Long-term Expertise with Artificial Objects Increases Visual Competition with Early Face Categorization Processes

Bruno Rossion[1], Daniel Collins[2], Valérie Goffaux[1], and Tim Curran[2]

Abstract

■ The degree of commonality between the perceptual mechanisms involved in processing faces and objects of expertise is intensely debated. To clarify this issue, we recorded occipito-temporal event-related potentials in response to faces when concurrently processing visual objects of expertise. In car experts fixating pictures of cars, we observed a large decrease of an evoked potential elicited by face stimuli between 130 and 200 msec, the N170. This sensory suppression was much lower when the car and face stimuli were separated by a 200-msec blank interval. With and without this delay, there was a strong correlation between the face-evoked N170 amplitude decrease and the subject's level of car expertise as measured in an independent behavioral task. Together, these results show that neural representations of faces and nonface objects in a domain of expertise compete for visual processes in the occipito-temporal cortex as early as 130–200 msec following stimulus onset. ■

INTRODUCTION

Human adults are extremely efficient in recognizing individual faces as a result of a combination of innate biological constraints to the visual system and of extensive and prolonged experience with the category of faces throughout development (Carey, 1992). Yet, even when the face processing system is fully matured, recognizing members of a natural or an artificial nonface object category may rely on face-related mechanisms following visual expertise with the object category. The evidence supporting this claim comes from behavioral, neuroimaging, and electrophysiological measures in humans (Tarr & Cheng, 2003).

Behavioral evidence has suggested that processing of faces and objects of expertise shows common characteristics such as configural/holistic processing (Gauthier & Tarr, 1997) and automatic subordinate-level processing (Tanaka & Taylor, 1991), but these commonalities do not necessarily imply common underlying neural processes. Neuroimaging evidence has been considered to address more directly the hypothesis of a common neural mechanism. In the lateral part of the middle fusiform gyrus, an area shown by neuroimaging studies to respond more to faces than other object (e.g., Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997; Sergent, Otha, & MacDonald,

1992), experts at discriminating novel objects and real-world categories, such as birds and cars, present a larger activation to members of these categories than novices (Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). Taking advantage of their excellent temporal resolution, human scalp electrophysiological studies have found results suggesting that the perceptual categorization mechanisms for faces in the occipito-temporal cortex, as indexed by the N170 event-related potential (ERP) component (Bentin, Allison, Puce, Perez, & McCarthy, 1996), are also recruited when processing objects of expertise (Busey & Vanderkolk, 2005; Rossion, Gauthier, Goffaux, Tarr, & Crommelinck, 2002; Tanaka & Curran, 2001). However, given the limited spatial resolution of functional magnetic resonance imaging (fMRI), and especially ERPs, it remains possible that functionally separate yet anatomically nearby neural networks process faces separately from other objects of expertise.

Recent studies have attempted to obtain evidence for common processing stages by measuring the extent to which concurrent expert processing interferes with face recognition mechanisms (Rossion, Kung, & Tarr, 2004; Gauthier, Skudlarski, Gore, & Anderson, 2003). In a working memory task that alternated between images of cars and faces, it was found that experts processed cars more holistically than novices, and that holistic processing of cars interfered with the N170 ERP to faces for experts more than for novices (Gauthier et al., 2003). Similarly, during concurrent visual stimulation in ERPs,

---

[1]Université catholique de Louvain, Belgium, [2]University of Colorado at Boulder

early face-related visual processes are substantially suppressed when subjects are fixating novel members of a nonface object category on which they had been trained (Rossion et al., 2004).

Although these previous ERP interference experiments provide evidence for shared perceptual mechanisms contributing to face and expert-object processing, fundamental questions remain. Gauthier et al. (2003) found evidence for a shared holistic processing mechanism, but did not find overall differences between the N170 to faces for car experts versus novices. Rossion et al. (2004) found that training with artificial objects did reduce strongly the size of the N170 to concurrently processed faces, but the relevance of these limited training studies to real-world expertise is debatable. The present research addressed these issues with three purposes in mind. First, we tested whether long-term visual expertise developed ''naturally'' with a category of familiar man-made objects—cars—would compete with perceptual mechanisms for faces. Second, the correlation between behavioral expertise and neuronal competition was investigated. Third, we tested whether the competition critically depends on the concurrent stimulation of the two categories, faces and objects of expertise, or if it can be observed when the two stimuli are separated by a short blank interval. To address these issues, we measured the amplitude of the N170 ERP to faces in 19 car novices and 19 experts. Subjects were first presented either with a central fixation cross, a car picture, or a scrambled version of a car stimulus. In Experiment 1 (concurrent stimulation), the central stimulus remained on the center of the screen while the face stimulus of interest appeared in the right or in the left visual field. In Experiment 2, the central stimulus disappeared 200 msec before the presentation of the face. The order of experiments was counterbalanced across subjects.

METHODS

Subjects

Twenty self-reported car experts and 20 novices with normal or corrected vision gave informed consent and participated in this study, approved by the Human Research Committee at University of Colorado at Boulder, for payment or partial course credit. A novice subject was excluded for poor electroencephalogram (EEG) signal-to-noise ratio, and the last expert subject recruited was removed to have an equal number of subjects ($n = 19$) in the two groups for analyses. All subjects were men.

Car Expertise Test

Subjects' car expertise was tested like in previous work (Gauthier et al., 2003), yielding a quantitative estimate of their ability relative to their performance with birds (used

here as a baseline for novice-level performance). Subjects matched sequentially presented (256×256) gray-scale images of cars and birds on the basis of their model or species (224 trials.) The first image was presented for 1000 msec followed by a mask for 500 msec, and then the second image appeared and remained till the subject made a response or 5000 msec had passed. Matching stimuli were not physically identical but were different exemplars of the same bird species or the same make/ model of car from different years. An index of car expertise (delta d') was computed as the d' difference between the car and bird conditions. Self-reported car experts yielded a $\Delta d'$ of 1.46, which was significantly larger than the $\Delta d'$ of car novices (0.58) ($p < .0001$). Although individual subjects may certainly vary in their face recognition abilities, face expertise was not measured, as we assumed that car experts and novices, as a group, would be matched on their ability to process faces.

EEG Experiments

Stimuli

Faces. Thirty color photographs of full front faces of undergraduate students, 15 men and 15 women without glasses, facial hair, and make-up, were used in this study. They were extracted from a whole set of faces used previously in several studies (e.g., 10, 24). All face photographs were edited in Adobe Photoshop 4.0 to remove backgrounds and haircut, and everything below the chin. They were all of neutral facial expression. On average, the size of each face photograph was 4 cm wide (around 2.29° at 100 cm from the monitor) and 5 cm height (2.86°).

Cars and scrambled cars. A total of 30 cars were used. They were mostly European and Japanese cars, also used in previous behavioral and ERP studies. Car pictures subtended a size of about 4.01° of visual angle. A ''scrambled'' version of each car picture was created by spatially quantizing (or pixellating) the car picture in approximately 6 × 3 pixels (see Figure 1). Spatial quantization is a way to remove relevant fine-scale information—similar to applying a low-pass filter—which can be done up to a point where the stimulus cannot be recognized. Low-level properties, such as overall size, luminance, and color parameters of the whole image, are preserved, making this stimulus a good control compared to the original car pictures, even though spatial quantization introduces high-spatial frequency components created by the edges and corners of the block (Bachmann & Kahusk, 1997; Sergent, 1986).

Procedure

Experiment 1. A trial included the presentation of an object (car, scrambled car) or a fixation cross for

Figure 1. Time line of the stimulation events in Experiment 1. The face picture was presented either on the left or the right side of the central stimulus (car, scrambled car, fixation). ISI = interstimulus interval.
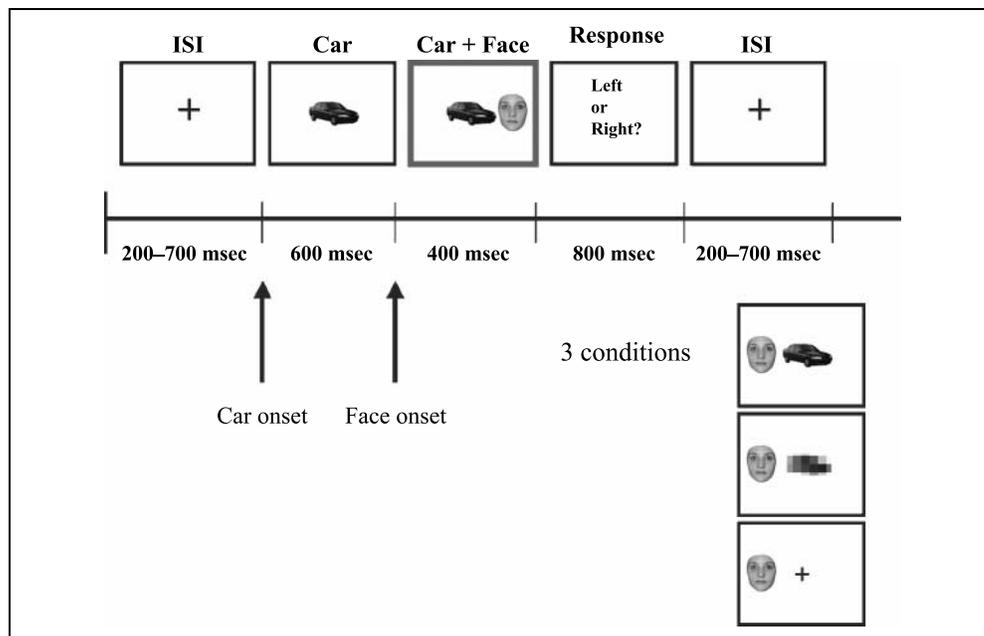
1000 msec in the center of the screen. Six hundred milliseconds, on average (randomized between 500 and 700 msec), after the onset of the first picture, a face photograph was presented for 400 msec, either on the left or on the right of the object (Figure 1). The center of the face stimulus appeared at 7.5 cm from the center of the screen (4.29°). The offset of the two stimuli was simultaneous, and a blank screen was then presented for ~1250 msec (randomized delay between 1000 and 1500 msec). There were 120 trials of each condition (car + face; scrambled car + face; fixation cross + face). The face appeared on the each side (left or right) in half of the trials, giving 60 trials/condition for ERP averages. The order of all trials was fully randomized so that subjects could not anticipate whether a car, a scrambled car, or a fixation cross would be presented at the next trial, and whether the face would appear on the left or on the right visual field. The subject's task was to press a left key or a right key if the face appeared on the left or on the right of the object. Subjects were instructed to keep their eye gaze and attention on the center of the screen and respond as accurately and as fast as possible to the presentation of the face. This irrelevant task was used to avoid any attentional bias in favor of one of the object categories throughout the experiment while keeping the subject's attention and motivation at a good level during the whole experiment.

Experiment 2. The procedure was identical to Experiment 1, except that a blank screen of 200 msec replaced the first stimulus (car, scrambled car, cross) before the appearance of the second stimulus. Half of the subjects performed the Experiment 1 first, and the other half started with Experiment 2.

EEG Recording

Scalp electrical activity (EEG) was recorded with a 128-channel geodesic Sensor Net (Electrical Geodesics, Eugene, OR; Tucker, 1993) connected to an AC-coupled, 128-channel, high-input impedance amplifier (200 MW, Net Amps, Electrical Geodesics). Amplified analog voltages (0.1–100 Hz bandpass) were digitized at 250 Hz. Recorded voltages were initially referenced to a vertex channel. Voltages were re-referenced off-line into a common average reference. Individual sensors were adjusted until impedances were less than 50 kΩ. Vertical and horizontal eye movements (electrooculogram [EOG]) were recorded by electrodes placed on the external canthi of the eyes (horizontal movements) in the inferior and superior areas of the ocular orbit for vertical eye movements.

EEG/ERP Analyses

Electroencephalogram data were analyzed using Eeprobe 3.0 (ANT) running on Red Hat Linux 7.0. After filtering of the EEG with a 1–30 Hz bandpass filter, EEG and EOG artifacts were removed using a [−35; +35 μV] deviation over 200 msec intervals on all electrodes. A 1-Hz high pass Hanning filter (201 points) was used to reduce the effect of stimulus anticipation on the EEG preceding the presentation of the face photograph. In case of too many blink artifacts, they were corrected by a subtraction of vertical EOG (VEOG) propagation factors based on PCA-transformed EOG components. Epochs beginning 200 msec prior to the face stimulus onset and continuing for 800 msec were extracted, corrected from baseline deviations from 0 using 200 msec prestimulus window, and averaged for each condition separately. Averaged

ERPs were re-referenced using a common average reference. Only correct response trials were averaged.

Statistical Analysis

The N170 amplitude was measured at four occipito-temporal locations around T6/T5 (#97/#58), where it was the largest in response to faces, in the two hemispheres (electrode numbers = RH: #102, #101, #100, #97; LH: #50, #57, #63, #58). The N170 peaked, on average, at about 160 msec when the faces were presented in the contralateral visual field, and was delayed of about 20 msec (~180 msec) for ipsilateral presentations. N170 latency was similar whether faces were preceded by a cross, a car, or a scrambled car (Table 1). To account for latency differences between hemispheres, mean amplitude was averaged within 20 msec temporal windows centered on the mean latency of the electrode where the N170 was maximal (T6/T5) by condition. The analysis of variance (ANOVA) included the factors group (self-reported expert vs. novices), condition (car, scrambled, fixation cross), visual field (left/right), hemisphere (left, right), and electrode site (four sites). Correlation analyses were performed between a behavioral index of expertise (similarly to previous studies, $\Delta d' = d'$ for matching cars $- d'$ for matching cars; see, e.g., Gauthier et al., 2003) and the amount of reduction of N170 amplitude to faces when the central stimulus was a car as opposed to either the scrambled car condition or the fixation cross.

RESULTS

Experiment 1: Concurrent Presentation

Behavioral Results

During the EEG recording, participants' level of attention was monitored by requiring identification of the location of the face stimulus, appearing in the right or left visual field, by pressing one of two corresponding keys with the right hand. They were near ceiling performing this task (between 97.8% and 99.3%) in all conditions (3 sessions × 2 visual fields × 3 preceding cues). Mean response times (RTs) in all conditions were extremely fast and similar, means ranging between 286 and 304 msec. There was a main effect of presentation on RTs, faces presented in the right visual field being detected faster than faces presented in the left visual field [$F(1,36) = 27.4$, $p < .001$]. All other effects were not significant, although there were trends for experts to be faster than novices [$F(1,36) = 3.2$, $p = .08$] and an overall slight slowing down for the pixellated car condition [$F(2,72) = 2.98$, $p = .06$]. Most importantly, there was no interaction between the groups (experts vs. novices) and the conditions [$F(2,72) = 0.2$ $p = .8$].

Electrophysiological Results

On average, the N170 in response to lateralized faces took place at about 160 msec when the face stimulus was presented in the contralateral hemisphere, and was delayed of about 20 msec (~180 msec) in the ipsilateral hemisphere (Terasaki & Okazaki, 2002). The N170 latency did not differ, whether the face stimulus was preceded by cars, scrambled cars, or a fixation cross (Figure 2; Table 1).

There was a highly significant Group × Condition interaction [$F(2,72) = 5.69$, $p < .005$], which was due to the substantial reduction of the N170 mean amplitude in experts as compared to novices when subjects were viewing cars ($p < .05$), but not when viewing scrambled cars ($p = .43$) or a fixation cross ($p = .37$). Thus, the N170 in response to the lateralized face stimuli undergoes a major decrease in amplitude when presented in the context of a central car picture, significant only in experts (Figures 2–4; Table 1).

To obtain a quantitative measure of expertise, each subject completed a separate sequential matching task

Table 1. Latency and Amplitude Values of the N170 in Response to Faces for the Two Groups of Subjects in the Three Experimental Conditions for Experiment 1

Context	Novices						Experts					
	Car		Scrambled Car		Fixation Cross		Car		Scrambled Car		Fixation Cross	
Hemisphere	L	R	L	R	L	R	L	R	L	R	L	R
(A) Left Visual Field Presentation												
Latencies (msec)	192	160	184	156	184	156	192	148	180	152	180	148
Amplitudes (μV)	**−2.13**	**−4.22**	−3.31	−5.15	−2.68	−4.88	**−0.99**	**−3.08**	−2.68	−5.73	−2.57	−4.93
(B) Right Visual Field Presentation												
Latencies (msec)	156	184	152	184	156	180	152	184	152	180	152	180
Amplitudes(μV)	**−3.91**	**−3.32**	−5.16	−4.15	−4.58	−3.92	**−3.38**	**−2.32**	−4.59	−4.46	−3.38	−3.78

Note the massive reduction of N170 amplitude for experts in the ''car context'' condition (in **bold**).
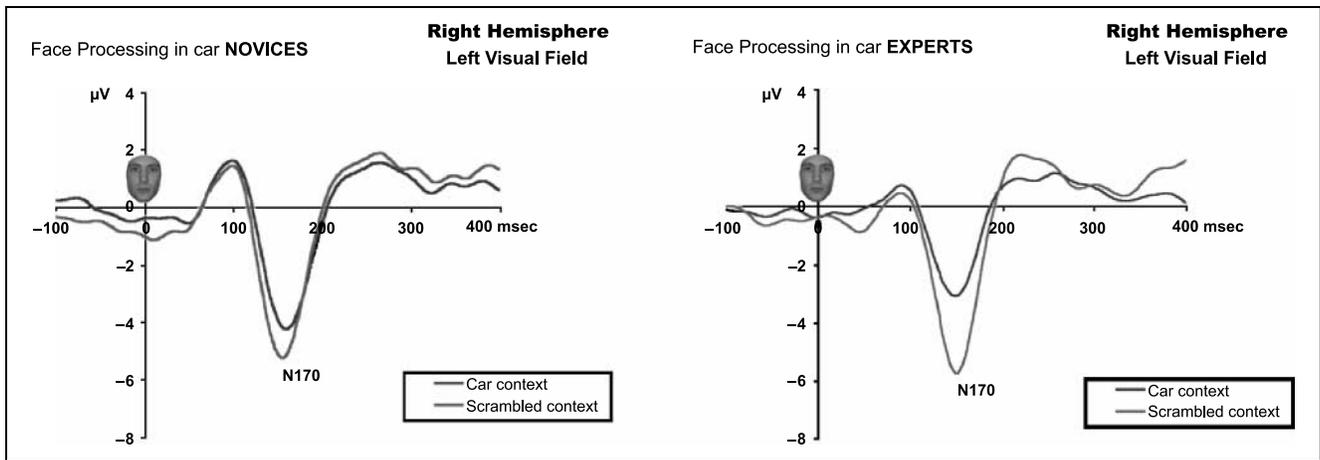
Figure 2. ERPs (right hemisphere, channel T6) in response to faces presented in the left visual field, in novices and experts. Note the massive reduction of the N170 in response to faces in car experts, when a car picture is presented (car context), relative to a scrambled car picture. Note that earlier (i.e., P1), differences between groups do not interact with the conditions of interest (cars vs. scrambled cars).

in which pairs of cars where presented for same/different judgments (Gauthier et al., 2003). Same trials included cars with the same make and model that differed in production year, gray-scale level, and viewpoint. An index of car expertise was computed as the performance difference between this car-matching task and a control task of matching bird species. We computed the correlation between the index of car expertise and the difference of the N170 amplitude evoked by faces in the context of processing cars versus scrambled cars or fixation. In the right hemisphere, there were highly significant correlations between the two parameters for contralateral presentation (Cars vs. Scrambled: $r = .46$, $p < .003$; and Cars vs. Fixation: $.54, p < .0004$; Figure 5;

Table 2). In the right hemisphere, there were also significant correlations for the presentation in the ipsilateral (right) visual field ($r = .44$, $p < .006$ for Cars–Scrambled; $r = .23$, $p > .10$ for Cars–Fixation).

Experiment 2: Delayed Presentation

Behavioral Results

As in Experiment 1, subjects were almost at ceiling performing the lateralized target detection task (between 96.8% and 98.9%) in all conditions (3 sessions × 2 visual fields × 3 preceding cues). Mean RTs in all conditions were extremely fast and similar, means ranging
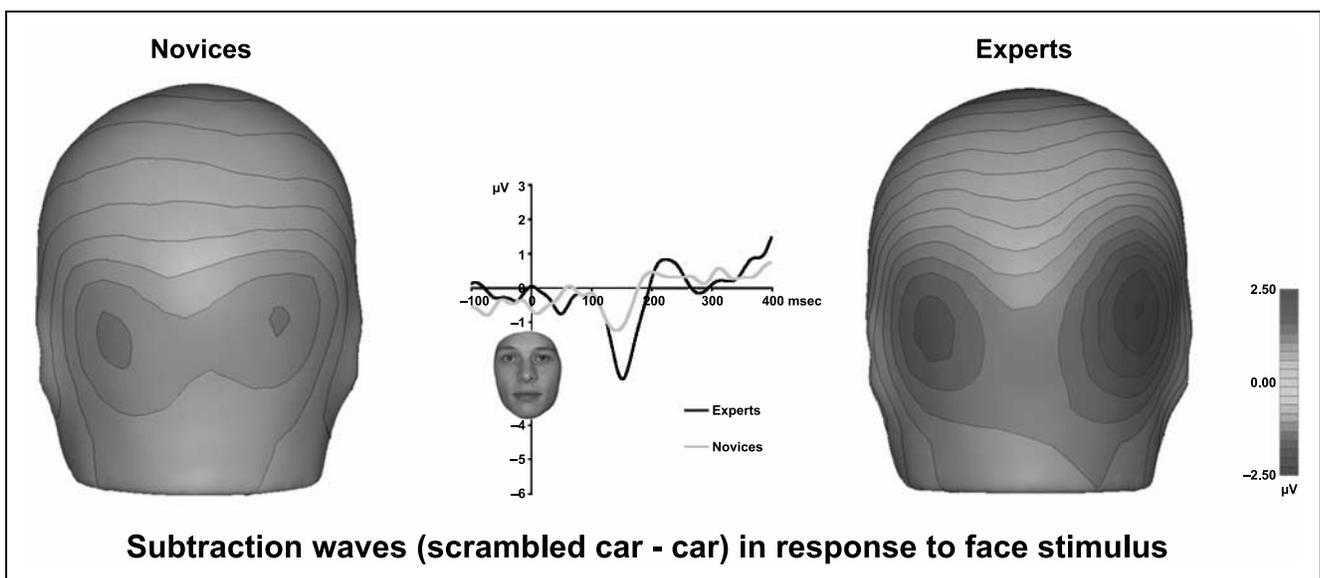

Figure 3. Topographical maps of the subtraction waveforms illustrating the car expertise effect on the face N170. The response to faces when a car picture is present is subtracted from the response to faces in the context of a scrambled car. The larger amplitude difference in experts is reflected as negative.
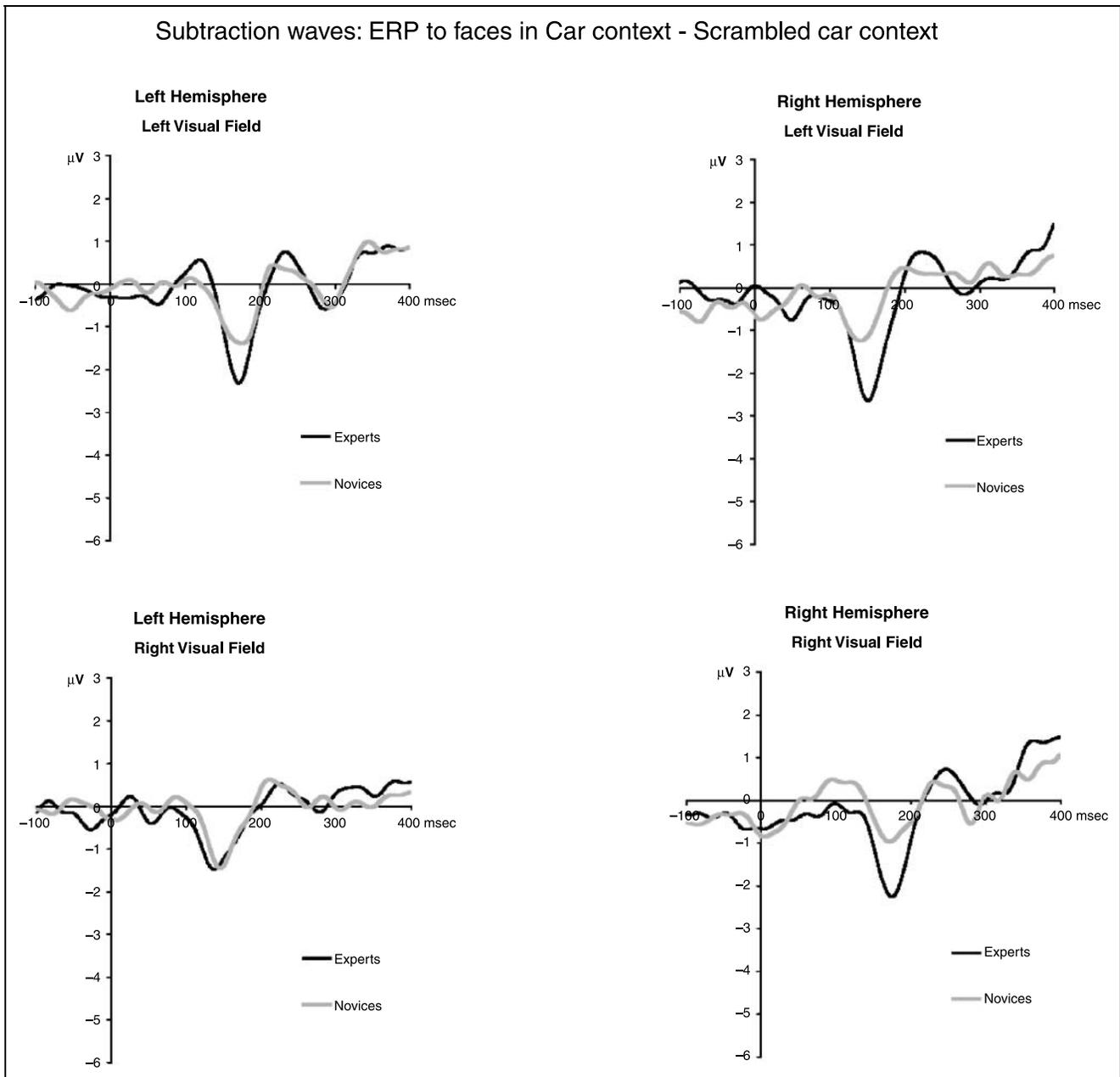
Figure 4. Subtraction waveforms for all hemifield presentations in the right and left hemispheres (T6 and T5 channels). Note that the larger decrease of N170 amplitude in experts is observed for contralateral and ipsilateral presentations in the right hemisphere, but only for contralateral presentations in the left hemisphere.

between 277 and 292 msec. There was a main effect of side of presentation on RTs, faces presented in the right visual field being detected faster than faces presented in the left visual field [$F(1,36) = 10.76, p < .0023$]. All other effects were not significant (all $ps > .23$).

Electrophysiological Results

Experiment 2 was identical to Experiment 1, except that the offset of the cue (car, scrambled car, fixation) was 200 msec before the onset of the face stimulus. As for Experiment 1, the N170 in response to lateralized faces took place at about 160 msec when the face stimulus was presented in the contralateral hemisphere, and was delayed of about 20 msec (~180 msec) in the ipsilateral hemisphere. The N170 latency did not differ whether preceded by cars, scrambled cars, or a fixation cross (Figure 6; Table 3).

The amplitude of the N170 to faces appeared decreased when cars were presented as a cue, especially in car experts, but the difference was much smaller than in Experiment 1, even though the same stimuli and subjects were tested (Figure 6). The Group × Condition interaction failed to reach significance [$F(2,72) = 1.86, p = .16$]. There was a significant interaction between
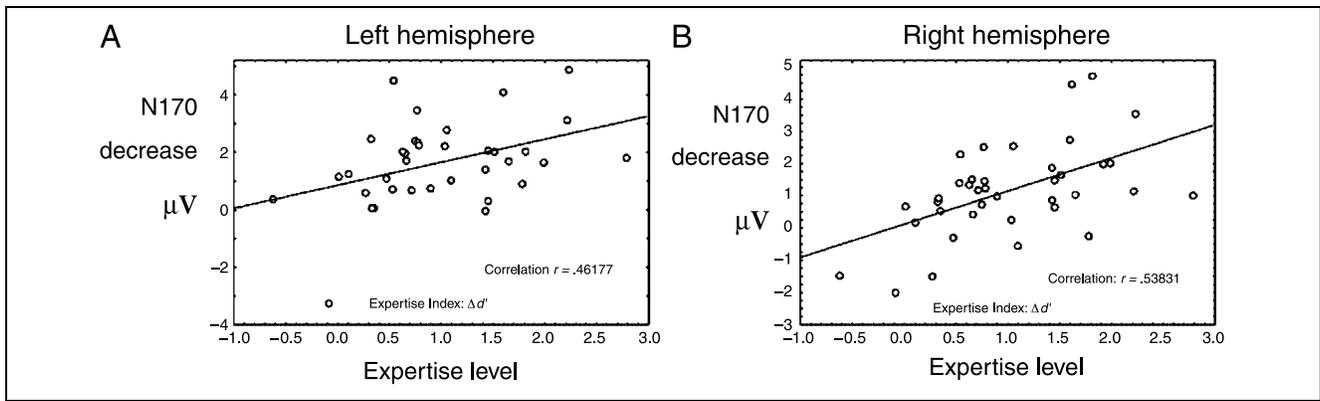
Figure 5. Correlation measures between the behavioral index of expertise, and the decrease of N170 amplitude to faces competing with car stimuli. (A) Right visual field stimulation, left hemisphere recording; (B) Left visual field stimulation, right hemisphere recording.

group, condition, and hemisphere [$F(2,72) = 3.89$, $p = .024$]. In the left hemisphere, there was no interaction between group and condition ($p = .75$), but the interaction was significant in the right hemisphere ($p = .02$). However, there was no difference between novices and experts in any of the conditions, whether the cue was a car picture ($p = .19$), a scrambled car picture ($p = .33$), or a fixation cross ($p = .8$).

Finally, in order to directly compare Experiments 1 and 2, we ran a global ANOVA analyses with the two experiments as factor. There was no significant Experiment × Condition × Expertise effect [$F(2,72) = 1.08$, $p = .345$]. However, there was an interaction between experiment, condition, expertise, and hemisphere [$F(2,72) = 3.00$, $p = .05$]. This reflected the fact that, in the right hemisphere, there were differences between experts and novices when subjects were viewing cars (i.e., face-N170 amplitude reduction in experts) in Experiment 1, whereas no such difference was found in Experiment 2, as described in the analyses above.

To sum up, although the N170 in response to the lateralized face stimuli appears to undergo a larger decrease in amplitude when presented in the context

Table 2. Experiment 1

Correlation Measures	Cars–Scrambled	Cars–Fixation
RH–LVF presentation	**$r = .46, p < .003$**	**$r = .54, p < .0004$**
RH–RVF presentation	**$r = .44, p < .006$**	$r = .23, p = .16$
LH–RVF presentation	$r = .2, p = .22$	$r = .18, p = .27$
LH–LVF presentation	$r = .16, p = .33$	$r = .24, p = .15$

Correlation measures between the behavioral index of expertise and the decrease of N170 amplitude to faces competing with car stimuli, relative to control conditions (scrambled car and fixation cross).
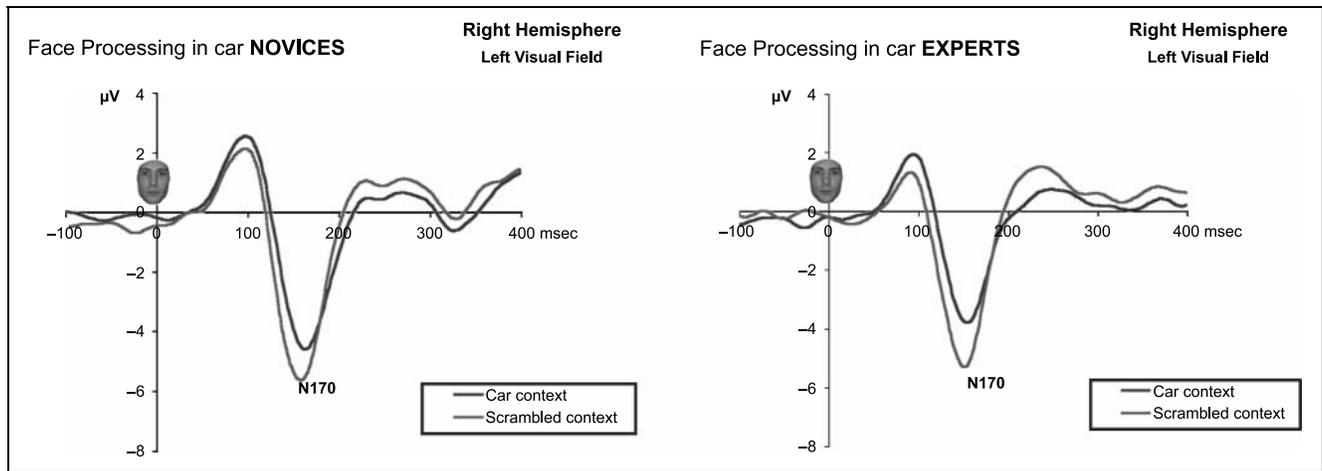
of a central car picture than control stimuli in experts (Figure 6; Table 3), the effect was not significant with 200 msec intervened between the car and the face. However, correlation analyses on the basis of the behavioral measure of expertise showed significant relationships with the amount of face-N170 reduction when presented in the context of processing a car picture relative to scrambled and fixation conditions (Table 4).

DISCUSSION

Our findings can be summarized in three points. First, long-term visual experience developed at discriminating and recognizing individual members of a category of familiar man-made objects—cars—leads to a decrease of the early (~130–200 msec) neural response to faces during conditions of concurrent visual presentation. Second, the suppression effect is more pronounced in the right hemisphere (where it reaches more than 2 µV, between 20% and 40% of the EEG signal), and correlates significantly with a behavioral index of visual expertise measured independently. Third, when a 200-msec blank interval is inserted between the stimulation of cars and faces, the evoked activity to the latter stimuli has substantially recovered from suppression even though the correlation between the neural and behavioral measures remains significant. The most straightforward interpretation of these observations is that when perceiving members of a category of man-made objects such as cars, experts rely on visual processes in the occipito-temporal cortex that compete with those involved in processing faces, leading to a suppression of these perceptual face processes.

Neurophysiological Mechanisms of the N170 Competition Effect

When two visual stimuli are present at the same time within a neuron's receptive field, the response of the neuron is a weighted average of the responses to the individual stimuli when presented alone (Reynolds,

Figure 6. ERPs (right hemisphere, channel T6) in response to faces presented in the left visual field, in novices and experts, in Experiment 2. There was a small, nonsignificant reduction of the N170 in response to faces in car experts, when a car picture is presented (car context), relative to a scrambled car picture. Compare the reduction of the effect of expertise in this experiment with the effect observed when the car picture remains on the screen during the presentation of the face stimulus (Figure 2).

Chelazzi, & Desimone, 1999; Moran & Desimone, 1985). That is, if a preferred visual stimulus for a cell is presented together with a poor visual stimulus, the cell's response is reduced compared to that elicited by the single good stimulus. This sensory suppressive interaction among multiple visual stimuli has been observed at the single-cell level in several visual areas in the ventral stream of the monkey brain (V2, V4, infero-temporal cortex [IT]; Reynolds et al., 1999; Rolls & Tovee, 1995; Miller, Gochin, & Gross, 1993; Moran & Desimone, 1985) and is generally interpreted as an expression of competition for neural representation (Kastner & Ungerleider, 2001; Desimone, 1998). In IT, it was described for neurons responding preferentially to faces (Rolls & Tovee, 1995; Miller et al., 1993). In the same vein, fMRI studies performed on human subjects reported a reduction of blood-oxygen-level-dependent signal in several extrastriate visual areas of the ventral stream (V2, V4, TEO, TE) when presenting complex shapes simultaneously compared to sequential presentation (Kastner & Ungerleider, 2001; Kastner, De Weerd, Desimone, & Ungerleider, 1998).

Compared to these observations of sensory suppressive interactions between visual stimuli at the single-cell level or in fMRI, the interest of the present work is twofold. First and most interestingly, these results show that a competition between different object shapes (i.e., a car and a face stimulus) can be dramatically increased with visual experience, and is thus not only dependent on the visual structure of the stimuli (Beck & Kastner, 2005). Second, by virtue of the excellent temporal resolution offered by ERP recordings and the spatial sampling of the whole system, the results demonstrate that visual competition between faces and objects of expertise takes place as early as 130 msec in the human brain, during a limited time window, in occipito-temporal areas.

The sensory suppression observed here in neurophysiological recordings on the occipito-temporal human

Table 3. Latency and Amplitude Values of the N170 in Response to Faces for the Two Groups of Subjects in the Three Experimental Conditions for Experiment 2

Context	Novices						Experts					
	Car		Scrambled Car		Fixation Cross		Car		Scrambled Car		Fixation Cross	
Hemisphere	L	R	L	R	L	R	L	R	L	R	L	R
(A) Left Visual Field												
Latencies (msec)	188	164	180	160	180	156	184	156	180	152	176	148
Amplitudes (μV)	**−3.36**	**−4.06**	−4.18	−5.62	−3.47	−5.02	**−2.16**	**−3.8**	−3.33	−5.32	−2.90	−5.00
(B) Right Visual Field												
Latencies (msec)	156	184	152	180	152	180	152	180	148	176	148	180
Amplitudes (μV)	**−4.54**	**−4.49**	−5.34	−5.56	−5.05	−4.54	**−3.11**	**−3.91**	−4.15	−5.18	−3.63	−4.64

Table 4. Experiment 2

Correlation Measures	Cars–Scrambled	Cars–Fixation
RH–LVF presentation	**r = .42, p < .009**	**r = .32, p < .05**
RH–RVF presentation	r = .27, p < .10	**r = .35, p = .029**
LH–RVF presentation	**r = .40, p = .013**	r = .25, p = .13
LH–LVF presentation	**r = .38, p = .018**	**r = .48, p = .0019**

Correlation measures between the behavioral index of expertise and the decrease of N170 amplitude to faces competing with car stimuli, relative to control conditions (scrambled car and fixation cross).

scalp suggests two possible mechanisms at the neurophysiological level.[1] First, cells responding preferentially to faces in occipito-temporal areas may also respond to nonface objects of expertise (i.e., cars) in experts. The face-N170 is thought to originate from multiple brain areas located in the occipito-temporal cortex, including the fusiform gyrus, the superior temporal sulcus, and the middle and inferior temporal gyri (e.g., Herrmann, Ehlis, Muehlberger, & Fallgater, 2005; Henson et al., 2003), where cortical surface potentials in response to faces—N200s—have been observed at roughly the same latency (Allison, Puce, Spencer, & McCarthy, 1999). Single-cell recordings in the monkey inferotemporal cortex show that cells in these areas are organized in columns that may be highly selective to face stimuli (e.g., Tsao, Freinwald, Tootell, & Livingstone, 2006; Tanaka, 1996; Perrett, Hietanen, Oram, & Benson, 1992; Desimone, 1991). However, it is yet unclear whether these neurons are tuned to respond to faces only (i.e., are ''domain-specific''), or if they may also fire in response to members of a nonface object category following extensive visual experience with this category. Single neurons in the monkey IT can tune their response to novel visually similar objects—bars or ''amoeba'' shapes—following expertise training (Baker, Behrmann, & Olson, 2002; Logothetis & Pauls, 1995). These neurons appeared to share a number of properties with face-selective neurons such as viewpoint-selectivity (Logothetis, Pauls, Bulthoff, & Poggio, 1994) and a strong sensitivity to the removal of parts of the stimulus (Logothetis & Pauls, 1995). However, recordings in these studies are made in more anterior and ventral areas of IT than in the regions where most face-selective cells have been reported, and the response of these cells to face stimuli is unknown. Second, increased sensory suppression between objects of expertise and faces may be due to competitive interactions from distinct populations of cells through local lateral inhibitory connections (Allison, Puce, & McCarthy, 2002; Wang, Fujita, & Muruyama, 2000). In the monkey brain, local inhibition contributes to generating the specificity of IT neurons to complex stimuli. Furthermore, blocking inhibition in IT mostly reveals responses of a cell to new stimuli differing to a preferred stimulus in a systematic way along certain parameters (contrast, piece of shape, etc....), suggesting that local competition between preferred stimuli at the single-cell level is not randomly organized but depends on the object visual features (Wang et al., 2000).

In summary, the reduced amplitude of the N170 potential observed in response to faces here may result from the recruitment of face cells for nonface objects of expertise, or to an increased local competition from distinct populations of cells coding for car stimuli, following extensive visual expertise training. From a functional point of view, the observation of a competition between the early processing of faces and cars strongly suggests that the same perceptual mechanisms are used for both categories in car experts, whether the effect is due to an overlap at the cellular level, or to a competition between different populations of neurons that are intermingled. Thus, the perceptual mechanisms reflected by the N170 do not appear to be dedicated to visual stimuli with a facial configuration, but they can be, in a large part, recruited for nonface objects following expertise training. Furthermore, given that these results have been observed with rather different object shapes (Rossion et al., 2004), this indicates that the tuning of the perceptual mechanisms is quite broad with respect to object geometry.

Competition during Concurrent and Delayed Stimulus Presentation

Competition effects were much larger when the presentation of faces and cars temporally overlapped (Experiment 1) than when they were presented at separate times (Experiment 2). This suggests that the competition between the processes recruited for the two stimuli is maximized when the object of expertise is still eliciting a sustained and ongoing activation in high-level visual areas. Moreover, a concurrent stimulation paradigm allows eliciting a competition between perceptual processes/representations, rather than relying on interference effects between visual short-term memory representations and visual processes for the incoming visual stimulus (Jacques & Rossion, 2004, 2006; Rossion et al., 2004). Yet, given the poor spatial resolution of scalp field potential recordings, the two competing stimuli have to be presented with an onset asynchrony (as in the face adaptation paradigm in ERPs; see Kovacs et al., 2006). A period of about half a second is chosen here as a compromise between a null (i.e., simultaneous presentation) or a very short stimulus onset asynchrony (SOA), for which the evoked potentials associated with the face stimulus could not be isolated, or a much longer SOA, which could lead to a reduced competition

given the progressive decrease of neural response to the sustained presentation of the first stimulus (i.e., neural habituation; Ringo, 1996). The presence of an SOA between the two competing stimuli is a fundamental difference with respect to single-cell recordings studies of competing visual shapes, in which the activity of different neurons can be measured independently by presenting the stimuli in their respective receptive fields or by presenting different preferential stimuli (e.g., Reynolds et al., 1999; Rolls & Tovee, 1995; Miller et al., 1993; Moran & Desimone, 1985). The ERP paradigm also differs to neuroimaging experiments where the activation of separate visual areas can be sampled independently (Beck & Kastner, 2005; Kastner et al., 1998). The optimal SOA duration to observe a marked competition effect in scalp ERPs is an empirical matter, but may also depend on the task at hand. For instance, neurons in the inferotemporal cortex discharge in response to a first stimulus for a longer period of time if the second stimulus is relevant relative to the first, such as in cued visual search tasks (Chelazzi, Duncan, Miller, & Desimone, 1998), suggesting that the effects of competition described here may be larger when the subject has to compare the two events rather than concentrating on the second one.

Spatial Attention and Visual Competition

An alternative explanation to the suppression of response to faces occurring between 130 and 180 msec in the occipito-temporal cortex may be that subjects reduced attention to the portion of the visual field where the face stimulus appears, when another meaningful stimulus (i.e., cars) is presented in the central part of the visual field. There is indeed evidence that early visual-evoked activity in the extrastriate cortex can be suppressed when attention is shifted either from the location of the evoking stimulus to another location in the visual field (Luck et al., 1994), or from the object of interest presented at the same location (Chen, Seth, Gally, & Edelman, 2003). However, an attentional account of the present results is very unlikely for several reasons. First, the reduction of amplitude in response to faces occurred in interaction with expertise. Whereas it may be argued that the object of expertise is more salient to experts, the reverse may be equally true (i.e., cars may be more salient to novices given their lesser experience with this category). In other words, the direction of the effect is unclear in an attentional account, whereas an expertise-mediated competition account predicts a reduction of face-related ERPs for car experts during the presentation of the car picture, as was observed. Second, there was no interaction between expertise and the conditions in accuracy/RT in the behavioral task, suggesting that experts did not have less resources than novices to process the face targets when car pictures were presented in the center. Admit-

tedly, this could be due to ceiling effect in a very simple visuomotor task, which was not sensitive enough to detect differences between the two groups, but again there is no evidence for an attentional confound here. Finally, the effect found here occurs in a narrow time window, between 130 and 180 msec after stimulus onset, whereas effects of attention on early visual components typically start at 80 msec, at the level of the P1, and are generally sustained (Luck, Woodman, & Vogel, 2000). As a matter of fact, when the effect of attention and competition for face processing are directly compared, the effect of attention starts clearly at the P1 level (about 80 msec onset), much earlier than the N170 competition effect (see Jacques & Rossion, in press). In the present study, if anything, novices showed a *larger* P1 than experts in Experiment 1 (Figure 2) and—less clearly—in Experiment 2 (Figure 6). These group differences could be due to intersubjects' variability in P1 amplitude, or functional differences (i.e., novices paying more attention overall), but, critically, there was no difference between the conditions (competing stimulus) for experts or novices at this latency (see Figure 7, with complementary analyses). Had experts paid more attention than novices to the car pictures in the present study, one should have observed a specific increase of amplitude for the P1 component in experts, *in that condition only*.

In sum, there is no evidence that differences in spatial attention to the face stimulus presented laterally may account for the results observed here, which can be attributed to inherent competition between visual processes for competing object representations. Yet, there is behavioral and neurophysiological evidence that visual attention may mediate the underlying suppressive interaction between two stimuli of interest (i.e., faces and objects of expertise) when they are shown concurrently in the visual field (Chelazzi et al., 1998; Desimone, 1998). According to this biased competition model (Kastner & Ungerleider, 2001; Desimone, 1998), objects in the visual field compete for the response of cells in the visual cortex, and these competitive interactions are stronger in a given cortical area when competing stimuli activate cells in the same region of the cortex (e.g., IT). Visual attention may bias one stimulus over the other one, by virtue of both bottom-up (i.e., salience) and feedback top-down (i.e., behavioral relevance) mechanisms. The biased competition by means of visual attention can be expressed as a reduction of the amount of spikes for the unattended object (e.g., Chelazzi et al., 1998), or a reduction of the cell's receptive field (Rolls, Aggelopoulos, & Zheng, 2003).

Evidence for Nonmodular Face Mechanisms

Our results indicate that visual expertise with a nonface category of man-made objects, cars, leads to the recruitment of perceptual mechanisms that are normally
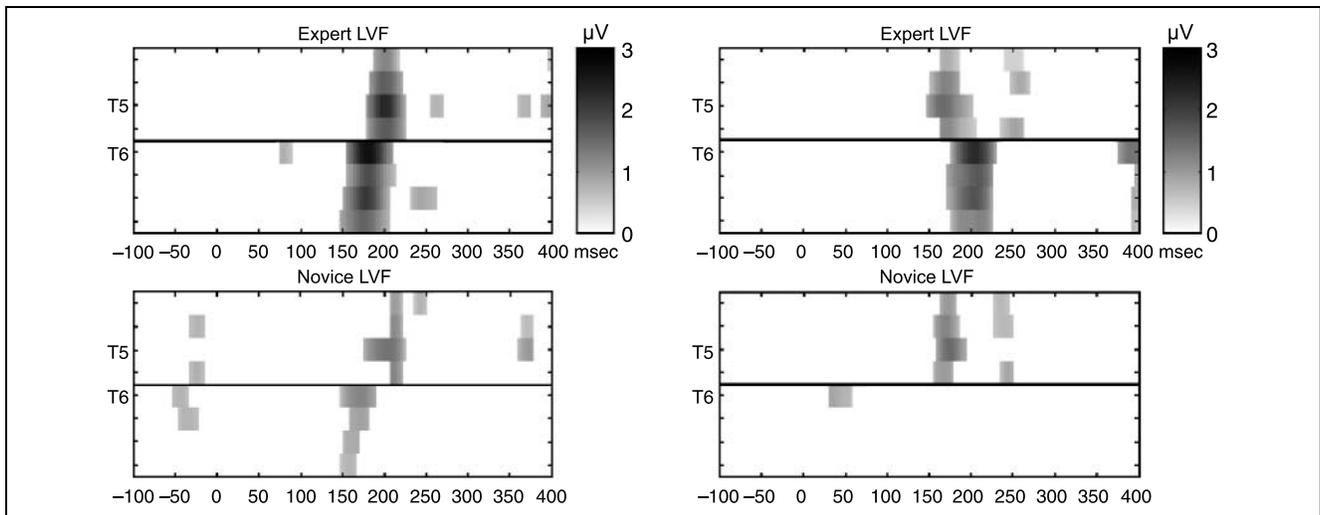
Figure 7. Significant differences between ERP elicited by faces in the scrambled car versus car context independently for experts and novices and for the left and right visual fields (LVF and RVF) assessed at each time points between −100 and 400 msec (Experiment 1). Significance thresholds were computed using a nonparametric bootstrap permutation methods (2000 permutations, two-tailed test, $p < .01$). The figure represents time regions of significant differences computed separately for experts/novices, LVF/LVF, and the eight inferotemporal electrodes (four in each hemisphere) where the N170 was maximal (see Methods). Gray-scale patches represent the magnitude of the difference between the conditions in absolute value as a function of time, whenever significantly different from zero. A value is considered significantly different from zero if four consecutive samples (16 msec at 250 Hz sampling rate) are below the $p < .01$ level. These results show that the difference between car and scrambled car conditions in experts is limited to the 150–200 msec time window (as it is for novices, but with much less significance) and neither affects the preceding component P1 nor the later components.

involved in processing face stimuli. At the functional level, this conclusion stands whether the exact same populations of neurons code for both object categories, or if close populations of neurons carrying identical functions (i.e., a degenerate network; Tononi, Sporns, & Edelman, 1999) compete with each other. The main effects of expertise were found in the right hemisphere, whether face stimuli were presented in the left or right visual field. This concurs with previous observations disclosing such competition effects (Jacques & Rossion, 2004, 2006, in press; Rossion et al., 2004; Gauthier et al., 2003), and with other effects of expertise (e.g., Busey & Vanderkolk, 2005; Gauthier et al., 2000; although a stronger left hemisphere effect was found in Rossion et al., 2002). More generally, it is in line with a wealth of evidence supporting the right hemisphere dominance in face processing (e.g., Zangenehpour & Chaudhuri, 2005; Le Grand, Mondloch, Maurer, & Brent, 2003; Kanwisher et al., 1997; Sergent et al., 1992; Sergent & Signoret, 1992), which appears to be related to holistic/configural processes for faces (Schiltz & Rossion, 2006; Le Grand et al., 2003; Rossion et al., 2000; Hillger & Koenig, 1991).

The question of whether faces are handled by a modular system has been the topic of much controversy in the field for more than half a century. It was raised originally by the observation of face-specific disorders following brain damage (prosopagnosia, Bodamer, 1947; see e.g., Sergent & Signoret, 1992), single neurons responding exclusively to face stimuli in the nonhuman primate visual cortex (Gross, Rocha-Miranda, & Bender, 1972; e.g., Perrett et al., 1992; Desimone, 1991; Perrett,

Rolls, & Caan, 1982), and behavioral evidence in humans for disproportionate effects of stimulus transformation such as upside-down inversion on faces (Yin, 1969). Whereas a number of researchers have interpreted these and subsequent findings as supporting the view that faces are handled by fixed specific processing mechanisms (e.g., Kanwisher, 2000), others have emphasized the role of visual experience and processing requirements in shaping face and object differences, claiming that nonface objects can be handled by face-related mechanisms in the human adult brain (Tarr & Cheng, 2003). Ultimately, the resolution of this debate will have to come from both conceptual clarifications and agreements inside and outside this field, and from empirical evidence. In the present study, we found compelling evidence for clear-cut and substantial modulation of early face categorization processes following visual expertise with nonface objects. These findings do not contradict the statement that ''faces are special'' for the human adult visual system, but they do support the view that a substantial proportion of visual processes related to faces can be also recruited for other objects following long-term visual expertise, leading to expertise-mediated effects of competition in the occipito-temporal cortex between 130 and 200 msec following stimulus onset.

Acknowledgments

Note

1. It should be kept in mind that the relationship between neuronal activity at the single-cell level and observed on far-field potentials is indirect. It is generally acknowledged that ERPs originate mostly from postsynaptic depolarization generated along the apical dendrites of cortical pyramidal cells, not from spike trains at the level of the neuron's axon. However, an increase in spike rate in a synchronized population of cells will be associated with an increase of activity at the postsynaptic level (i.e., current flows), leading to recordable far field potentials.

REFERENCES

Allison, T., Puce, A., & McCarthy, G. M. (2002). Category-sensitive excitatory and inhibitory processes in human extrastriate cortex. *Journal of Neurophysiology, 88,* 2864–2868.

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

Bachmann, T., & Kahusk, N. (1997). The effects of coarseness of quantisation, exposure duration, and selective spatial attention on the perception of spatially quantised (''blocked'') visual images. *Perception, 26,* 1181–1196.

Baker, C. I., Behrmann, M., & Olson, C. R. (2002). Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nature Neuroscience, 5,* 1210–1216.

Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience, 8,* 1110–1116.

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

Bodamer, J. (1947). Die-Prosop-agnosie. *Archiv fuer Psychiatrie und Nervenkrankheiten, 179,* 6–54 [H. D. Ellis & M. Florence, Trans. (1990). *Cognitive Neuropsychology, 7,* 81–105].

Busey, T., & Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. *Vision Research, 45,* 431–448.

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 335,* 95–103.

Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology, 80,* 2918–2940.

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, U.S.A., 100,* 3501–3506.

Desimone, R. (1991). Face-selective cells in the temporal cortex of monkeys. *Journal of Cognitive Neuroscience, 3,* 1–8.

Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 353,* 1245–1255.

Gauthier, I., Curran, T., Curby, K. M., & Collins, T. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6,* 428–432.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3,* 191–197.

Gauthier, I., & Tarr, M. J. (1997). Becoming a ''Greeble'' expert: Exploring the face recognition mechanism. *Vision Research, 37,* 1673–1682.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform area increases with expertise in recognizing novel objects. *Nature Neuroscience, 6,* 568–573.

Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology, 35,* 96–111.

Henson, R. N., Goshen-Gottstein, Y., Ganel, T., Otten, L. J., Quayle, A., & Rugg, M. D. (2003). Electrophysiological and haemodynamic correlates of face perception, recognition and priming. *Cerebral Cortex, 13,* 793–805.

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

Hillger, L. A., & Koenig, O. (1991). Separable mechanisms in face processing: Evidence from hemispheric specialization. *Journal of Cognitive Neuroscience, 3,* 42–58.

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

Jacques, C., & Rossion, B. (2006). The time course of visual competition to the presentation of centrally fixated faces. *Journal of Vision, 6,* 154–162.

Jacques, C., & Rossion, B. (in press). Electrophysiological evidence for temporal dissociation between spatial attention and sensory competition during human face processing. *Cerebral Cortex.*

Kanwisher, N. (2000). Domain-specificity in face perception. *Nature Neuroscience, 3,* 758–763.

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

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science, 282,* 108–111.

Kastner, S., & Ungerleider, L. G. (2001). The neural basis of biased competition in human visual cortex. *Neuropsychologia, 39,* 1263–1276.

Kovacs, G., Zimmer, M., Banko, E., Harza, I., Antal, A., & Vidnyansky, Z. (2006). Electrophysiological correlates of

visual adaptation to faces and body parts in humans. *Cerebral Cortex, 16,* 742–753.

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2003). Expert face processing requires visual input to the right hemisphere during infancy. *Nature Neuroscience, 6,* 1108–1112.

Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered representations in the primate. *Cerebral Cortex, 5,* 270–288.

Logothetis, N. K., Pauls, J., Bulthoff, H. H., & Poggio, T. (1994). View dependent object recognition by monkeys. *Current Biology, 4,* 401–414.

Luck, S. J., Hillyard, S. A., Mouloua, M., Woldorff, M. G., Clark, V. P., & Hawkins, H. L. J. (1994). *Journal of Experimental Psychology: Human Perception and Performance, 20,* 87–104.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences, 4,* 432–440.

McCarthy, G., Puce, A., Gore, J., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience, 9,* 605–610.

Miller, E. K., Gochin, P. M., & Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Research, 616,* 25–29.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784.

Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 335,* 23–30.

Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research, 47,* 329–342.

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience, 19,* 1736–1753.

Ringo, J. L. (1996). Stimulus specific adaptation in inferior temporal and medial temporal cortex of the monkey. *Behavioural Brain Research, 76,* 191–197.

Rolls, E. T., Aggelopoulos, N. C., & Zheng, F. (2003). The receptive fields of inferior temporal cortex neurons in natural scenes. *Journal of Neuroscience, 23,* 339–348.

Rolls, E. T., & Tovee, M. J. (1995). The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. *Experimental Brain Research, 103,* 409–420.

Rossion, B., Dricot, L., De Volder, A., Bodart, J.-M., Crommelinck, M., de Gelder, B., et al. (2000). Hemispheric asymmetries for whole-based and parts-based face processing in the

human fusiform gyrus. *Journal of Cognitive Neuroscience, 12,* 793–802.

Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelinck, M. (2002). Expertise training with novel objects leads to face-like electrophysiological responses. *Psychological Science, 13,* 250–257.

Rossion, B., Kung, C. C., & Tarr, M.-J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, U.S.A., 101,* 14521–14526.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage, 32,* 1385–1394.

Sergent, J. (1986). Microgenesis of face perception. In H. D. Ellis, M. A. Jeeves, F. Newcombe, & A. M. Young (Eds.), *Aspects of face processing* (pp. 17–33). Dordrecht: Martinus Nijhoff.

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

Sergent, J., & Signoret, J.-L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex, 2,* 375–388.

Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience, 19,* 109–139.

Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science, 12,* 43–47.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23,* 457–482.

Tarr, M. J., & Cheng, Y. (2003). Learning to see faces and objects. *Trends in Cognitive Sciences, 7,* 23–30.

Terasaki, O., & Okazaki, M. (2002). Transcallosal conduction time measured by visual hemifield stimulation with face images. *NeuroReport, 13,* 97–99.

Tononi, G., Sporns, O., & Edelman, G. M. (1999). Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Sciences, U.S.A., 96,* 3257–3262.

Tucker, D. M. (1993). Spatial sampling of head electrical fields: The geodesic sensor net. *Electroencephalography and Clinical Neurophysiology, 87,* 154–163.

Tsao, D. Y., Freinwald, W. A., Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science, 311,* 670–674.

Wang, Y., Fujita, I., & Muruyama, Y. (2000). Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nature Neuroscience, 3,* 807–813.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 41–145.

Zangenehpour, S., & Chaudhuri, A. (2005). Patchy organization and asymmetric distribution of the neural correlates of face processing in monkey inferotemporal cortex. *Current Biology, 15,* 993–1005.