Vision Research 51 (2011) 1297–1311

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

ERP evidence for the speed of face categorization in the human brain: Disentangling the contribution of low-level visual cues from face perception

Bruno Rossion *, Stéphanie Caharel

Institute of Psychology, Institute of Neuroscience, Université Catholique de Louvain, Belgium

A R T I C L E I N F O

Article history:
Received 11 August 2010
Received in revised form 15 March 2011
Available online 28 April 2011

Keywords:
ERP
N170
P1
Face perception
Object processing
Low and high vision
Amplitude spectrum
Low-level visual cues

A B S T R A C T

How fast are visual stimuli categorized as faces by the human brain? Because of their high temporal resolution and the possibility to record simultaneously from the whole brain, electromagnetic scalp measurements should be the ideal method to clarify this issue. However, this question remains debated, with studies reporting face-sensitive responses varying from 50 ms to 200 ms following stimulus onset. Here we disentangle the contribution of the information associated with the phenomenological experience of a face (phase) from low-level visual cues (amplitude spectrum, color) in accounting for early face-sensitivity in the human brain. Pictures of faces and of a category of familiar objects (cars), as well as their phase-scrambled versions, were presented to fifteen human participants tested with high-density (128 channels) EEG. We replicated an early face-sensitivity – larger response to pictures of faces than cars – at the level of the occipital event-related potential (ERP) P1 (80–100 ms). However, a similar larger P1 to phase-scrambled faces than phase-scrambled cars was also found. In contrast, the occipito-temporal N170 was much larger in amplitude for pictures of intact faces than cars, especially in the right hemisphere, while the small N170 elicited by phase-scrambled stimuli did not differ for faces and cars. These findings show that sensitivity to faces on the visual evoked potentials P1 and N1 (N170) is functionally dissociated: the P1 face-sensitivity is driven by low-level visual cues while the N1 (or N170) face-sensitivity reflects the perception of a face. Altogether, these observations indicate that the earliest access to a high-level face representation, that is, a face percept, does not precede the N170 onset in the human brain. Furthermore, they allow resolving apparent discrepancies between the timing of rapid human saccades towards faces and the early activation of high-level facial representations as shown by electrophysiological studies in the primate brain. More generally, they put strong constraints on the interpretation of early (before 100 ms) face-sensitive effects in the human brain.

1. Introduction

How fast can a visual stimulus be categorized as a face by the human brain? This question is of interest because measurements of processing speed in the visual system can constrain models of its functional organization. For instance, human observers can release a button when a face is present in a centrally presented natural scene as fast as about 250–300 ms following stimulus onset (Rousselet, Mace, & Fabre-Thorpe, 2003). When two scenes are presented simultaneously in the left and right visual fields, observers are even capable of initiating a saccade towards the scene containing a face as fast as 100–110 ms after stimulus onset (Crouzet, Kirchner, & Thorpe, 2010; for original studies with animal categorization tasks, see Kirchner and Thorpe (2006) and Thorpe, Fize, and Marlot (1996)). Such short latencies between stimulus onset and correct behavioral responses suggest that these tasks are performed on the basis of a single feedforward sweep through the visual system (Thorpe & Fabre-Thorpe, 2002). Moreover, information might not even be able to reach high-level visual areas in such a short time scale, suggesting that extremely fast perceptual decision about the presence of a face in a visual scene may rely on evidence from low-level visual cues. For instance, it has been shown that saccades towards faces can be based merely on Fourier amplitude spectrum (AS) of the images rather than on information that is associated with the phenomenological experience of seeing a face in the stimulus (a high level visual representation), that is, phase information (Honey, Kirchner, & VanRullen, 2008; although see Cerf, Harel, EinhÃuser, & Koch, 2008).

To date, these observations of extremely fast behavioral face detection have been difficult to reconcile with data obtained with the classical method used to infer the time-course of processing in the human brain at a global scale, namely event-related potentials (ERPs). As other visual stimuli, flashed faces elicit a sequence of ERPs – obtained by averaging segments of EEG time-locked to

---

* Corresponding author. Address: Institute of research in Psychology (IPSY), and Institute of Neuroscience (IoNS), Centre for Cognitive and Systems Neuroscience, Université Catholique de Louvain, 10 Place du Cardinal Mercier, 1348 Louvain-la-Neuve, Belgium. Fax: +32 10 47 37 74.

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

0042-6989/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2011.04.003

stimulus onset (Cobb & Dawson, 1960) – on the scalp, with the most prominent visual EPs taking place before 200 ms. These visual EPs have been described as (1) the C1, a component peaking at about 70 ms that is observed only for large stimuli at occipito-parietal sites of the midline and reverses polarity for stimuli presented in the upper/lower visual field, suggesting a major source in the primary visual cortex (Clark, Fan, & Hillyard, 1995; Di Russo, Martinez, Sereno, Pitzalis, & Hillyard, 2002; Jeffreys & Axford, 1972; but see Ales, Yates, & Norcia, 2010 for recent evidence that the C1 may arise from a combination of activation in striate and early extrastriate areas); (2) the P1 (called C2 by Jeffreys & Axford, 1972 and also referred to as P100 in some studies), a 100 ms peak component over central and lateral posterior electrode sites and thought to originate from dorsal and ventral extrastriate visual areas (Di Russo et al., 2002) and (3) a N1 component, or a complex of posterior negative components, peaking between 130 and 200 ms and whose distribution and response properties may vary substantially with the kind of visual material presented. Precisely, the visual N1 is particularly large in response to faces and most prominent on lateral occipito-temporal electrode sites.[1] When elicited by faces, this lateral N1 has been termed the N170 (Bentin, McCarthy, Perez, Puce, & Allison, 1996). Although magnetoencephalography (MEG) is sensitive to only a subset of the sources generating the EEG scalp components (i.e., tangential sources), MEG studies have also reported a sequence of similar visual responses or ERMf components, and in particular a M1/M100 and M170 in response to faces (e.g., Halgren, Raij, Marinkovic, Jousmaki, & Hari, 2000; Linkenkaer-Hansen et al., 1998; Liu, Harris, & Kanwisher, 2002; Okazaki, Abrahamyan, Stevens, & Ioannides, 2008; Sams, Hietanen, Hari, Ilmoniemi, & Lounasmaa, 1997; Watanabe, Kakigi, & Puce, 2003).

Considering this sequence of electrophysiological events reflecting visual processes, a question of interest is the following: along that sequence, what is the latency at which there is evidence that faces are processed differently from other visual stimuli, and can this time frame be related to behavioral markers of face detection?

While no modulation of the C1 elicited by faces as opposed to other visual stimuli has been reported to our knowledge, there have been reports of face repetition effects at extremely short latencies, sometimes as early as 45–80 ms (George, Jemel, Fiori, & Renault, 1997; Morel, Ponz, Mercier, Vuilleumier, & George, 2009; Mouchetant-Rostaing, Giard, Bentin, Aguera, & Pernier, 2000a; see also Seeck et al., 1997 for intracranial reports) or even at 30–60 ms (Braeutigam, Bailey, & Swithenby, 2001). However, such face repetition effects have been found at extremely early latencies, sometimes before or within the time-range of early primary visual cortex activation. Moreover, they differ between studies in terms of their scalp topography and timing, and several studies have found that these effects are not specific to faces or even to high-level visual stimuli (e.g., George et al., 1997; Mouchetant-Rostaing, Giard, Delpuech, Echallier, & Pernier, 2000b). Consequently, their significance with respect to the speed of face processing is doubtful.

In contrast, it is precisely because of its consistent larger response to faces than other visual stimuli that most studies have focused on the N170, a component whose positive counterpart on the central electrode sites – the vertex positive potential (VPP) – was the focus of early studies (Bötzel & Grusser, 1989; Jeffreys; 1989; see Joyce & Rossion, 2005). Like the VPP, the N170 (Bentin et al., 1996; Bötzel, Schulze, & Stodieck, 1995; George, Evans, Fiori, Davidoff, & Renault, 1996) has an average peak latency of about 160 ms over lateral occipital electrode sites in most studies, but its latency varies between 140–200 ms in individual brains. Its larger amplitude to faces than other visual stimuli may start as early as 120–130 ms (Bentin et al., 1996; Bötzel et al., 1995; Itier & Taylor, 2004a, 2004b; Rossion et al., 2000; Rousselet, Husk, Bennett, & Sekuler, 2008; for a review see Rossion & Jacques, 2008). Considering the faster visual responses recorded in the monkey than the human brain (e.g., Schroeder, Molholm, Lakatos, Ritter, & Foxe, 2004), this time-frame could well be compatible with the mean onset latency of neurons responding selectively to faces found in the monkey infero-temproal cortex (IT), i.e., about 90–100 ms (e.g., Kiani, Esteky, & Tanaka, 2005). That is, the mean onset latency of the N170 face effect – i.e., the larger response to faces than other visual categories – appears to correspond roughly to the time at which face-selective cells activity could be initiated in the human brain (Rossion & Jacques, 2008). Supporting this view, a large and typical N170 response is elicited by very different kinds of stimuli as long as they are readily perceived (i.e., interpreted) as faces by the visual system (i.e., photographs, schematic faces, line drawings, faces made by object parts, inverted faces, half-faces, isolated eyes, . . . see Rossion and Jacques (2011) for a review).

However, if one considers that roughly 20 ms are necessary to activate the brain stem structures involved in oculomotor control, this 120–130 ms onset time-frame for the N170 is still incompatible with the speed of the 100–110 ms fastest saccades towards faces recently recorded (Crouzet et al., 2010). This reasoning suggests that the earliest saccadic responses towards faces are based on low-level visual information characterizing faces rather than face perception *per se*. In this context, it is particularly interesting that the P1 component, peaking at about 90–100 ms, has also been reported as being larger to pictures of faces than objects in a number of studies (e.g., Eimer, 1998; Goffaux, Gauthier, & Rossion, 2003; Herrmann, Ehlis, Muehlberger, & Fallgatter, 2005; Itier & Taylor, 2004a, 2004b; M1 in MEG studies, e.g., Halgren et al., 2000; Liu et al., 2002; Okazaki et al., 2008). These P1 face effects, starting shortly before 100 ms, are much less consistent than the N170 face effects, and so far, to our knowledge, they have been observed only with face photographs. Nevertheless, P1 sensitivity to faces is found in a sufficiently large number of studies, and thus should not be ignored when addressing the question of the speed of face categorization in the human brain (Rossion & Jacques, 2008). Moreover, there is recent evidence that P1 amplitude is correlated with neural activation in a right hemisphere face-sensitive area of the lateral inferior occipital cortex as identified in fMRI (Sadeh, Podlipsky, Zhdanov, & Yovel, 2010), and that early (60–100 ms) TMS-induced disruption of activation in this area may impair face processing (Pitcher, Walsh, Yovel, & Duchaine, 2007). Finally, intracerebral recordings from epileptic patients undergoing investigation prior to surgery show that category-related information, in particular about faces vs. object categories can be present in the response of areas in the ventral visual pathways as early as 100 ms after stimulus onset (Liu, Agam, Madsen, & Kreiman, 2009).

Considering these observations altogether, it can be hypothesized that these early face-related activities, in particular the face-sensitivity observed on the human scalp, i.e., the P1 face-effect, would reflect low-level visual cues that differentiate faces from other objects, most notably amplitude spectrum and perhaps category-related color cues. In contrast, the typical larger N170 response to faces than to objects would be largely driven by the shape of the stimulus, that is, the information that is associated with perceptual awareness of a face (Rossion & Jacques, 2008, 2011).

Previous EEG or MEG studies support this view, but only indirectly. For instance, early M1 differences between full color photographs of faces and objects may disappear when schematic

---

[1] T5/T6, or P7/P8 in 64-channel systems in the 10–20 electrode convention, or lower channels when available (see Rossion & Jacques, 2008).

black-and-white faces are used (Halgren et al., 2000). Also, Rousselet, Husk, Bennett, & Sekuler, 2005; Rousselet et al., 2008) showed that electrophysiological responses to gray-level faces, houses and noise textures with identical amplitude spectra did not differ before the N170 time-window, suggesting that early P1/M1 differences as found in other studies with such stimuli may be due to AS and or color differences. In a following study, the same authors (Rousselet, Pernet, Bennett, & Sekuler, 2008) showed that ERPs to faces and phase-scrambled faces did not differ until about 127 ms following stimulus onset. In MEG, Tanskanen, Nasanen, Montez, Paallysaho, and Hari (2005) showed that contrary to the M170, the M1 component does not increase with the visibility of a face (see also Jemel et al., 2003), but is the largest when the face was covered by noise presented in the spatial frequency range (8–16 cycles/image) that is optimal for face perception. Thus, that study did not only show that the M1 amplitude was unrelated to face perception *per se*, but that it was positively related to low-level cues that characterize faces in particular (i.e., spatial frequencies in 8–16 cycles/image range). Altogether, these studies suggest that P1/M1 face-sensitivity is related to low-level visual cues, most notably amplitude spectrum (the energy – contrast – at each spatial frequency scale and orientation), while N170/M170 face-sensitivity would rather be related to the phase information, which is associated with the phenomenological experience of a face (a high-level representation).

However, to our knowledge, no study has shown a P1/M1 early difference between faces and objects that can be attributed to low-level visual cues, while at the same time showing that the next difference in time between faces and objects (N170/M170) is rather due to information associated with the phenomenological experience of a face. That is, no study has shown a dissociation between low-level visual cues and high-level visual processes (i.e., face perception) as accounting for P1 vs. N170 face effects respectively. Such observations would help clarifying the time-course of face categorization, indicating that both low-level and high-level information contribute to electrophysiological face-sensitive effects at different time-scales and help reconciling evidence from electrophysiological and behavioral studies of fast face detection.

Here, to test this dissociation, we performed a simple ERP experiment in which we contrasted the presentation of segmented pictures of faces and cars. The face and car stimuli were equalized for overall luminance, but not for amplitude spectra. They were also presented in color, which were left different between faces and cars (Fig. 1, Table 1). Hence, although overall luminance of the two sets of color stimuli was controlled (see stimulus information in methods Table 1), the stimuli were *not* normalized for low-level cues, unlike in the studies of Rousselet et al. (2005, 2008), Rousselet, Pernet, et al. (2008)) mentioned above. Again, this was done on purpose since we aimed at testing whether, to what extent and when, low-level cues such as amplitude spectrum and color contribute to P1 and N170 face-sensitivity effects.

In order to dissociate the contribution of low-level visual cues from face perception in generating early electrophysiological differences between faces and cars, we presented observers with the exact same stimuli, in their phase-scrambled version. This operation randomizes the phase information of the stimuli but preserves their overall power spectra (see e.g., Jacques & Rossion, 2004; Sadr & Sinha, 2004; VanRullen, 2006). Thus, we had a simple 2 × 2 design, crossing category (faces vs. cars) and shape (intact vs. scrambled). Based on the studies reviewed above, we had several predictions. First, faces should elicit a larger P1 than cars, in agreement with evidence reported in several previous studies. However, this effect should also be found for phase-scrambled stimuli: a larger P1 to phase-scrambled faces than phase-scrambled cars (novel prediction). That is, a main effect of category on the P1 amplitude, with no significant interaction between category and shape. Second, the N170 should be larger for faces than cars, also in agreement with the literature. However, contrary to the P1 this N170 effect should not be accounted for by low-level visual cues. That is, phase-scrambled faces should not lead to a larger N170 than phase-scrambled cars. Or, at the very least, there should be an interaction between category and shape so that the N170 difference between faces and cars would be larger for intact than phase-scrambled stimuli.

2. Materiels and methods

2.1. Participants

Fifteen paid volunteers (seven females, two left-handed, mean age = 23.6 ± 3.83; range: 29–34 years old) were included in the study. All participants had normal (self-reported) or corrected vision. Written informed consent was obtained from all participants prior to the experiment.

2.2. Stimuli

Two sets of 43 colored photographs of full front faces (21 males) and cars were used (Fig. 1A). Faces were presented without glasses, facial hair or make-up, and with neutral expression. All face pictures were trimmed to remove their variable backgrounds, clothing and hairline using Adobe® Photoshop® 7.0. Car pictures were also edited to remove background. The RGB stimuli (8-bits per channel) had a resolution of 72 pixels/inch. Mean normalized pixel intensity (0–1) was slightly lower for cars (0.58 ± 0.01) than faces (0.59 ± 0.01) but this difference was negligible (for detailed information about the stimuli in the three RGB channels, see Table 1). Two additional sets of stimuli were made by scrambling the faces and the cars using a Fourier phase randomization procedure (Jacques & Rossion, 2004; Nasanen, 1999; Rousselet, Pernet, et al., 2008; Sadr & Sinha, 2004; VanRullen, 2006) (Fig. 1B). The phase randomization procedure replaces the phase spectrum of the images with random values, keeping the AS of the image unaltered (unlike quantization or other scrambling manipulations applied to faces for making control stimuli, e.g., Herrmann et al., 2005). It yields images that preserve almost completely the global low-level properties of the original image (i.e., luminance, global contrast – see Table 1 for details –, color, spatial frequency amplitude spectrum) while completely degrading shape.

All stimuli subtended ~3.72° × 4.24° of visual angle. Segmented faces and cars were placed on a darker grey background than the general background of the monitor so that the size of the overall visual stimulation was exactly the same for faces and cars compared to their scrambled counterparts.

2.3. Procedure

After electrode-cap placement, participants were seated in a light- and sound-attenuated room, at viewing distance of 100 cm from a computer monitor. Stimuli were displayed using E-prime 1.1, on a light grey background. In each trial, a fixation point displayed at the centre of the screen for 100 ms, followed approximately 300 ms (randomized between 200 and 400 ms) later by the presentation of the test stimulus during 300 ms. The offset of this stimulus was followed by an inter-trial interval of about 1700 ms (1600–1800 ms). Participants were asked to judge whether the presented stimulus was an object (face or car) or else a "texture" (scrambled versions), and gave their response by pressing one of two keys with their dominant hand. This task was irrelevant with respect to the goals of the experiment, and was used to maintain participants' level of attention constant during the task. Participants were not aware that half of the

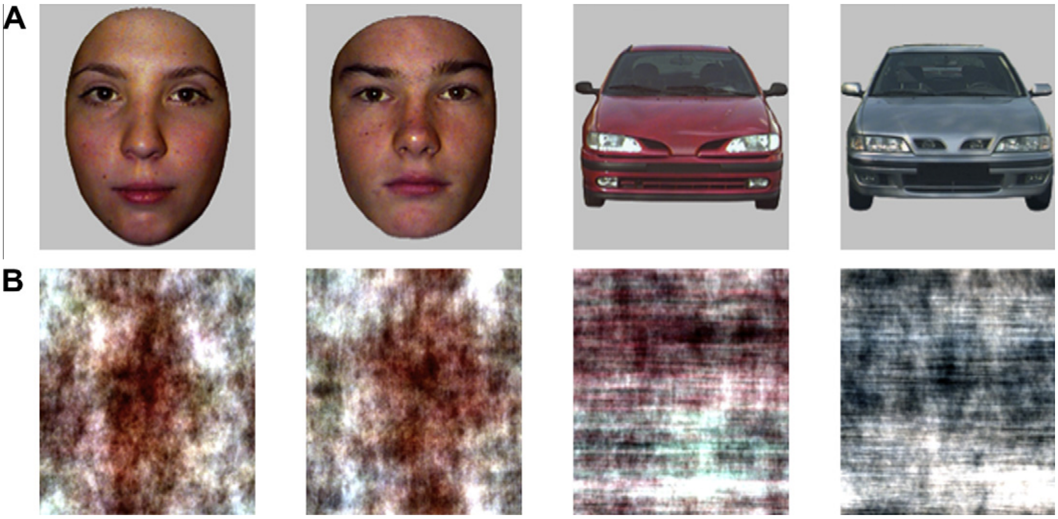

Fig. 1. Examples of stimuli used in the experiment. Face and car stimuli (A) and their corresponding Fourier phase-scrambled version (B).

Table 1
Above. Average pixel intensity values (±SD) in the three color channels (RGB) for the pictures of faces, cars and their phase-scrambled versions used in this study. Below. Average standard deviation of pixel intensity (global contrast). Average pixel intensity was not different between pictures of faces and cars, but mean standard deviations of pixel intensities (global contrast) were slightly higher for pictures of cars than faces. There are also minor differences in the mean luminance and global contrast between intact and phase-scrambled images. These differences are due to the fact that the phase scrambling procedure generates pixel values outside the [0 255] range even though the mean and standard deviation of the pixel values distribution are equal to that from the intact image. These out-of-range values are clipped respectively to 0 or 255 when writing the actual image, which result in a minor change in the mean and standard deviation of the phase scrambled image. On average, less than 5% of the pixels were clipped that way.

		Faces	Scrambled faces	Cars	Scrambled cars
Average Pixel intensity (across images)	**Global**	**0.59 ± 0.01**	**0.58 ± 0.01**	**0.58 ± 0.01**	**0.57 ± 0.01**
	Red	0.70 ± 0.01	0.69 ± 0.03	0.58 ± 0.03	0.58 ± 0.03
	Blue	0.52 ± 0.02	0.51 ± 0.02	0.58 ± 0.01	0.57 ± 0.01
	Green	0.55 ± 0.02	0.55 ± 0.02	0.57 ± 0.02	0.57 ± 0.02
Average standard deviation of pixel intensity (across images)	Global	0.220 ± 0.002	0.213 ± 0.003	0.238 ± 0.004	0.230 ± 0.004
	Red	0.162 ± 0.005	0.157 ± 0.005	0.239 ± 0.018	0.230 ± 0.016
	Blue	0.250 ± 0.004	0.241 ± 0.005	0.240 ± 0.008	0.232 ± 0.009
	Green	0.279 ± 0.005	0.262 ± 0.006	0.236 ± 0.011	0.228 ± 0.011

phase-scrambled stimuli were made of faces, and the other half of cars, and no participant noticed anything about the relation between these phase-scrambled stimuli and the intact stimuli. They were instructed to maintain eye gaze fixation to the centre of the screen (tip of the nose for faces, or mid-distance between the headlights of the cars) and during the whole trial and to respond as accurately and as fast as possible. Participants performed four blocks of 86 trials (344 trials in total with 86 trials per condition; the 43 stimuli in each set repeated two times each). The order of conditions was randomized within each block.

2.4. EEG recording

EEG was recorded from 128 Ag/AgCl electrodes mounted in an electrode cap (Waveguard, ANT; 2D map of all electrode positions can be accessed here: http://www.ant-neuro.com/products/caps/waveguard/layouts/128/). Electrode positions included the standard 10–20 system locations and additional intermediate positions. Vertical and horizontal eye movements were monitored using four additional electrodes placed on the outer canthus of each eye and in the inferior and superior areas of the right orbit. During EEG recording, all electrodes were referenced to the left mastoid reference, and electrode impedances were kept below 10 kΩ. EEG was digitalized at a 1000 Hz sampling rate and a digital anti-aliasing filter of 0.27∗ sampling rate was applied at recording (at 1000 Hz sampling rate, the usable bandwidth is 0–∼270 Hz). After a 0.1 Hz high-pass and 30 Hz low-pass filtering of the EEG, trials contaminated with eye movements or other artifacts (⩾±75 μV in −200 to 800 ms) were marked and rejected (mean number of trials was 76.5, range across conditions: 75–78). Incorrect trials and trials containing EEG artifacts were removed. For each participant, averaged epochs ranging from −200 to 800 ms relative to the onset of the stimulus and containing no EEG artifact were computed for each condition separately and baseline corrected using the 200 ms pre-stimulus time window. Participants' averages were then re-referenced to a common average reference.

2.5. Statistical analyses

2.5.1. Behavior

Accuracy (percent correct) and mean correct response times (trimmed mean, values above of below 2 SDs removed) were computed for each of the four conditions and were submitted to a repeated-measures analysis of variances (ANOVA) with *Shape* (Normal vs. Scrambled) and *Category* (Face vs. Car) as within-subject factors.

2.5.2. Electrophysiology

Two clear visual components elicited by the different sets of stimuli were analyzed: the P1 (maximal at approximately 95 ms), and the N170 (maximal at approximately 150 ms) (Fig. 2). Amplitude values of these components were measured at the different pairs of occipito-temporal electrodes in the left and right hemisphere where they were the most prominent (for the P1: PO7/8, PPO9/10 h, POO9/10 h, PO5/6, PO9/10, and for the N170: P7/8, PO7/8, TPP9/10 h, PO9/10, P9/10, I1/2, POO9/10 h, PPO9/10 h)

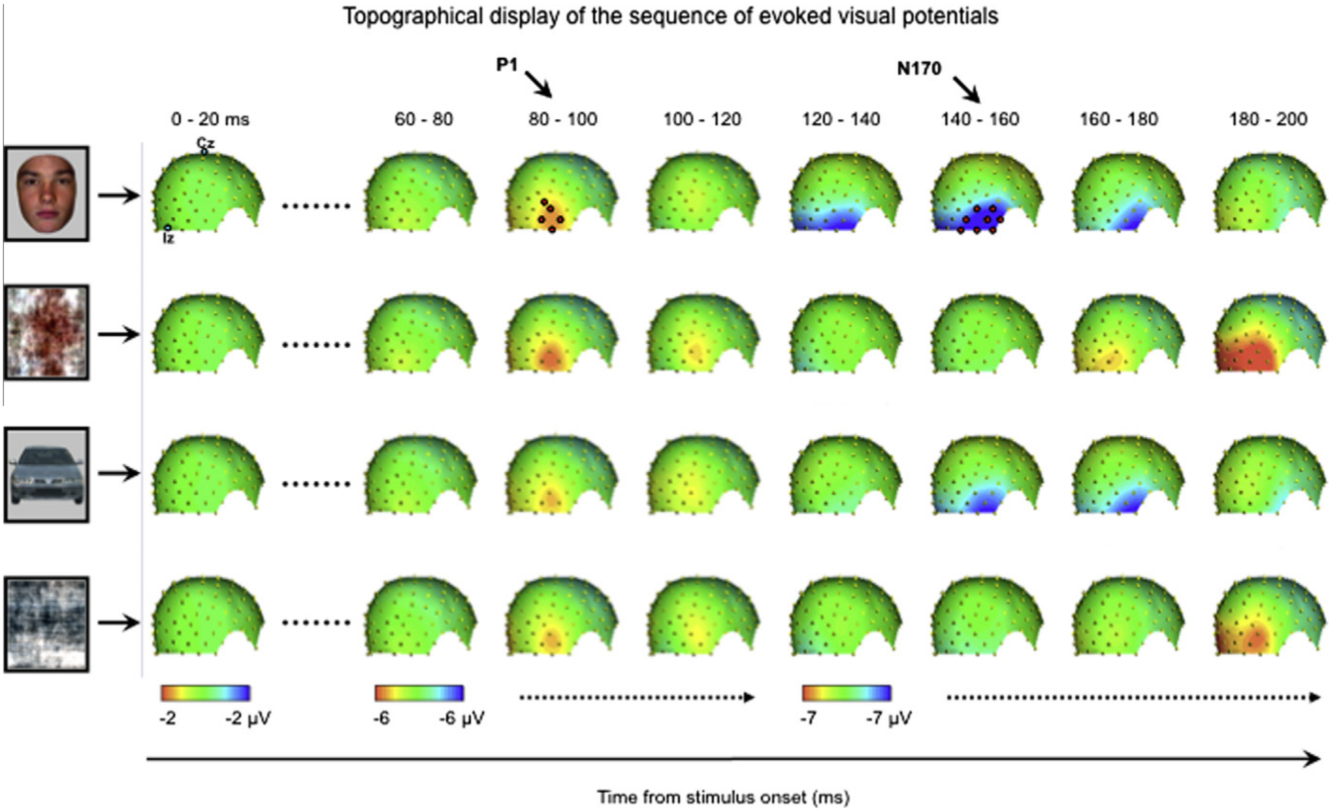

Fig. 2. Topographical maps from grand-averaged data of all (n = 15) participants representing the scalp distribution for each condition (Face, Scrambled Face, Car, Scrambled Car) within 20 ms temporal window from 0 to 200 ms after stimulus onset. The electrodes sites represented in a red circle were defined based on the topographical maps and used to analyze the P1 and N170 components. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Fig. 2). Amplitudes were quantified for each condition as the mean voltage measured within 30 ms windows centered on the grand average peak latencies of the components' maximum. Peak latency of both components were extracted automatically at the maximum amplitude between 60 and 140 ms for the P1, and at the minimum value between 110 and 190 ms for the N170 at the two pairs of occipito-temporal electrodes (PO9/10, PPO9/10 h) where these components were maximal. The amplitude and latency values of each component were then submitted to separate repeated-measures analysis of variance with *Shape*, *Category*, *Hemisphere*, and *Electrode* as within-subject factors. All effects with two or more degrees of freedom were adjusted for violations of sphericity according to the Greenhouse–Geisser correction. In order to track precisely the time course of *Category* or *Shape* effects at occipito-temporal channels, ERP differential waveforms between two conditions were computed at each time point between −200 and 400 ms at two pairs of occipito-temporal electrodes (PO9/10, PPO9/10 h). Conditions were compared by pairs (faces vs. cars; scrambled faces vs. scrambled cars; difference of the differential waveforms computed individually) on each time-point (paired t-test, df = 14). Differences were considered to be significant if they reached $p < .05$ for 10 consecutive time-points (10 ms).

3. Results

3.1. Behavioral results

Participants performed the categorization task with a mean accuracy rate of 97.6% across the four conditions (range between 97.1% and 98.2%), and similar RTs (between 409 and 431 ms). There were no significant differences in accuracy rates between conditions (all $ps > 0.12$). For correct response times (RTs), participants were faster to classify normal than phase-scrambled stimuli ($F(1, 14) = 6.63$; $p = .022$). The main effect of *Category* did not reach significance ($F(1, 14) = 4.17$; $p = .06$). However, there was a significant interaction between *Shape* and *Category* ($F(1, 14) = 8.14$; $p = .0012$), because participants were faster for faces (409 ms) than cars (428 ms) in normal shape ($F(1, 14) = 6.82$; $p = .02$), but there was no significant difference for their scrambled counterparts ($F(1, 14) = 0.65$; $p = .43$) (431 and 429 ms, respectively).

3.2. Eletrophysiological data

3.2.1. Event-related components

Fig. 2 illustrates the sequence of visually evoked responses on the scalp following stimulus presentation, and Fig. 3 displays the superimposition of all electrode channels between −100 ms and 400 ms following stimulus onset, for all four conditions. The first (Fig. 3) electrophysiological event recorded on the scalp was characterized by a large positivity (P1) recorded bilaterally over lateral occipital sites. It peaked shortly before 100 ms. It was observed for all conditions, but appeared of stronger magnitude in response to faces and scrambled faces than cars and scrambled cars (Figs. 2 and 3; Table 2), at least in the right hemisphere (Figs. 2 and 4). It was followed by a large negative deflection starting between 120 and 140 ms, mainly for intact stimuli, and peaking at about 150–160 ms, the N170. The N170 was observed mainly for intact stimuli, being extremely small in negative amplitude for phase-scrambled stimuli (Figs. 2 and 4; Table 2). The topography of the P1 and N170 differed clearly also, with the P1 being maximal at lateral occipital sites, while the N170 (for faces and cars) was maximal at lower and more anterior temporal channels (Figs. 2 and 4). Also, the P1 showed little or no lateralization, while the N170 appeared much larger in the right hemisphere, in particular for faces (Fig. 4). These two components of interest were most prominent at

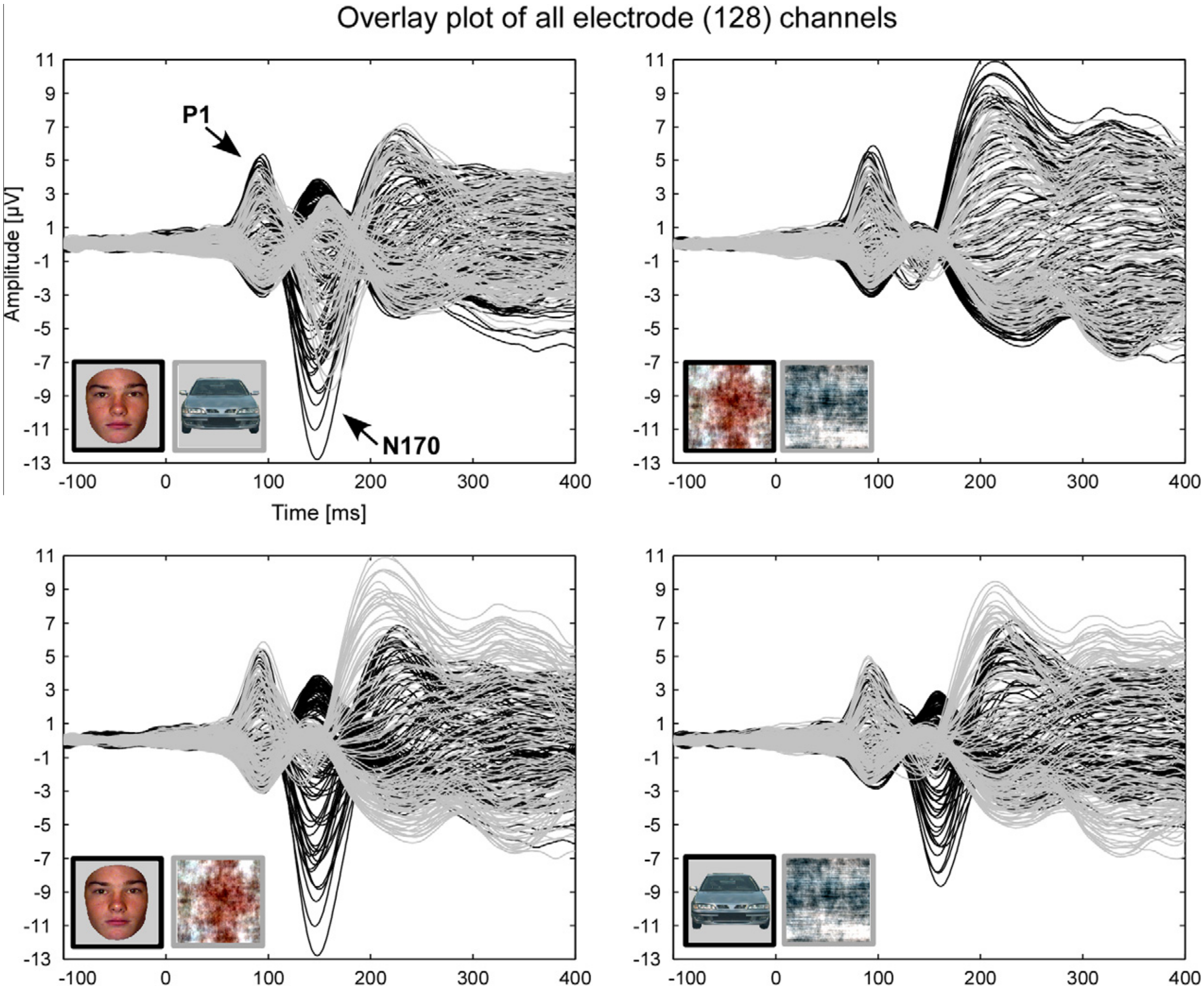

Fig. 3. Grand average ERPs elicited by all condition pairings across all channels ($N = 128$).

Table 2
Mean amplitude (±standard deviation) on the channels of interest for the P1 and the N170 amplitude in the left and the right hemisphere for the four experimental conditions.

		P1 amplitude (µV)		N170 amplitude (µV)	
		LH	RH	LH	RH
Intact shapes	Face	3.89 ± .55	4.38 ± .79	−5.49 ± .91	−7.85 ± 1.43
	Car	3.54 ± .58	3.78 ± .87	−4.60 ± .92	−5.30 ± 1.12
Scrambled shapes	Face	4.10 ± .67	4.77 ± .92	−.71 ± 1.04	−.32 ± 1.20
	Car	4.05 ± .73	3.74 ± .84	−.89 ± .92	−.72 ± 1.05

posterior electrodes sites and followed by large variations of amplitude across the whole scalp (Fig. 3), which were not the focus of the present study and were not analyzed further.

3.2.2. P1 analysis

The P1 latency did not vary significantly between conditions, as indicated by the absence of main effects of *Shape* ($F(1,14) = 2.19$; $p = .16$), or *Category* ($F(1,14) = 1.56$; $p = .23$), and of interaction between the two factors ($F(1,14) = 3.02$; $p = .13$).

With respect to P1 amplitude, there was a significant interaction between *Category* and *Hemisphere* ($F(1,14) = 6.28$; $p = .025$) because the P1 was larger for faces than cars in the right hemisphere ($F(1,14) = 7.05$; $p = .0188$), but not in the left hemisphere ($F(1,14) = 0.78$; $p = .391$) (Figs. 4 and 5). There were no other significant effects or interactions on the P1 amplitude (all $ps > .05$), no effect of *Shape* ($F(1,14) = 0.93$; $p = .35$), and most importantly, no interaction between *Category* and *Shape* ($F(1,14) = .08$; $p = .78$) (Figs. 4 and 5).

In summary, the P1 was larger in response to faces than cars, but this difference was similar in amplitude for intact and phase-scrambled stimuli. Interestingly, there was a highly significant correlation across individual participants between the face vs. car amplitude difference observed for intact and phase-scrambled stimuli, in the right hemisphere only ($r = 0.61$; $t = 2.78$, $p < 0.01$) (Fig. 6) (left hemisphere ($r = 0.21$, NS).

3.2.3. N170 analysis

The N170 latency was shorter in response to faces than cars ($F(1,14) = 46.86$, $p < .0001$), and in response to scrambled compared to normal shapes ($F(1,14) = 12.04$, $p = .0037$). There was a significant *Shape* × *Category* interaction ($F(1,14) = 10.30$, $p = .006$) due to a smaller difference between faces and cars in scrambled versions ($F(1,14) = 7.72$, $p = .015$, earlier for scrambled faces than

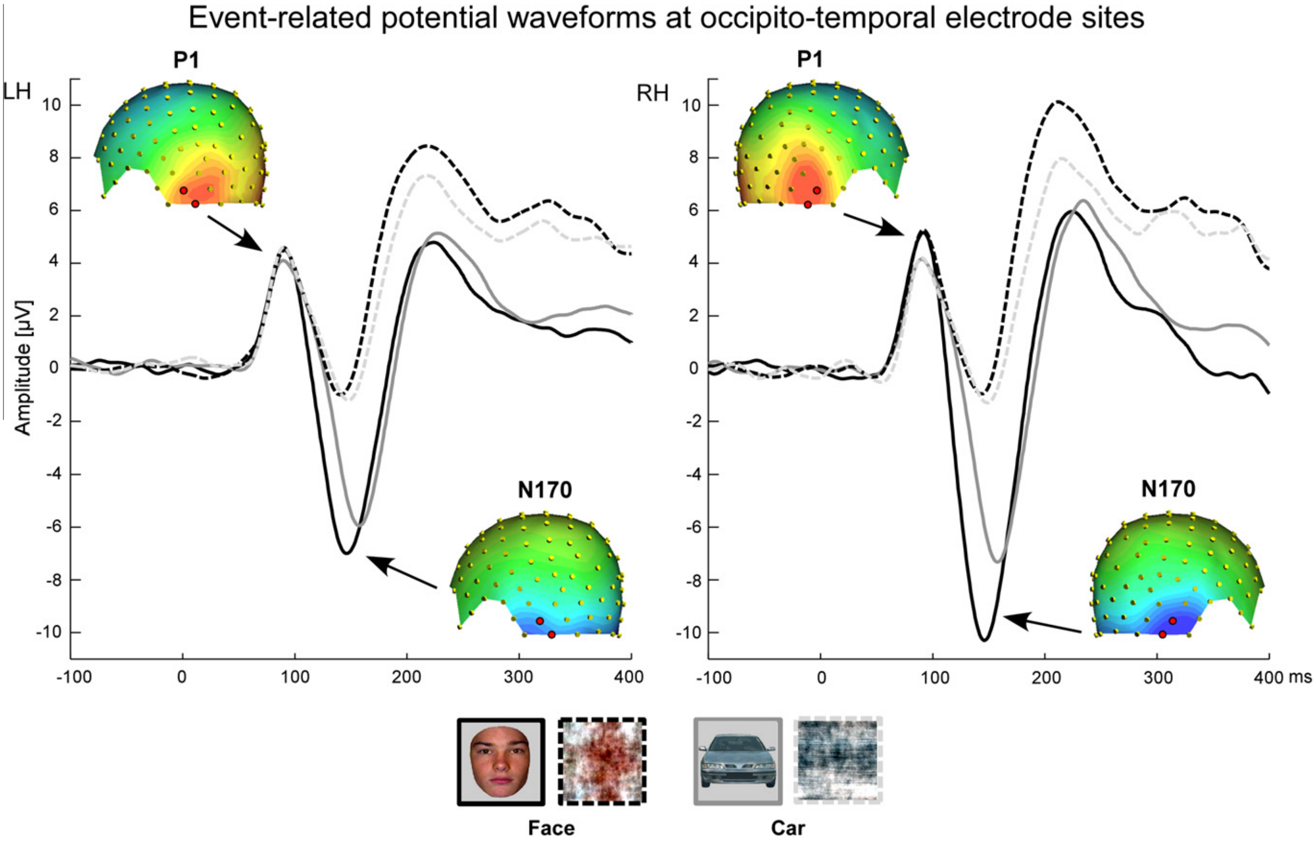

Fig. 4. Grand average ERP waveforms elicited by all four conditions at left and right pooled occipito-temporal electrode sites (waveforms averaged for electrodes PPO9/10 h, PO9/10) where the P1 and N170 components peaked maximally. The maps represent the scalp distribution of the P1 (at around 95 ms) and the N170 (at around 150 ms) components for intact face stimuli.

scrambled cars) as compared to the difference found for intact versions ($F(1, 14) = 35.75$, $p = .00003$; earlier for faces than cars).

The N170 component was much larger for normal than scrambled stimuli ($F(1, 14) = 44.09$, $p < .0001$) (Figs. 3 and 4), and also larger for faces than cars ($F(1, 14) = 9.95$, $p = .007$). There were three significant two-ways interactions between *Category*, *Shape* and *Hemisphere* factors ($F(1, 14) = 5.72$, $p = .031$; *Category* × *Hemisphere*: $F(1, 14) = 4.65$, $p = .049$; *Shape* × *Category*: $F(1, 14) = 27.75$, $p = .0001$).

However, and most interestingly, these effects on N170 amplitude were qualified by a significant triple interaction between *Shape*, *Category* and *Hemisphere* ($F(1, 14) = 8.32$, $p = .012$) (Figs. 3 and 4). This interaction was due to a larger amplitude in response to normal faces than cars particularly in the right hemisphere (RH) ($F(1, 14) = 24.63$, $p = .0002$; Left hemisphere (LH): $F(1, 14) = 4.85$, $p = .045$), while there was no such effect for scrambled stimuli (RH: $F(1,14) = .38$, $p = .54$; LH: $F(1, 14) = 3.22$, $p = .09$). This interaction reveals also that the N170 amplitude to normal faces was significantly larger in the right than in the left hemisphere ($F(1, 14) = 5.48$, $p = .034$), while for the normal cars and for both scrambled shapes, the interhemispheric differences were not significant (all $ps > .05$).

Finally, the N170 was larger on the antero-lateral compared to the more postero-medial electrodes (main effect of *Electrodes*, $F(7, 98) = 6.88$, $\varepsilon = .29$, $p = .0035$) which was qualified by a significant two-ways interaction between *Shape* and *Electrode* ($F(7, 98) = 10.5$, $\varepsilon = .34$, $p = .0001$). For normal shapes, the N170 was larger on antero-lateral than postero-medial electrodes ($F(7, 98) = 11.25$, $\varepsilon = .32$, $p = .0001$), while this was not true for scrambled shapes ($F(7, 98) = 1.85$, $\varepsilon = .37$, $p = .16$).

3.2.4. Peak-to-peak analyses

To isolate the N170 differences from any differences already present at the level of the P1, peak-to-peak analyses were performed on latencies and amplitudes of these components at the same two pairs of occipito-temporal electrodes (PO9/10, PPO9/10 h) where these components were maximal (P1 values were subtracted from N170 values.

These analyses indicated shorter N170-P1 latencies for scrambled than normal shapes ($F(1, 14) = 35.72$, $p < .0001$), and for faces than cars ($F(1, 14) = 24.75$, $p = .0002$). There were no other significant effects or interactions.

With respect to amplitude, the P1-N170 difference was larger for normal than scrambled stimuli ($F(1, 14) = 47.39$, $p < .0001$), and was also larger for faces than cars ($F(1, 14) = 8.20$, $p = .012$). However, this difference between normal faces and cars was significant in the right hemisphere only ($F(1, 14) = 19.95$, $p = .0007$; LH: $F(1, 14) = 3.54$, $p = .08$). There was no such difference for scrambled stimuli (RH: $F(1, 14) = .98$, $p = .34$; LH: $F(1, 14) = .41$, $p = .53$) (*Shape* × *Category* × *Hemisphere* interaction: $F(1, 14) = 5.99$, $p = .028$; *Shape* × *Category* interaction: $F(1, 14) = 19.79$, $p = .0005$).

In summary, the N170 was much larger for intact as compared to scrambled stimuli. Furthermore, it was larger to intact faces than cars in the right hemisphere, but there was no such difference between the two kinds of scrambled stimuli.

3.2.5. Time-points analyses at occipito-temporal sites

Time-points analyses performed on occipito-temporal sites showed a significant amplitude difference between faces and cars first between 87 and 107 ms following stimulus onset (P1 time-window), in the right hemisphere. However, an early faces vs. cars

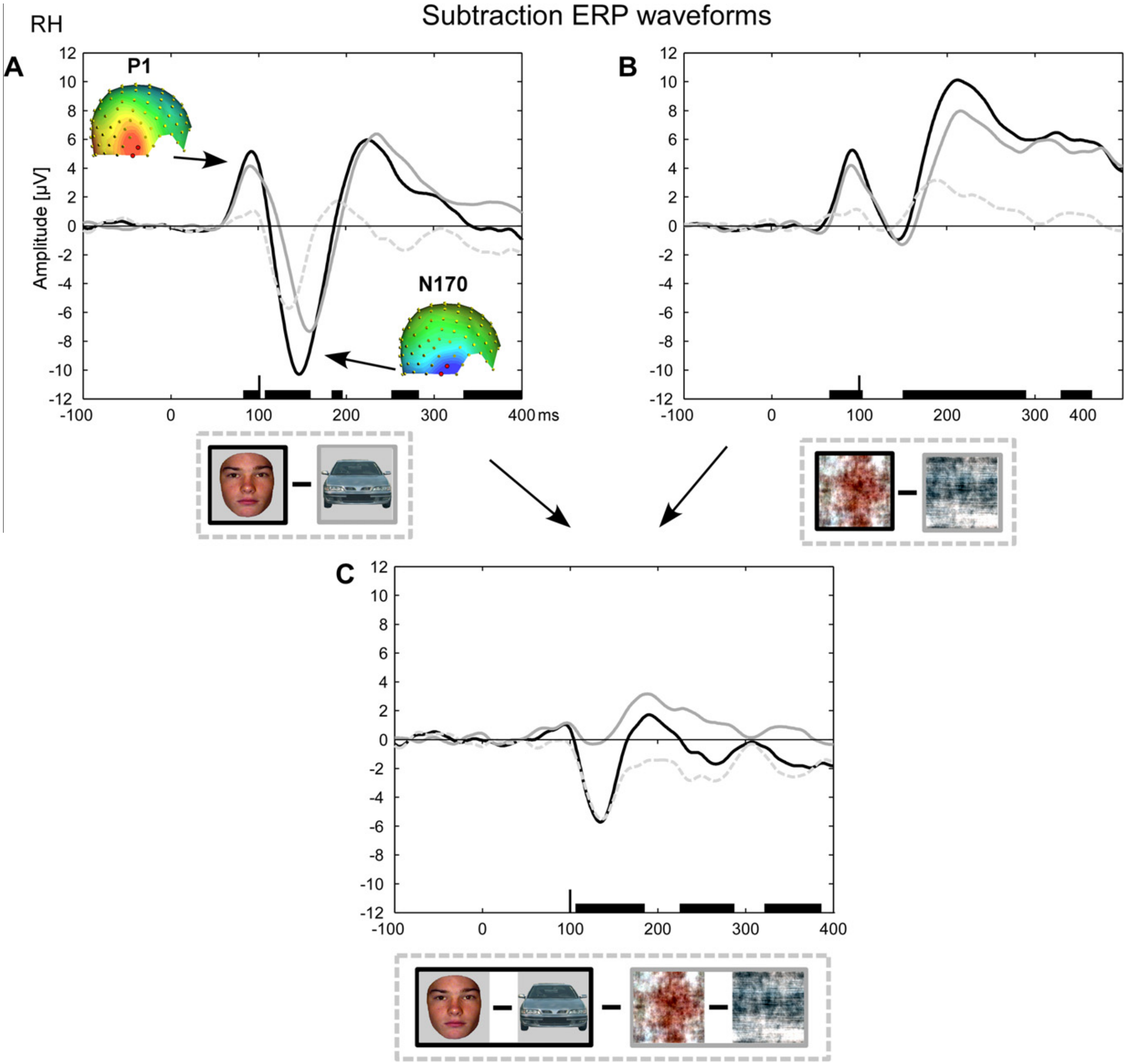

Fig. 5. ERP waveforms at right occipito-temporal electrode sites (waveforms averaged for electrodes PPO10 h and PO10). (A) ERP waveforms elicited for faces, cars, and the subtraction 'faces – cars' showing a difference before 100 ms (P1 range) and during the N170 time-window. Time points at which two conditions differed significantly ($p < .05$) are represented by horizontal black thick line along the horizontal axis. (B) Waveforms at the same channels for scrambled faces, scrambled cars and their subtraction waveform, also showing the difference before 100 ms between these two conditions. There were late (post N170) differences also, which not found for intact stimuli and were not the focus of this study. (C) The two subtraction waveforms are superimposed on each other, together with the interaction (difference of the difference), revealing only face-sensitivity effects that cannot be accounted for by low-level cues, after 100 ms.

difference for scrambled stimuli was also found at around the same latency or even earlier (between 68 and 104 ms) (Fig. 5). At a second temporal window, between 108 and 157 ms (N170), there was a significant difference between normal faces and cars. However, there were no such effects for scrambled stimuli. In the left hemisphere, there were no significant differences between conditions before 100 ms, confirming amplitude analysis of the P1 component. From 107 ms on, the effects were similar to the effects observed in the right hemisphere.

In addition, subtraction analyses between the differential waveforms elicited for faces and cars (ERP elicited by faces *minus* ERP elicited by cars) for normal shape and scrambled stimuli were performed (Fig. 5). This analysis enabled to identify precisely the time-point at which the *Category* and *Shape* factors interact. It identified a first difference between 107 and 187 ms over the right occipito-temporal cortex, that is during the N170 time-window (109–159 ms in the LH).

3.2.6. Analysis of spatial standard deviation across all channels

Even though the effects appeared restricted to the locations of interest where the components were the most prominent, the difference between these two differential waveforms [(Face-Car) – (Srambled Face–Scrambled Car)] was also analyzed at each time point using the spatial standard deviation across all electrodes of the scalp (Fig. 7). This measure, which is referred by some authors as the global field power (GFP, Lehmann & Skrandies, 1980) provides a compact description of the signal across the scalp. It is assumed that stronger electric fields lead to larger values and that the peaks coincide with maximum activation of the underlying generator. This analysis showed that the first period of significance

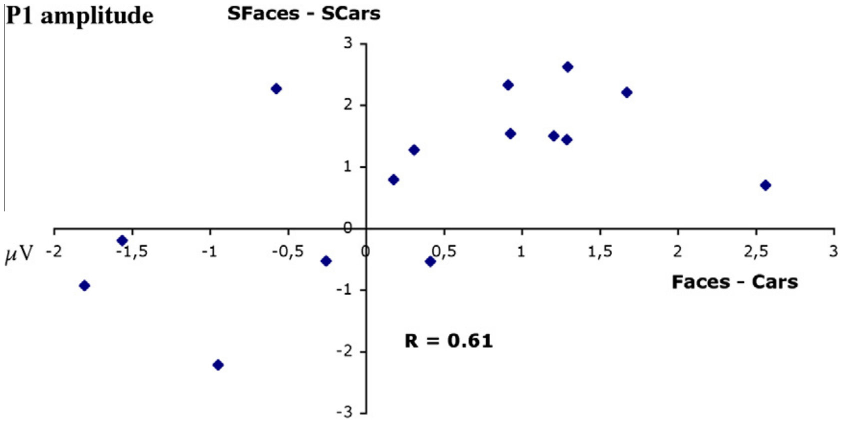

Fig. 6. Illustration of the significant correlation between the P1 amplitude difference for intact faces and cars (*X*-axis) and the difference for scrambled faces and cars (*Y*-axis) across the 15 participants.

occurred between 127 and 155 ms post-stimulation (Fig. 7), well after the P1 temporal window.

In summary, millisecond-by-millisecond statistical analyses revealed that the first difference between faces and cars occurs as early as 80–90 ms following stimulus onset, during the early P1 time-window, in the right hemisphere. However, this difference is similar in amplitude for intact and phase-scrambled shapes, meaning that this early category effect appears to be related to low-level visual cues but not to the perception of faces *per se*. In contrast, the second time-window is related to an effect found specifically for intact stimuli, so that it is driven mostly by the phase information, which is associated with the phenomenological experience of a face (or car).

4. Discussion

To summarize, the lateral occipital P1 component, with an average peak latency of 90–100 ms here, was larger in amplitude for pictures of faces than cars. However, this difference was similarly found for phase-scrambled versions of these categories, which contained no information about the structure of the stimulus, and were not perceived as faces or cars. The significant correlation across individual participants between the face vs. car P1 amplitude difference for intact and phase-scrambled stimuli in the right hemisphere supports the view that the P1 face-sensitivity is accounted for by low-level visual cues. Also, the P1 amplitude did not differ between intact and phase-scrambled stimuli. In contrast, the next visual component, the occipito-temporal N170 was almost non-existent for phase-scrambled stimuli, and larger in amplitude for both intact faces and cars. It was also much larger for faces than cars, in particular in the right hemisphere, as previously reported. In contrast to the P1, there was no N170 amplitude difference between phase-scrambled versions of these stimuli.

The results of this experiment are clear and their interpretation is quite straightforward. Nevertheless, we believe that they carry significant information regarding the time course of face categorization in the human brain.

4.1. Face-sensitivity before 100 ms is driven by low-level visual cues

The larger P1 for pictures of faces than cars is in line with previous observations of larger P1 to faces than objects in general (e.g., Eimer, 1998; Goffaux et al., 2003; Herrmann et al., 2005; Itier & Taylor, 2004a, 2004b; Liu et al., 2002; Okazaki et al., 2008). Here we show for the first time that this P1 amplitude difference between faces and cars is similar for both intact and scrambled stimuli. That is, the P1 face effect can be accounted for by low-level visual cues preserved by phase-scrambling. It has nothing to do with phase information, that is, the kind of information associated with the phenomenological experience of a face, or a face percept. As indicated in the introduction, this observation is in agreement with previous studies showing that such early P1 face effects can disappear when schematic black-and-white drawings (Halgren et al., 2000), or two-tone images of faces ("Mooney faces", George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) are presented instead of face photographs. Such P1 face effects are also absent when faces and nonface objects are equalized according to AS (Rousselet et al., 2005, 2008).

In previous studies, face-sensitivity on the P1 was reported either on medial (O1, O2) or lateral occipital electrodes, or both (Eimer, 1998, 2000; Goffaux et al., 2003; Herrmann et al., 2005; Itier & Taylor, 2004a, 2004b; Liu et al., 2002; Okazaki et al., 2008; Sadeh et al., 2010). Here, the main face effect on the P1 – independently of phase-scrambling – was observed on lateral occipital electrodes, were the P1 was the largest. However, there was no evidence of an effect associated specifically with high-level face stimulation elsewhere on the scalp, in this time-range.

One may argue that with MEG rather than EEG, early (M1) face-sensitivity might not be fully attributable to low-level visual cues. However, the available evidence from MEG studies rather support the present findings, showing that P1 and M1 face-effects are highly similar (e.g., Linkenkaer-Hansen et al., 1998), and that M1 face-sensitivity is also related to low-level cues (Halgren et al., 2000; Tanskanen et al., 2005). Also worth mentioning is the fact the P1 face effect was found also in the left hemisphere in some studies (e.g., Eimer, 1998; Goffaux et al., 2003; Herrmann et al., 2005; Itier & Taylor, 2004a, 2004b), while there was no evidence of such an effect in the left hemisphere here (Fig. 4). Nevertheless, the strong effect found exclusively in the right hemisphere here suggests that the well-known RH advantage in face processing (e.g., Landis, Regard, Bliestle, & Kleihues, 1988; Parkin & Williamson, 1987; Sergent, Ohta, & Macdonald, 1992) is not only related to high-level representations of faces, but partly to low-level visual cues. Interestingly, this observation is compatible with the fact that faster saccades can be made towards faces presented in the left compared to the right visual field (Crouzet et al., 2010).[2] More generally, and to answer the question asked at the beginning of this paper, the present finding reveals that the pre-100 ms face effects

---

[2] Although this left visual field (LVF) advantage was true both for scenes containing faces or vehicles in the study of Crouzet et al., 2010, this indicates that earlier differences in saccadic RTs between faces and objects can be found for LVF than right VF stimulation.

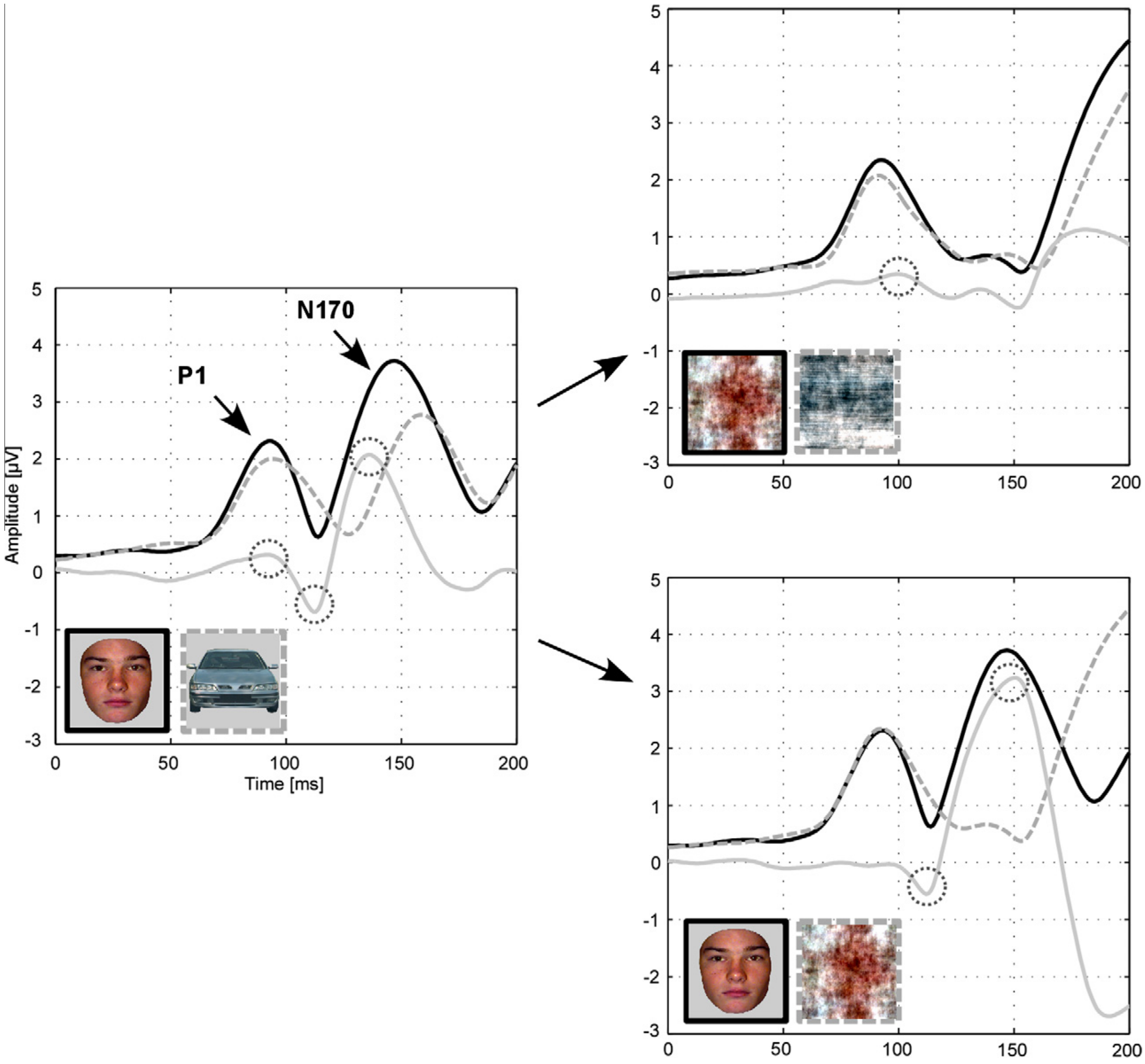

Fig. 7. Time course of the spatial standard deviation across all electrodes of the difference between faces and cars both in intact and scrambled shapes. Time points at which the two differential waveforms differed significantly ($p < .05$) are represented by circles in dashed line around the peak.

observed in ERPs that are compatible with the timing of the fastest saccadic behavioral RTs are directly related to low-level visual cues. Thus, they strongly suggest that the fastest saccades towards faces are also based primarily on such low-level cues, most probably AS information. This claim is in line with studies showing that rapid image recognition can be biased by priming using AS (amplitude spectrum) information (Guyader, Chauvin, Peyrin, Hérault, & Marendaz, 2004; Kaping, Tzvetanov, & Treue, 2007).

4.2. Implications for the interpretation of early (<100 ms) face-sensitivity effects in the human brain

Taken together with these previous sources of evidence, the present observations put strong constraints on the interpretation of previously reported early P1/M1 face-related effects: such effects are likely to be driven by low-level visual cues and should therefore not be interpreted as reflecting the activation of high-level face representations. For instance, Liu et al. (2002) found a larger M1 to faces than objects, which was interpreted as a face detection stage. In one experiment of that study, the M1 was larger for faces with parts rearranged than when parts in their correct locations were masked by black ovals, leading the authors to conclude that the M1 reflected an early face detection stage based on facial parts. In fact, the present study suggests that the information driving the differential amplitude of such early M1 responses has nothing to do with face parts *per se*. Rather, this effect might be due to low-level visual cues such as spatial frequencies that were minimized in that study when black ovals covered facial parts. Supporting this point, there is evidence that amplitude spectrum differences between faces and other complex visual stimuli are conveyed primarily by internal facial features (Keil, 2008).

In the same vein, the P1 component can be larger in amplitude for inverted than upright faces (e.g., Itier & Taylor, 2002, 2004a, 2004b; M1 in MEG also: see Linkenkaer-Hansen et al., 1998), again showing in some studies the same kind of effect that has been consistently found at the level of the N170/M170 (e.g., Itier & Taylor, 2004a, 2004b; Linkenkaer-Hansen et al., 1998; Rossion et al., 1999, 2000; Sagiv & Bentin, 2001). While some authors have interpreted this effect as reflecting an initial P1 stage of holistic/configural encoding for faces (Itier & Taylor, 2002, 2004a, 2004b; see also Halit, de Haan, & Johnson, 2000), the present results suggest that this interpretation was premature, since high-level face representations might not even have been activated yet at that latency in the human brain. Of course, regarding face inversion, the P1 amplitude increase cannot be easily accounted for by low-level visual cues as manipulated here, since spatial frequency spectrum and color are preserved by inversion. However, P1 amplitude and latency are highly sensitive to the upper vs. lower visual field stimulation (Di Russo et al., 2002), so that simple parameters such as the balance of contrast in the upper vs. lower visual field for upright and inverted faces may well account for such early P1 face inversion effects (Jacques & Rossion, 2009a). Supporting this claim, matching individual faces presented at multiple orientations (0–360°, 30° steps) leads to the same tuning functions for behavior and EEG amplitude variations with angle of face rotation only *after* the P1 peak, in the time-window of the N170 (Jacques & Rossion, 2007; see also Jemel, Coutya, Langer, & Roy, 2009).

The present data also suggest that other well-known effects, such as P1 amplitude or even latency variations for faces with different facial expressions, in particular faster and larger responses to fearful faces (Batty & Taylor, 2003; Pizzagalli, Regard, & Lehmann, 1999; Pourtois, Dan, Grandjean, Sander, & Vuilleumier, 2005; Vlamings, Goffaux, & Kemner, 2009) may not be directly related to the perception of facial expressions *per se* but to low-level cues that could be associated with different facial expressions. Similar modulations of electrophysiological responses well before 100 ms, for instance to facial expression (Liu & Ioannides, 2010) should also be interpreted with great caution. Finally, a number of studies in social neuroscience have reported modulations by socially relevant categories of faces either on the P1 or its negative counterpart on central sites (e.g., sex-typical vs. atypical faces, Freeman, Ambady, & Holcomb, 2010; black vs. white faces, Ito & Urland, 2003, 2005; Kubota & Ito, 2007). Typically, these early effects are interpreted as demonstrating an early activation of social categories. However, what our observations, and those of previous studies as cited above, suggest is that these early effects precede face perception *per se*. It follows that such effects might reflect low-level cues that vary systematically with different sets of faces and can be picked up by the perceptual system to categorize faces socially.

Finally, beyond P1/M1 effects, our observations also put constraints on the interpretation of early (<100 ms) face-sensitive effects as observed with other methods. Most notably, disruption of face-selective processes after TMS applied at that latency to the right lateral occipital cortex (Pitcher et al., 2007) may be also based on low-level visual cues that differ between faces and nonface objects (e.g., houses) rather than high-level face representations *per se*.

In summary, our findings suggest that early face-sensitivity on the human scalp (P1 effects, around or before 100 ms) is driven by low-level visual cues, but not by structured information associated with the percept of a face. Admittedly, this may be true only when human observers do not know in advance if a visual stimulus to be presented is going to be a face or another kind of stimulus (as here). However, there are experimental situations in which participants of a given experiment are presented only with face stimulation, or are told to pay attention to faces only, or even to a subcategory of faces only (e.g., fearful faces, asian faces, likeable faces, etc.). In this case, one cannot exclude that attentional factors and expectancies (Summerfield & Egner, 2009) lead to increased early face-sensitivity effects, perhaps around or before the P1 time-range, which may be driven and modulated in a top-down fashion by pre-activated representations of faces in high-level visual areas, or areas of the anterior temporal and prefrontal cortices (e.g., Gamond et al., 2011; for recent neuroimaging evidence without precise time information see Chiu, Esterman, Han, Rosen, & Yantis, 2011; Esterman & Yantis, 2010).

4.3. Spatial frequencies, color and other potential low-level visual cues

What kinds of low-level visual cues can account for the early P1 face effects observed here? As mentioned, global amplitude spectrum of the images is the primary candidate, since it is preserved by the phase-scrambling of the images (unlike randomization of pixels or other forms of scrambling, which may lead to smaller P1 amplitudes than full faces, e.g., Herrmann et al., 2005). There is evidence that faces and nonface natural images vary in their AS distribution, with a steeper spectrum decrease for faces with spatial frequencies compared to natural images (Keil, 2008). Also, man-made categories have more energy in cardinal (i.e., vertical and horizontal) orientations compared to natural categories (Torralba & Oliva, 2003). This AS difference between faces and objects in itself may account for the early P1 face-sensitivity, and suggest that the human brain may be particularly tuned to this low-level information that characterizes faces, facilitating rapid face detection. That is, this information might play a functional role in face categorization, and should not necessarily be controlled in a given experiment (see discussion in Rossion & Jacques, 2008). For instance, AS has been claimed to contribute to face pop-out in visual search (VanRullen, 2006; but see Hershler & Hochstein, 2006) and may well contribute to fast saccades towards faces (Crouzet & Thorpe, 2010). The present observations cannot be taken as supporting the view that AS in itself is sufficient to fixate to and to categorize stimuli as faces accurately and rapidly (see Cerf et al., 2008), but they show that evidence for face categorization can be accumulated early on from such low-level visual cues.

Even though early saccades to faces were found on grayscale images (Crouzet et al., 2010), another low-level cue that might play a role here also in generating early P1 face effects is color. Indeed, there is evidence from electrophysiological data in human and monkeys that color can play a role (i.e., speed up processing) in early visual scene and face categorization (Edwards, Xiao, Keysers, Földiák, & Perrett, 2003; Goffaux et al., 2005, respectively). Moreover, even with highly degraded images, complex face categorization tasks such as sex judgments, can be made based on differential color distributions (Nestor & Tarr, 2008). Here, the stimuli we used varied in color information, which could be diagnostic in two ways. First, the characteristic color of a face – which is somewhat preserved in phase-scrambled stimuli (i.e., the mean and standard deviation within each RGB color channel is preserved) – can easily be contrasted with the more variable characteristics of color for car stimuli. Second, within a given face picture, there are more variations in color than within an artificial stimulus such as a car, usually painted in a uniform color.

Finally, one cannot exclude that other low-level cues may account for the early P1 difference between faces and other stimuli such as cars as used here. Here, luminance is not a candidate because it was controlled between the different kinds of stimuli presented. Global contrast was only slightly different for pictures of cars (intact or scrambled) and faces, and if anything it was in fact larger for pictures of cars than faces, so that it is also very unlikely to have contributed to the P1 effect. However, Scholte and colleagues (2009) found that parameters reflecting the distribution of contrast in natural images (roughly reflecting the number and the strength of edges in the image) can explain a large part of the variance in early ERP activity recorded over occipital sites, starting at about 80 ms, and with a peak at about 125 ms. While phase-scrambling reduces these low-level cues, some differences between faces and cars in terms of the distribution of contrast remain (see Table 1), which may also partly contribute to the early P1 effect as found here.

4.4. N170 face effect: activation of high-level face representations

Contrary to the P1, we observed a much larger N170 for faces or cars than phase-scrambled stimuli, indicating that the N170 response was *not* driven primarily by low-level visual cues. Rather, a large N170, or electrophysiological activity starting well after 100 ms, appears to be associated with the organization of visual cues into structured patterns. This result agrees with previous observations of a much larger N170 for faces compared to phase-scrambled faces (Rousselet, Pernet, et al., 2008) and extend them to nonface objects. It indicates that contrary to the P1, there must be a contribution of visual processes in areas that are sensitive to object shape, such as the lateral occipital complex (LOC; Malach et al., 1995), in generating the N170.

Moreover, with respect to the main goal of the study, we replicated once again the larger N170 amplitude in response to pictures of faces than a highly familiar visual category, cars. This effect is similar, albeit of a greater magnitude, to the effect found on the P1. Most importantly, contrary to the P1, this amplitude difference was not accounted for, even partly, by low-level cues such as amplitude spectrum and color: the N170 was not larger at all for scrambled faces than scrambled cars. This observation complements Rousselet et al. (2008)'s finding that N170 amplitude differences between faces and objects are preserved even when images differ only in phase information, that is, when amplitude spectra are equated across image categories. Here, the interest of the study is precisely that stimuli were *not* equalized in amplitude spectra, and also gave rise to differential P1 amplitude for faces and cars that were accounted for entirely by low-level cues. Hence, the present study does not only demonstrate that N170 face effect are still present when differential low-level cues between faces and objects are removed, but indicate that these low-level cues do not contribute at all to the N170 face effect.

To be more precise, our analyses indicate that while pre-100 ms face-sensitive effects are accounted for entirely by low-level cues, post-100 ms these cues do not play any role in the sensitivity to faces. That said, one cannot fully exclude from our study that other more complex low-level visual cues, such as those related to distribution of contrast in images, may also play a role in face categorization even between 100 and 200 ms following stimulus onset (see Scholte et al., 2009).

4.5. Why is the N170 larger (and earlier) to faces?

If low-level visual cues do not explain or even contribute to the larger N170 amplitude for faces than objects, then where does this effect come from?

From a neurophysiological point-of-view, the N170 corresponds to a massive increase of EEG power that is phase-locked and time-locked to stimulus onset (Rousselet, Husk, Bennett, & Sekuler, 2007; Rossion & Jacques, 2008). Faces – because of innate constraints or more likely as a by-product of visual experience – are overrepresented in high-level areas of the visual cortex. Besides the LOC, which responds to all object shapes without presenting clear category differences in its overall response, there are many areas of the visual cortex that respond preferentially – or even exclusively – to face stimuli. Sergent et al. (1992) were the first to note the particularly large cortical volume devoted to faces in their initial neuroimaging study, a finding confirmed by intracranial recordings made over the ventral and lateral occipito-temporal cortex (Allison, Ginter, et al., 1994; Allison, McCarthy, Nobre, Puce, & Belger, 1994; Allison, Puce, Spencer, & McCarthy, 1999). FMRI studies have now defined these areas more precisely in the middle part of the lateral fusiform gyrus ("fusiform face area", FFA; e.g., Kanwisher, McDermott, & Chun, 1997; Puce, Allison, Gore, & McCarthy, 1995), the lateral inferior occipital gyrus (subset of ventral LOC, "occipital face area", e.g., Gauthier et al., 2000), the posterior part of the superior temporal sulcus (pSTS, e.g., Puce, Allison, Bentin, Gore, & McCarthy, 1998) and the anterior section of the infero-temporal cortex (AiT, Rajimehr, Young, & Tootell, 2009) (for reviews see Fox, Iaria, & Barton, 2009; Haxby, Hoffman, & Gobbini, 2000). Recent studies have shown an even wider network of face-selective areas in humans, all along the ventral stream and the superior temporal sulcus (Pinsk et al., 2009; Tsao, Moeller, & Freiwald, 2008; Weiner & Grill-Spector, 2010). In comparison, besides body parts (Peelen & Downing, 2007; Weiner & Grill-Spector, 2010), there is no such known clustering of category-selective responses for other categories than faces. Considering this, it is likely that processes carried out within and between these face-preferential areas generate large current flows contributing to the particularly large N170 field potential recorded on the scalp for faces. Moreover, these areas are strongly right lateralized (Sergent et al., 1992), an observation that is in agreement with the largest N170 face effect over the right than the left hemisphere as found here and in previous studies (e.g., Bentin et al., 1996; Rossion, Joyce, Cottrell, & Tarr, 2003).

From a functional point-of-view, there is now a wide body of evidence indicating that the increase of N170 amplitude to faces is driven by high-level information, that is, information associated with the *perception* – in the sense of an interpretation – of the stimulus as a face (Rossion & Jacques, 2011). For instance, the N170 is larger to the exact same two-tone "Mooney" images, or Arcimboldo paintings, when they are presented at upright orientation – and thus generally perceived as faces – than when they are presented upside-down (George et al., 2005; Jeffreys, 1993; Rossion & Jacques, 2008). In other cases, whenever a stimulus contains enough information (either in the local elements, or in their global configuration, or both) to *be interpreted* as a face by the visual system, the N170 is large in amplitude (see Bentin & Golland, 2002; Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, 2002). This is true for face photographs obviously, but also for schematic faces, faces with features rearranged, inverted faces, faces cut in two halves, isolated eyes, faces with contrast inverted, faces without eyes, etc., all kinds of stimuli that are readily interpreted as being facelike (e.g., Bentin et al., 1996; Eimer, 1998; George et al., 1996; Itier & Taylor, 2002; Jacques & Rossion, 2010; Rossion et al., 1999; Sagiv & Bentin, 2001). However, when a transformation removes most of the diagnostic information to even perceive the stimulus as a face, the N170 is reduced in amplitude (e.g., an isolated nose or mouth, Bentin et al., 1996; superimposed random noise in frequency bands critical for face perception, Tanskanen et al., 2005; masking features with noise *and* breaking face configuration by inversion, Schneider, DeLong, & Busey, 2007). These observations all support the view that the N170 face effect reflects the perception of faces *qua* faces in high-level visual cortex.

Finally, in the present study, the N170 was not only larger, but it also peaked significantly earlier for faces than cars. This observation is not novel (e.g., Itier & Taylor, 2004a, 2004b), but unlike amplitude effects it is not always systematically found: there are several studies in which the N170 latency does not peak earlier for faces than other visual stimuli (e.g., Rossion et al., 2000). There are different – non-exclusive – possible accounts for this latency effect. First, if visual heterogeneity between items is larger for nonface objects than faces, for instance if cars present larger shape and color differences than faces, there could be an increase of latency jitter between trials for cars compared to faces, causing latency differences between categories in the averaged N170 response (Regan, 1989). However, this should be associated with a widening of the N170 component to nonface objects (see Rossion & Jacques, 2008), which was not observed here. Second, high-level representations might genuinely be activated earlier for face than objects, an account compatible with shorter mean onset latencies of infero-temporal cortex neurons in response to faces than nonface visual stimuli in the monkey brain (Kiani et al., 2005). It is also in agreement with observations that saccades towards faces are faster than to other objects (Crouzet et al., 2010; see also Cerf et al., 2008), that faces are detected more rapidly than other objects in visual scenes and search arrays (Hershler, Golan, Bentin, & Hochstein, 2010; Hershler & Hochstein, 2005) and that classifying the stimuli as meaningful objects was significantly faster for faces than cars in the present study (behavioral results). Third, contrary to the amplitude face effect, a faster N170 for faces might be driven partly by low-level visual cues. Indeed, the N170 latency was also longer for scrambled cars than scrambled faces (Fig. 4). Moreover, subtracting slight P1 latency differences between conditions cancelled out the interaction between category and shape on the N170 (-P1) latencies. Altogether, these observations suggest that low-level visual cues play a significant role in the shorter N170 latency for faces than cars (and possibly other visual objects). This observation does not contradict Kiani et al. (2005)'s findings, but suggest that the shortest onset latencies to activate face-selective cells may come from earlier inputs from lower-visual areas, so that low-level visual cues play a role in this latency difference observed in higher visual areas.

5. Conclusions

In many EEG (or MEG) studies of face stimulation, the visual evoked potentials P1 and N1 – the latter being called the N170 in the field of face processing – show similar response properties. Most notably, both components can be found being larger in response to faces than objects, and both to be larger and delayed in response to inverted than upright faces. We previously showed that the effects of face inversion on these components can be dissociated by means of a parametric manipulation of face orientation (Jacques & Rossion, 2007). The results of the present study point to an even more fundamental dissociation between the P1 and N170, showing that the P1 face-sensitivity is essentially a response to low-level visual cues of the stimuli, while the N170 face-sensitivity is driven by the perception of the stimulus as a face, independently of low-level visual cues such as amplitude spectrum and color. These findings allow apparent discrepancies between the human ability to saccade extremely rapidly towards faces and the early activation of high-level facial representations as shown by electrophysiological studies in the primate brain. They also put strong constraints on the interpretation of early (before 100 ms) face-sensitive effects in the human brain.

Acknowledgments

This research was supported by an ARC Grant 07/12-007 (Communauté Française de Belgique – Actions de Recherche Concertées) Bruno Rossion is supported by the Belgian National Fund for Scientific Research (Fonds de la Recherche Scientifique – FNRS). We thank Corentin Jacques for making the stimuli, and for his commentaries on a previous version of this paper, together with Guillaume Rousselet and an anonymous reviewer.

References

Ales, J. M., Yates, J. L., & Norcia, A. M. (2010). V1 is not uniquely identified by polarity reversals of responses to upper and lower visual field stimuli. *NeuroImage, 52*, 1401–1409.

Allison, T., Ginter, H., McCarthy, G., Nobre, A., Puce, A., Luby, M., et al. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology, 71*, 821–825.

Allison, T., McCarthy, G., Nobre, A., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex, 5*, 544–554.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex, 9*, 415–430.

Batty, M., & Taylor, MJ. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research, 17*, 613–620.

Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: The influence of perceptual experience on early visual processing of faces. *Cognition, 86*(1), 1–14.

Bentin, S., McCarthy, G., Perez, E., Puce, A., & Allison, T. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience, 8*, 551–565.

Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., & von Cramon, D. Y. (2002). Priming visual face-processing mechanisms: Electrophysiological evidence. *Psychological Science, 13*(2), 190–193.

Bötzel, K., & Grusser, O. J. (1989). Electric brain potentials-evoked by pictures of faces and non-faces – A search for face-specific eeg-potentials. *Experimental Brain Research, 77*, 349–360.

Bötzel, K., Schulze, S., & Stodieck, S. R. G. (1995). Scalp Topography And Analysis Of Intracranial Sources Of Face-Evoked Potentials. *Experimental Brain Research, 104*, 135–143.

Braeutigam, S., Bailey, A. J., & Swithenby, S. J. (2001). Task-dependent early latency (30–60 ms) visual processing of human faces and other objects. *NeuroReport, 12*, 1531–1536.

Cerf, M., Harel, J., EinhÃuser, W., & Koch, C. (2008). Predicting human gaze using low-level saliency combined with face detection. In J. C. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems*. Cambridge, MA: MIT Press. 20.

Chiu, Y. C., Esterman, M., Han, Y., Rosen, H., & Yantis, S. (2011). Decoding task-based attentional modulation during face categorization. *Journal of Cognitive Neuroscience, 23*, 1198–1204.

Clark, V. P., Fan, S., & Hillyard, S. A. (1995). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping, 2*, 170–187.

Cobb, W. A., & Dawson, G. D. (1960). The latency and form in man of the occipital potentials evoked by bright flashes. *Journal of Physiology – London, 152*, 108–121.

Crouzet, S., & Thorpe, S. J. (2010). Power spectrum cues underlying ultra-fast saccades towards faces. In *Poster presented at the Annual Meeting of the Vision Science Society Meeting*.

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. 16.

Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping, 15*, 95–111.

Edwards, R., Xiao, D., Keysers, C., Földiák, P., & Perrett, D. (2003). Color sensitivity of cells responsive to complex stimuli in the temporal cortex. *Journal of Neurophysiology, 90*, 1245–1256.

Eimer, M. (1998). Does the face-specific N170 component reflect the activity of a specialized eye processor? *NeuroReport, 9*, 2945–2948.

Eimer, M. (2000). Effects of face inversion on the structural encoding and recognition of faces – Evidence from event-related brain potentials. *Cognitive Brain Research, 10*, 145–158.

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

Fox, C. J., Iaria, G., & Barton, J. (2009). Defining the face-processing network: Optimization of the functional localizer in fMRI.. *Human Brain Mapping, 30*, 1637–1651.

Freeman, J. B., Ambady, N., & Holcomb, P. J. (2010). The face-sensitive N170 encodes social category information. *NeuroReport, 21*, 24–28.

Gamond, L., George, N., Lemaréchal, J. D., Hugueville, L., Adam, C., & Tallon-Baudry, C. (2011). Early influence of prior experience on face perception. *NeuroImage, 54*, 1415–1426.

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

George, N., Evans, J., Fiori, N., Davidoff, J., & Renault, B. (1996). Brain events related to normal and moderately scrambled faces. *Cognitive Brain Research, 4*, 65–76.

George, N., Jemel, B., Fiori, N., Chaby, L., & Renault, B. (2005). Electrophysiological correlates of facial decision: Insights from upright and upside-down Mooney-face perception. *Cognitive Brain Research, 24*, 663–673.

George, N., Jemel, B., Fiori, N., & Renault, B. (1997). Face and shape repetition effects in humans: A spatio-temporal ERP study. *NeuroReport, 8*, 1417–1423.

Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Cognitive Brain Research, 16*, 416–424.

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P. G., & Rossion, B. (2005). Diagnostic colors contribute to the early stages of scene categorization: Behavioral and neurophysiological evidence. *Visual Cognition, 12*, 878–892.

Guyader, N., Chauvin, A., Peyrin, C., Hérault, J., & Marendaz, C. (2004). Image phase or amplitude? Rapid scene categorization is an amplitude-based process. *CR Biologies, 327*, 313–318.

Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex, 10*, 69–81.

Halit, H., de Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *NeuroReport, 11*, 1871–1875.

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

Herrmann, M. J., Ehlis, A. C., Muehlberger, A., & Fallgatter, A. J. (2005). Source localization of early stages of face processing. *Brain Topography, 18*, 77–85.

Hershler, O., Golan, T., Bentin, S., & Hochstein, S. (2010). The wide window of face detection. *Journal of Vision, 10*(20:21), 1–14.

Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop-out effect for faces. *Vision Research, 45*, 1707–1724.

Hershler, O., & Hochstein, S. (2006). With a careful look: Still no low-level confound to face pop-out. *Vision Research, 46*, 3028–3033.

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

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *NeuroImage, 15*, 353–372.

Itier, R. J., & Taylor, M. J. (2004a). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex, 14*, 132–142.

Itier, R. J., & Taylor, M. J. (2004b). Source analysis of the N170 to faces and objects. *NeuroReport, 15*, 1261–1265.

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85*, 616–626.

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, & Behavioral Neuroscience, 5*, 21–36.

Jacques, C., & Rossion, B. (2004). Concurrent processing reveals competition between visual representations of faces. *NeuroReport, 15*, 2417–2421.

Jacques, C., & Rossion, B. (2007). Early electrophysiological responses to multiple face orientations correlate with individual discrimination performance in humans. *NeuroImage, 36*, 863–876.

Jacques, C., & Rossion, B. (2009a). Spatio-temporal dissociation between low- and high-level effects of stimulus inversion on early face-sensitive electrophysiological responses. *Journal of Vision, 9*(8), 533.

Jacques, C., & Rossion, B. (2010). Misaligning face halves increases and delays the N170 specifically for upright faces: Implications for the nature of early face representations. *Brain Research, 1318*, 96–109.

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research, 78*, 193–202.

Jeffreys, D. A. (1993). The influence of stimulus orientation on the vertex positive scalp potential evoked by faces. *Experimental Brain Research, 96*, 163–172.

Jeffreys, D. A., & Axford, J. G. (1972). Source locations of pattern-specific components of human visual evoked potentials. I. Component of striate cortical origin. *Experimental Brain Research, 16*, 1–21.

Jemel, B., Coutya, J., Langer, C., & Roy, S. (2009). From upright to upside-down presentation: A spatio-temporal ERP study of the parametric effect of rotation on face and house processing. *BMC Neuroscience, 10*, 100.

Jemel, B., Schuller, A. M., Cheref-Khan, Y., Goffaux, V., Crommelinck, M., & Bruyer, R. (2003). Stepwise emergence of the face-sensitive N170 event-related potential component. *NeuroReport, 14*, 2035–2039.

Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology, 116*, 2613–2631.

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

Kaping, D., Tzvetanov, T., & Treue, T. (2007). Adaptation to statistical properties of visual scenes biases rapid categorization. *Visual Cognition, 15*, 12–19.

Keil, M. S. (2008). Does face image statistics predict a preferred spatial frequency for human face processing. *Proceedings of the Royal Society B: Biological Sciences, 275*, 2095–2100.

Kiani, R., Esteky, H., & Tanaka, K. (2005). Differences in onset latency of macaque inferotemporal neural responses to primate and non-primate faces. *Journal of Neurophysiology, 94*, 1587–1596.

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Visual Research, 46*, 1762–1776.

Kubota, J. T., & Ito, T. A. (2007). Multiple cues in social perception: The time course of processing race and facial expression. *Journal of Experimental Social Psychology, 43*, 738–752.

Landis, T., Regard, M., Bliestle, A., & Kleihues, P. (1988). Prosopagnosia and agnosia from noncanonical views. An autopsied case. *Brain, 11*, 1287–1297.

Latinus, M., & Taylor, M. J. (2005). Holistic processing of faces: Learning effects with Mooney faces. *Journal of Cognitive Neuroscience, 17*, 1316–1327.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and Clinical Neurophysiology, 48*, 609–621.

Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., & Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neuroscience Letters, 253*, 147–150.

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: Fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron, 62*, 281–290.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience, 5*, 910–916.

Liu, L., & Ioannides, A. A. (2010). Emotion separation is completed early and it depends on visual field presentation. *PLoS One, 5*(3), e9790.

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92*, 8135–8139.

Morel, S., Ponz, A., Mercier, M., Vuilleumier, P., & George, N. (2009). EEG-MEG evidence for early differential repetition effects for fearful, happy and neutral faces. *Brain Research, 13*(1254), 84–98.

Mouchetant-Rostaing, Y., Giard, M. H., Bentin, S., Aguera, P. E., & Pernier, J. (2000a). Neurophysiological correlates of face gender processing in humans. *European Journal of Neuroscience, 12*, 303–310.

Mouchetant-Rostaing, Y., Giard, M. H., Delpuech, C., Echallier, J. F., & Pernier, J. (2000b). Early signs of visual categorization for biological and non-biological stimuli in humans. *NeuroReport, 11*, 2521–2525.

Nasanen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Visual Research, 39*, 3824–3833.

Nestor, A., & Tarr, M. J. (2008). Gender recognition of human faces using color. *Psychological Science, 19*, 1242–1246.

Okazaki, Y., Abrahamyan, A., Stevens, C. J., & Ioannides, A. A. (2008). The timing of face selectivity and attentional modulation in visual processing. *Neuroscience, 152*, 1130–1144.

Parkin, A. J., & Williamson, P. (1987). Cerebral lateralization at different stages of facial processing. *Cortex, 23*, 99–110.

Peelen, M. V., & Downing, PE. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience, 8*, 636–648.

Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., et al. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *Journal of Neurophysiology, 101*, 2581–2600.

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

Pizzagalli, D., Regard, M., & Lehmann, D. (1999). Rapid emotional face processing in the human right and left brain hemispheres: An ERP study. *NeuroReport, 10*, 2691–2698.

Pourtois, G., Dan, E. S., Grandjean, D., Sander, D., & Vuilleumier, P. (2005). Enhanced extrastriate visual response to bandpass spatial frequency filtered fearful faces: Time course and topographic evoked-potentials mapping. *Human Brain Mapping, 26*, 65–79.

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

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

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

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

Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., et al. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology, 50*, 173–189.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *NeuroReport, 11*, 69–74.

Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage, 39*, 1959–1979.

Rossion, B., & Jacques, C. (2011). The N170: Understanding the time-course of face perception in the human brain. In S. Luck & E. Kappenman (Eds.), *The Oxford handbook of ERP components*. Berlin: Oxford University Press.

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *NeuroImage, 20*, 1609–1624.

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2005). Spatial scaling factors explain eccentricity effects on face ERPs. *Journal of Vision, 5*(10), 755–763.

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2007). Single-trial EEG dynamics of object and face visual processing. *Neuroimage, 36*, 843–862.

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *Journal of Vision, 8*, 12–13.

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. *Journal of Vision, 3*, 440–455.

Rousselet, G. A., Pernet, C. R., Bennett, P. J., & Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neuroscience, 9*, 98.

Sadeh, B., Podlipsky, I., Zhdanov, A., & Yovel, G. (2010). Event-related potential and functional MRI measures of face-selectivity are highly correlated: A simultaneous ERP-fMRI investigation. *Human Brain Mapping, 31*, 1490–1501.

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

Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part-based processes. *Journal of Cognitive Neuroscience, 13*, 937–951.

Sams, M., Hietanen, J. K., Hari, R., Ilmoniemi, R. J., & Lounasmaa, O. V. (1997). Face-specific responses from the human inferior occipito-temporal cortex. *Neuroscience, 77*, 49–55.

Schneider, B. L., DeLong, J. E., & Busey, T. A. (2007). Added noise affects the neural correlates of upright and inverted faces differently. *Journal of Vision, 7*, 4.

Scholte, H. S., Ghebreab, S., Waldorp, L., Smeulders, A. W., & Lamme, V. A. (2009). Brain responses strongly correlate with Weibull image statistics when processing natural images. *Journal of Vision, 9*(4), 1–15.

Schroeder, C. E., Molholm, S., Lakatos, P., Ritter, W., & Foxe, J. J. (2004). Human-Simian correspondence in the early cortical processing of multisensory cues. *Cognitive Processing, 5*, 140–151.

Seeck, M., Michel, C. M., Mainwaring, N., Cosgrove, R., Blume, H., Ives, J., et al. (1997). Evidence for rapid face recognition from human scalp and intracranial electrodes. *NeuroReport, 8*, 2749–2754.

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

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Science, 13*, 403–409.

Tanskanen, T., Nasanen, R., Montez, T., Paallysaho, J., & Hari, R. (2005). Face recognition and cortical responses show similar sensitivity to noise spatial frequency. *Cerebral Cortex, 15*, 526–534.

Thorpe, S. J., & Fabre-Thorpe, M. (2002). Fast visual processing and its implications. In M. Arbib (Ed.), *The handbook of brain theory and neural networks* (2nd ed.. Cambridge MA: MIT Press.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*, 520–522.

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

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *PNAS, 105*, 19514–19519.

VanRullen, R. (2006). On second glance: Still no high-level pop-out effect for faces. *Vision Research, 46*, 3017–3027.

Vlamings, P., Goffaux, V., & Kemner, C. (2009). Is the early modulation of brain activity by fearful facial expressions primarily mediated by coarse low spatial frequency information? *Journal of Vision, 9*(5), 12.1–13.

Watanabe, S., Kakigi, R., & Puce, A. (2003). The spatiotemporal dynamics of the face inversion effect: A magneto- and electro-encephalographic study. *Neuroscience, 116*, 879–895.

Weiner, K. S., & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage, 15*, 1559–1573.