. J. (2010). Fast saccades toward faces: Face detection in just 100 ms. *Journal of Vision*, 10(4), 16-16. doi:10.1167/10.4.16
- Crouzet, S. M., & Thorpe, S. J. (2011). Low-level cues and ultra-fast face detection. *Front Psychol*, 2, 342. doi:10.3389/fpsyg.2011.00342
- de Heering, A., & Rossion, B. (2015). Rapid categorization of natural face images in the infant right hemisphere. *eLife*, 4, e06564.
- Downing, P., Liu, J., & Kanwisher, N. (2001). Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia*, 39, 1329-1342.
- Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*, 9(2), 10-10.
- Eimer, M. (2000a). Attentional modulations of event-related brain potentials sensitive to faces. *Cognitive Neuropsychology*, 17, 103-116.
- Engell, A. D., & McCarthy, G. (2010). Selective attention modulates face-specific induced gamma oscillations recorded from ventral occipitotemporal cortex. *Journal of Neuroscience*, 30(26), 8780-8786. doi:10.1523/JNEUROSCI.1575-10.2010
- Frassle, S., Paulus, F. M., Krach, S., Schweinberger, S. R., Stephan, K. E., & Jansen, A. (2016). Mechanisms of hemispheric lateralization: Asymmetric interhemispheric recruitment in the face perception network. *Neuroimage*, 124(Pt A), 977-988. doi:10.1016/j.neuroimage.2015.09.055
- Furey, M. L., Tanskanen, T., Beauchamp, M. S., Avikainen, S., Uutela, K., Hari, R., & Haxby, J. V. (2006). Dissociation of face-selective cortical responses by attention. *Proceedings of the National Academy of Sciences of the United States of America*, 103(4), 1065-1070.
- Gao, X., Gentile, F., & Rossion, B. (2017). Fast Periodic Stimulation (FPS): A Highly Effective Approach In fMRI Brain Mapping. *bioRxiv*. doi:10.1101/135087
- Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., & D'Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory

- processing. *Proc Natl Acad Sci U S A*, 105(35), 13122-13126.
doi:10.1073/pnas.0806074105
- Gazzaley, A., Cooney, J. W., Rissman, J., & D'Esposito, M. (2005). Top-down suppression deficit underlies working memory impairment in normal aging. *Nature neuroscience*, 8(10), 1298-1300.
- Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *J. Neurosci.*, 14(11), 6336-6353.
- Hillger, L. A., & Koenig, O. (1991). Separable Mechanisms in Face Processing: Evidence from Hemispheric Specialization. *Journal of Cognitive Neuroscience*, 3(1), 42-58.
doi:10.1162/jocn.1991.3.1.42
- Ho, T. C., Brown, S., Abuyo, N. A., Ku, E.-H. J., & Serences, J. T. (2012). Perceptual consequences of feature-based attentional enhancement and suppression. *Journal of Vision*, 12(8), 15-15. doi:10.1167/12.8.15
- Holmes, A., Kiss, M., & Eimer, M. (2006). Attention modulates the processing of emotional expression triggered by foveal faces. *Neurosci Lett*, 394(1), 48-52.
- Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research*, 16(2), 174-184. doi:10.1016/s0926-6410(02)00268-9
- Jacques, C., Retter, T. L., & Rossion, B. (2016). A single glance at a face generates larger and qualitatively different category-selective spatio-temporal signatures than other ecologically-relevant categories in the human brain. *NeuroImage*, 137, 21-33.
- Jonas, J., & Rossion, B. (2016). Beyond the core face-processing network: intracerebral stimulation of a face-selective area in the right anterior fusiform gyrus elicits transient prosopagnosia. *Journal of Vision*, 16(12), 385-385.
- Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin Neurophysiol*, 111(10), 1745-1758.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *The Journal of Neuroscience*, 17(11), 4302-4311.
- Kastner, S., & Ungerleider, L. G. (2000). Mechanisms of visual attention in the human cortex. *Annual review of neuroscience*, 23(1), 315-341.
- Keil, A., Moratti, S., Sabatinelli, D., Bradley, M. M., & Lang, P. J. (2005). Additive Effects of Emotional Content and Spatial Selective Attention on Electrocutaneous Facilitation. *Cerebral Cortex*, 15(8), 1187-1197. doi:10.1093/cercor/bhi001
- Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences*, 9(2), 75-82. doi:<http://doi.org/10.1016/j.tics.2004.12.004>
- Liu-Shuang, J., Norcia, A. M., & Rossion, B. (2014). An objective index of individual face discrimination in the right occipito-temporal cortex by means of fast periodic oddball stimulation. *Neuropsychologia*, 52, 57-72.
doi:10.1016/j.neuropsychologia.2013.10.022
- Lochy, A., Van Belle, G., & Rossion, B. (2015). A robust index of lexical representation in the left occipito-temporal cortex as evidenced by EEG responses to fast periodic visual stimulation. *Neuropsychologia*, 66, 18-31.
- Lovén, J., Svärd, J., Ebner, N. C., Herlitz, A., & Fischer, H. (2013). Face gender modulates women's brain activity during face encoding. *Social cognitive and affective neuroscience*, 9(7), 1000-1005.

- Lueschow, A., Sander, T., Boehm, S. G., Nolte, G., Trahms, L., & Curio, G. (2004). Looking for faces: Attention modulates early occipitotemporal object processing. *Psychophysiology*, 41(3), 350-360. doi:10.1111/j.1469-8986.2004.00159.x
- Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Biol*, 14(9), 744-751. doi:10.1016/j.cub.2004.04.028
- Michel, F., Poncet, M., & Signoret, J. (1989). Les lésions responsables de la prosopagnosie sont-elles toujours bilatérales. *Revue de Neurologie*, 145, 764-770.
- Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proc Natl Acad Sci U S A*, 93(10), 4770-4774.
- Müller, M. M., Andersen, S., Trujillo, N. J., Valdés-Sosa, P., Malinowski, P., & Hillyard, S. A. (2006). Feature-selective attention enhances color signals in early visual areas of the human brain. *Proceedings of the National Academy of Sciences*, 103(38), 14250-14254. doi:10.1073/pnas.0606668103
- Müller, M. M., & Hübner, R. (2002). Can the spotlight of attention be shaped like a doughnut? Evidence from steady-state visual evoked potentials. *Psychological Science*, 13(2), 119-124.
- Müller, M. M., Picton, T. W., Valdes-Sosa, P., Riera, J., Teder-Sälejärvi, W. A., & Hillyard, S. A. (1998). Effects of spatial selective attention on the steady-state visual evoked potential in the 20–28 Hz range. *Cognitive Brain Research*, 6(4), 249-261.
- Müsch, K., Hamamé, C. M., Perrone-Bertolotti, M., Minotti, L., Kahane, P., Engel, A. K., . . . Schneider, T. R. (2014). Selective attention modulates high-frequency activity in the face-processing network. *Cortex*, 60, 34-51. doi:<http://dx.doi.org/10.1016/j.cortex.2014.06.006>
- Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottareau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research: A review. *Journal of Vision*, 15(6), 4-4. doi:10.1167/15.6.4
- O'Craven, K., Downing, P., & Kanwisher, N. (1999). fMRI evidence for objects as the units of attentional selection. *Nature*, 401, 584-587.
- Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, 45, 75-92.
- Parkin, A. J., & Williamson, P. (1987). Cerebral lateralisation at different stages of facial processing. *Cortex*, 23(1), 99-110.
- Pessoa, L., McKenna, M., Gutierrez, E., & Ungerleider, L. G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences*, 99(17), 11458-11463.
- Quek, G. L., & Finkbeiner, M. (2014). Face-sex categorization is better above fixation than below: Evidence from the reach-to-touch paradigm. *Cognitive, Affective, & Behavioral Neuroscience*, 14(4), 1407-1409. doi:10.3758/s13415-014-0282-y
- Quek, G. L., & Finkbeiner, M. (2016). The upper-hemifield advantage for masked face processing: Not just an attentional bias. *Attention, Perception, & Psychophysics*, 78(1), 52-68. doi:10.3758/s13414-015-0965-7
- Quek, G. L., & Rossion, B. (2017). Category-selective human brain processes elicited in fast periodic visual stimulation streams are immune to temporal predictability. *Neuropsychologia*. doi:<http://dx.doi.org/10.1016/j.neuropsychologia.2017.08.010>
- Regan, D. (1989). *Human brain electrophysiology: evoked potentials and evoked magnetic fields in science and medicine*. Amsterdam: Elsevier.

- Retter, T. L., & Rossion, B. (2016). Uncovering the neural magnitude and spatio-temporal dynamics of natural image categorization in a fast visual stream. *Neuropsychologia*, 91, 9-28. doi:10.1016/j.neuropsychologia.2016.07.028
- Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Defining face perception areas in the human brain: a large-scale factorial fMRI face localizer analysis. *Brain Cogn*, 79(2), 138-157. doi:10.1016/j.bandc.2012.01.001
- Rossion, B., Prieto, E. A., 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. doi:10.1016/j.neuroimage.2012.08.033
- Rossion, B., Torfs, K., Jacques, C., & Liu-Shuang, J. (2015). Fast periodic presentation of natural images reveals a robust face-selective electrophysiological response in the human brain. *J Vis*, 15(1), 15.11.18. doi:10.1167/15.1.18
- Rousselet, G. A., Mace, M. J., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *J Vis*, 3(6), 440-455. doi:10.1167/3.6.5
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84(1), 1-66.
- Sergent, J., & Corballis, M. C. (1989). Categorization of disoriented faces in the cerebral hemispheres of normal and commissurotomy subjects. *Journal of Experimental Psychology: Human Perception and Performance*, 15(4), 701.
- Sergent, J., Ohta, S., & Macdonald, B. (1992). Functional neuroanatomy of face and object processing A positron emission tomography study. *Brain*, 115(1), 15-36.
- Sergent, J., & Signoret, J.-L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex*, 2(5), 375-388.
- Sreenivasan, K. K., Goldstein, J. M., Lustig, A. G., Rivas, L. R., & Jha, A. P. (2009). Attention to faces modulates early face processing during low but not high face discriminability. *Attention, Perception, & Psychophysics*, 71(4), 837-846.
- Srinivasan, R., Russell, D. P., Edelman, G. M., & Tononi, G. (1999). Increased synchronization of neuromagnetic responses during conscious perception. *The Journal of Neuroscience*, 19(13), 5435-5448.
- Treue, S., & Trujillo, J. C. M. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399(6736), 575-579.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2001). Effects of attention and emotion on face processing in the human brain: An event-related fMRI study. *Neuron*, 30(3), 829-841.
- Wang, J., Clementz, B. A., & Keil, A. (2007). The neural correlates of feature-based selective attention when viewing spatially and temporally overlapping images. *Neuropsychologia*, 45(7), 1393-1399. doi:<http://dx.doi.org/10.1016/j.neuropsychologia.2006.10.019>
- Wiese, H., Schweinberger, S. R., & Neumann, M. F. (2008). Perceiving age and gender in unfamiliar faces: Brain potential evidence for implicit and explicit person categorization. *Psychophysiology*, 45(6), 957-969.
- Williams, M. A., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2005). Differential amygdala responses to happy and fearful facial expressions depend on selective attention. *Neuroimage*, 24(2), 417-425.
- Wojciulik, E., Kanwisher, N., & Driver, J. (1998). Covert visual attention modulates face-specific activity in the human fusiform gyrus: fMRI Study. *Journal of Neurophysiology*, 79(3), 1574-1578.

- Yamaguchi, M. K., Hirukawa, T., & Kanazawa, S. (2013). Judgment of gender through facial parts. *Perception*, 42(11), 1253-1265.
- Yi, D. J., Kelley, T. A., Marois, R., & Chun, M. M. (2006). Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain Research*, 1080(1), 53-62.
- Zanto, T. P., Hennigan, K., Östberg, M., Clapp, W. C., & Gazzaley, A. (2010). Predictive knowledge of stimulus relevance does not influence top-down suppression of irrelevant information in older adults. *Cortex; a journal devoted to the study of the nervous system and behavior*, 46(4), 564-574. doi:10.1016/j.cortex.2009.08.003
- Zhen, Z., Yang, Z., Huang, L., Kong, X.-z., Wang, X., Dang, X., . . . Liu, J. (2015). Quantifying interindividual variability and asymmetry of face-selective regions: a probabilistic functional atlas. *Neuroimage*, 113, 13-25.