Cognition 83 (2002) 1–29

COGNITION

www.elsevier.com/locate/cognit

Domain specificity versus expertise: factors influencing distinct processing of faces

David Carmel[a], Shlomo Bentin[a,b,*]

[a]*Department of Psychology, Hebrew University of Jerusalem, Jerusalem 91905, Israel*
[b]*Center for Neural Computation, Hebrew University of Jerusalem, Jerusalem 91905, Israel*

Received 1 June 2001; accepted 21 September 2001

Abstract

To explore face specificity in visual processing, we compared the role of task-associated strategies and expertise on the N170 event-related potential (ERP) component elicited by human faces with the ERPs elicited by cars, birds, items of furniture, and ape faces. In Experiment 1, participants performed a car monitoring task and an animacy decision task. In Experiment 2, participants monitored human faces while faces of apes were the distracters. Faces elicited an equally conspicuous N170, significantly larger than the ERPs elicited by non-face categories regardless of whether they were ignored or had an equal status with other categories (Experiment 1), or were the targets (in Experiment 2). In contrast, the negative component elicited by cars during the same time range was larger if they were targets than if they were not. Furthermore, unlike the posterior-temporal distribution of the N170, the negative component elicited by cars and its modulation by task were more conspicuous at occipital sites. Faces of apes elicited an N170 that was similar in amplitude to that elicited by the human face targets, albeit peaking 10 ms later. As our participants were not ape experts, this pattern indicates that the N170 is face-specific, but not specie-specific, i.e. it is elicited by particular face features regardless of expertise. Overall, these results demonstrate the domain specificity of the visual mechanism implicated in processing faces, a mechanism which is not influenced by either task or expertise. The processing of other objects is probably accomplished by a more general visual processor, which is sensitive to strategic manipulations and attention.

Keywords: Domain specificity; Expertise; Distinct processing of faces; ERPs; N170; N1

1. Introduction

Several lines of evidence suggest that the neural and perceptual processes involved in

---

[*] Corresponding author. Tel.: +972-2-5881090; fax: +972-2-5825659.
E-mail address: shlomo.bentin@huji.ac.il (S. Bentin).

0010-0277/02/$ - see front matter © 2002 Published by Elsevier Science B.V.
PII: S0010-0277(01)00162-7

face perception in humans[1] are distinct and segregated from those involved in the visual perception and identification of other objects.

One line of evidence is based on studies of face perception in infants. Using a visual tracking task or measuring the time spent by infants looking at different stimuli, several studies have shown that newborn babies (in some cases, a few minutes post-partum) prefer looking at face-like visual configurations rather than at other, equally complex stimuli (e.g. Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991; Maurer & Young, 1983; Simion, Valenza, Umilta, & Dalla Barba, 1998; Valenza, Simion, Cassia, & Umilta, 1996). Indeed, this pattern suggests the existence of an innate mechanism dedicated to face processing, but it has not always been replicated (e.g. Easterbrook, Kisilevsky, Hains, & Muir, 1999). Other studies have suggested that the apparent preference for faces versus non-faces in newborns reflects basic visual preferences (e.g. for higher stimulus density appearing in the upper than in the lower hemifields), rather than content-determined preferences (Mondloch et al., 1999; Simion, Cassia, Turati, & Valenza, 2001). However, even studies showing that newborns cannot distinguish between normally configured and scrambled faces without training found that they do distinguish between stimuli that contain and stimuli that do not contain physiognomic information (e.g. Easterbrook, Kisilevsky, Muir, & Laplante, 1999). Moreover, in contrast to the controversy regarding newborns' visual preferences, there is a general consensus that by about 2 months of age, the preference for faces is reliable (e.g. Maurer & Barrera, 1981). Some studies have even shown a within-category distinction (for the mother's face) (e.g. Bushnell, Sai, & Mullin, 1989; Pascalis, de Schonen, Morton, Deruelle, & Fabre-Grenet, 1995; Walton, Bower, & Bower, 1992). The reliable preference for faces 2 months after birth and the absence of such a clear preference in newborns may suggest that the face processing specificity, assumed in the adult perceptual system, may be induced by an experience-expectant developmental process (Nelson, 2001). This view is rooted in a neuroconstructivist approach to cognitive development (Karmiloff-Smith, 1995, 1997). Note, however, that neuroconstructivism theory accepts some form of innately specified starting points. Indeed, this approach highlights how tiny variations in an initial, perhaps innate predilection, could give rise to domain-specific mature structures (Karmiloff-Smith, 1998). Within such a framework, Johnson and Morton (1991) introduced a developmental theory of face domain specificity. According to their theory, humans are born with a subcortical mechanism that serves to orient rudimentary attention towards high-contrast elements typical to the configuration of human faces (Conspec). The influence of this mechanism lasts for about 2 months from birth, during which time it facilitates the development of a cortical mechanism (Conlern) which benefits from experience with faces, and accounts for the formation of a perceptual bias towards them (Johnson, 1997); this mechanism develops quickly into an adult expert system (Carey, 1992; Diamond & Carey, 1986).

Although the Johnson and Morton (1991) model (as well as other similar models such as de Schonen & Mathivet, 1990) reduced the role of innateness in face specificity to only a "head-start", they do not argue against domain specificity in adults. One can find analogous claims in the domain of spoken language. Although it is generally accepted that basic

---

[1] There is ample evidence for neural specificity in face perception in primates. Although suggestive and supporting neural specificity, we will limit the present introduction to studies in humans.

linguistic functions in adults rely on domain-specific neural structures, the issue of whether their domain specificity is innately determined or arises out of primary predispositions that are shaped by the nature of the experience that the organism is exposed to is still under debate (for an in-depth discussion of the nature/nurture debate see Elman et al., 1996).

The second line of evidence suggesting characteristic (albeit not unique) processing of faces includes a large corpus of studies showing a greater sensitivity to configural changes in face processing than in visual processing of other objects (e.g. Rhodes, 1988; Sergent, 1988; Tanaka & Sengco, 1997). By and large, performance studies demonstrated that faces, more than other objects, are processed “holistically” (e.g. Tanaka & Farah, 1993; for a recent review and detailed exposition of this view see Farah, Wilson, Drain, & Tanaka, 1998). Although global perceptual strategies have been occasionally reported for other stimuli such as cars (Tanaka et al., unpublished data, cited by Gauthier, Williams, Tarr, & Tanaka, 1998), and can be enhanced by increasing expertise with other well-defined visual categories (e.g. Gauthier & Tarr, 1997), only for faces does this strategy develop without explicit training (or might be innate). Indeed, a naturally developed “expert” processing mechanism may be exactly what distinguishes face perception from that of other stimulus types; the fact that extensive training might render strategies that are naturally used for face perception (such as global processing) available for other stimulus types does not weaken the claim of specificity (cf. Diamond & Carey, 1986). Furthermore, even after dozens of repeated exposures, configural manipulations can distinguish between the brain response to faces and to other objects (Rossion et al., 2000).

The above studies suggest that the cognitive system processes faces in a distinctive manner; evidence from two other lines of research points to the specificity of the neural mechanism used for face processing. One line includes neuropsychological studies of patients who seem to have lost (or never developed) the ability to recognize familiar faces (for a recent review see De Renzi, 1997). Although in most cases prosopagnosia (i.e. impaired face identification) is only the conspicuous aspect of a more general visual agnosia (e.g. Gauthier, Behrmann, & Tarr, 1999), there are a few reports of prosopagnosic patients whose ability to recognize objects was intact (e.g. Bentin, Deouell, & Soroker, 1999; Farah, Levinson, & Klein, 1995; McNeil & Warrington, 1993). Conversely, there are reports of patients suffering from associative object agnosia whose face recognition ability was spared (e.g. Moscovitch, Winocur, & Behrmann, 1997). The double dissociation between face and object recognition implies that the two abilities are functionally distinct, and suggests that they are anatomically segregated, in that focal damage can selectively impair either one. Furthermore, as suggested by a case who, following brain damage incurred at 1 day of age, suffered from severe prosopagnosia with fairly good object recognition, the ability to distinguish between faces and non-face objects does not require explicit experience with these categories (Farah, Rabinowitz, Quinn, & Liu, 2000). Although these data were criticized, particularly for not including reaction time measurements (Gauthier & Nelson, 2001), the overall pattern is clear and convincing. Indeed, the major question raised by this interesting case is why, despite brain plasticity that helps in developing almost normal linguistic functioning even in severe cases of right infantile hemiplegia (Kohn & Dennis, 1974), similar plasticity was not observed in that case.

The specificity of the neural substrate of face recognition has been further supported by studies in which the brain electrical or hemodynamic activity was recorded in association with the perception of human faces and other objects. Such studies implicated the ventral and inferior-temporal regions of the human brain (the "what" pathway of the visual system) in high-level visual processing of human faces (see reviews in Bentin, in press; McCarthy, 1999). For example, early investigations using PET showed that, compared with nonsense gratings and sinusoid shapes, both faces and objects activate extensive areas in the occipito-temporal cortex. However, whereas this activation was more pronounced in the right hemisphere for faces, it was larger in the left hemisphere for objects (Sergent, Ohta, & MacDonald, 1992). Face matching was also associated with selective rCBF increases in the fusiform gyrus in occipital and occipito-temporal cortex bilaterally and in a right prefrontal area in the inferior frontal gyrus (Haxby et al., 1994). More recently, fMRI studies provided a more detailed picture: faces activated the fusiform and middle occipital gyri, the lateral occipital sulcus and a more anterior region in the superior temporal sulcus (STS); this activation was larger in the right than in the left hemisphere (e.g. Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997). Note, however, that fMRI studies also showed distinct activation for semantic categories other than faces in regions adjacent to the face areas (e.g. Ishai, Ungerleider, Martin, & Haxby, 1999) and that activity in the fusiform face area (FFA) can be enhanced by stimuli other than faces by increasing expertise with them (Gauthier, Tarr, Anderson, Skudlarski, & Gore, 1999). However, as mentioned above, these findings should not necessarily weigh down the claim for neural specificity for face processing; they may merely suggest that semantic categories other than faces might activate distinctive regions in the ventral visual pathway, and that the FFA can be recruited to process stimuli for which humans may become experts. This argument, however, requires more substantiation.

Whereas neuroimaging studies are consistent in showing neuroanatomical specificity for face processing, they do not establish when such processing occurs. A tentative answer to this question is provided by event-related potential (ERP) studies. Recording subdurally, directly from the cortical surface, from more than 100 patients, Allison, Puce, Spencer, and McCarthy (1999) revealed that discrete regions in the human extra-striate cortex were activated by faces but not by other categories of visual stimuli such as cars, flowers, human hands, butterflies or printed words. The great majority of the activated sites were clustered on the ventral occipito-temporal cortex, predominantly in the fusiform gyrus, and some were found on the posterior lateral surface of the temporal lobes. The earliest face-specific component was a robust negative potential peaking around 200 ms (N200) from stimulus onset. Later face-specific activity was recorded from anterior fusiform regions and was tentatively associated with the identification (as opposed to the detection) of the face (see reviews in Allison et al., 1999; McCarthy, Puce, Belger, & Allison, 1999; Puce, Allison, & McCarthy, 1999). Hence, epicortical and intracranial electrophysiological recordings supported the localization of the face-specific neural mechanisms in the posterior lateral and posterior ventral regions suggested by hemodynamic studies, and revealed that faces are processed distinctively already at about 200 ms from stimulus onset. Far-field face-specific ERPs have also been recorded at the scalp. Being non-invasive, scalp ERPs are instrumental for the investigation of the normal brain function in humans. In accordance with the intracranial data, scalp ERP studies disclosed a negative component peaking at about 170 ms from stimulus onset (N170) whose amplitude was considerably larger in response to human faces than in response to equiluminant nonsense visual patterns, or meaningful non-face objects such as birds, furniture, human hands, cars, flowers, or butterflies (Bentin, Allison, Puce, Perez, & McCarthy, 1996; for converging evidence see, for example, George, Evans, Fiori, Davidoff, & Renault, 1996).[2] Although most stimuli elicit a local, negative or negative-going[3] potential at about the same latency, only faces consistently lead to a negative peak, differing from the activation for any other stimulus category. Furthermore, differences between the N170s elicited by non-face stimulus types are smaller and by far less consistent.

During the past 5 years, Bentin and his colleagues (as well as others) have extensively investigated the characteristics of the face-specific brain response manifested by the N170. In brief, they found that the N170 is not sensitive to face familiarity (Bentin & Deouell, 2000), its amplitude is slightly (but not dramatically) reduced if configuration of the inner components is distorted (Bentin et al., 1996), it can be elicited by isolated face components, particularly by eyes (in which case its amplitude is significantly augmented and its latency slightly delayed; Bentin et al., 1996; Bentin & McCarthy, unpublished data), and it is enhanced and delayed by face inversion (Rossion et al., 1999, 2000; Sagiv & Bentin, in press). In addition to photographs of natural faces, the N170 is also elicited by schematic drawings of faces, in which case inversion of the face significantly reduces its amplitude (Sagiv & Bentin, in press). Integrating ERP data from intracranial and scalp recordings as well as from available fMRI studies, Bentin (in press) suggested that visual stimuli including physiognomic information are distinctively integrated into a perceptual representation. We hypothesize that this process is performed by a complex neural mechanism that has the ability to process the face holistically but also to detect conspicuous physiognomic features (such as eyes). According to this model, the global process is performed by neural networks located in the middle fusiform gyrus, whereas the detection of face components relies on neural networks located more laterally, in the STS and the posterior inferior temporal gyrus (IT) (Bentin et al., 1996). Furthermore, we assume that whereas the activity of the fusiform global processor is manifested by the intracranially recorded N200, the N170, although associated both with the global and the component oriented activity, is influenced mostly by the latter. This hypothesis is based on the differential sensitivity of the N200 and the N170 to manipulations affecting the natural configuration of the face. Whereas, recorded from the FFA, the N200 is larger for full faces than for isolated eyes and is reduced by face inversion (at least in the right hemisphere), the N170 is considerably larger for eyes than for full faces, and is enhanced by inversion. By contrast, isolated components of schematic (rather than natural) faces do not elicit the N170 unless primed (Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, in press) and inversion of schematic faces reduces the amplitude of the N170 (Sagiv & Bentin, in press). This contrast can be accounted for assuming that for schematic faces, the N170 is affected solely by the holistic processor.

The concise review above indicates that face processing seems to implicate distinguish-

---

[2] A face-specific positive component peaking at the scalp vertex at 190 ms has been reported by Jeffreys & Tukmachi (1992).

[3] Meaning that, although positive, this component constitutes a negative peak relative to the period immediately before or after it.

able cognitive procedures that are carried out by localized neural substrates. This idea, however, has been challenged by findings showing that the FFA is activated by other objects as well as by faces, if these objects belong to a stimulus category with which the observer has acquired expertise through extensive training for within-category discrimination (e.g. Gauthier, Skudlarski, Gore, & Anderson, 2000; Gauthier, Tarr et al., 1999). In order to account for such findings, Gauthier and her collaborators suggested that faces are not processed by a domain-specific visual system; rather, the physiological measures suggesting specificity merely reflect the activity of a visual mechanism applied to perceptual categories for which there is expertise and not applied to other, less familiar stimulus categories. Additional support for the "expertise hypothesis" was provided by a recent ERP study, in which the N170 was larger in response to birds than to dogs in bird experts, but larger for dogs than for birds in dog experts (Tanaka & Curran, 2001). Indeed, it has been claimed that in order for the N170 to be regarded as a marker of a face-specific system it is not sufficient to show that its amplitude is larger in response to faces than to other stimulus categories (Rossion et al., 2000). Other factors to be considered are the existence or absence of similar N170 distinctions among other categories, and the interaction of these distinctions with task-associated strategies, such as attention and levels of categorization, as well as observer-associated factors such as level of expertise. To this end, the present study was aimed at exploring the influence of task-induced processing strategies, and the influence of expertise on the N170 associated with faces and other stimulus categories.

2. General methods

2.1. Participants

The same participants were tested with all three tasks. They were 12 right-handed undergraduates (nine females), with a mean age of 22.7 years. They did not suffer from any diagnosed neurological problem, and were not using any psychotropic medication or drugs. They participated in the experiment for payment or for course credit. All participants had normal or corrected to normal vision.

2.2. Stimuli

The stimuli were black and white photographs as follows: 210 photographs of human faces (frontal close-up portraits), 140 photographs of birds, 140 photographs of cars, 70 photographs of cupboards, 70 photographs of chairs and 70 photographs of faces of apes. In all photographs the background was a uniform gray. The height and width of the objects pictured were slightly different between categories, but all took up a similar portion of the frame. The photographs themselves were all the same size, subtending a visual angle of about 3°. See Fig. 1A for examples of the stimuli used.

2.3. ERP recording procedures

The EEG was recorded from 48 tin electrodes mounted on a custom-made cap. Fig. 1B

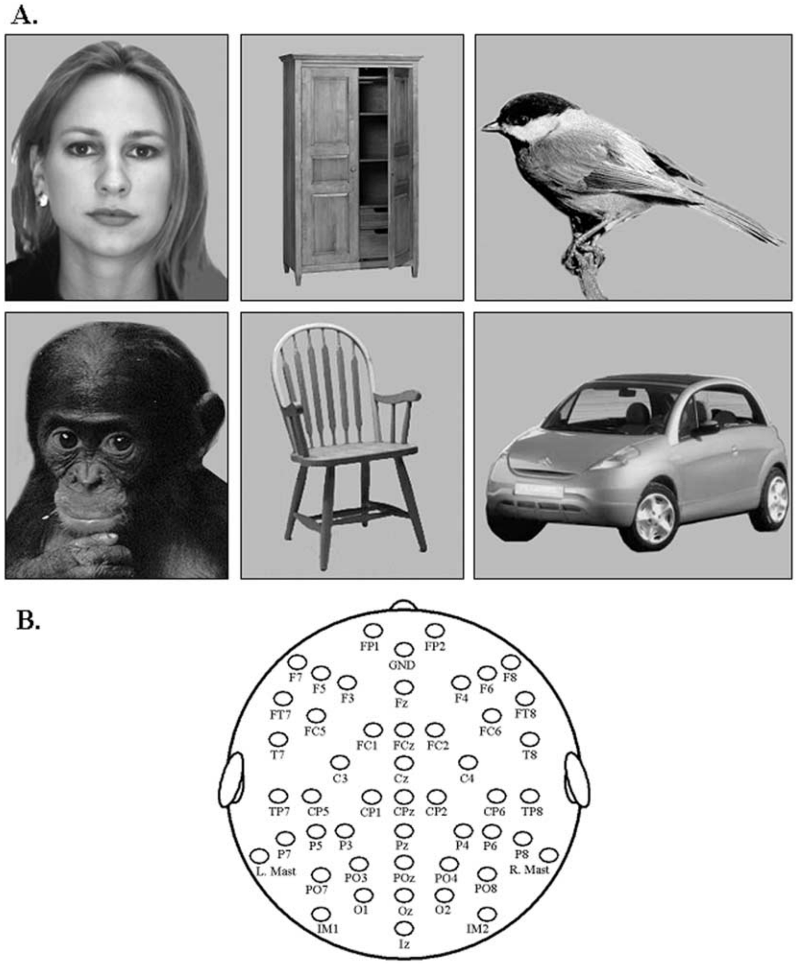

Fig. 1. (A) Examples of the six stimulus types used in the present study. (B) The distribution of the electrode cap that was used.

shows the electrode site distribution. EOG was recorded by two electrodes, one located on the outer canthus of the right eye and the other at the infraorbital region of the same eye. Both the EEG and the EOG were referenced to an electrode placed at the tip of the nose.

The EEG was continuously sampled at 250 Hz, amplified by 20 000 by a set of SAI battery-operated amplifiers with an analog band-pass filter of 0.1–70 Hz, and stored on disk for offline analysis.

ERPs resulted from averaging EEG epochs of 1000 ms starting 100 ms prior to stimulus onset. Average waveforms were computed for each subject in each of the ten conditions (four conditions in the Car Monitoring and in the Animacy Decision tasks, and two in the Human Face Monitoring task), and digitally filtered with a band-pass of 0.5–22 Hz. Trials contaminated by EOG and/or EEG artifacts were excluded from the average by an automatic rejection algorithm with a threshold amplitude of ±100 mV. No ERP was based on less than 50 trials.

The statistical analysis was based on within-participant factorial models in which the N170 amplitude and latency were dependent variables. These measures were defined as the amplitude and latency of the most negative peak between 120 and 210 ms from stimulus onset. Amplitudes were measured relative to the pre-stimulus onset baseline. Eight electrode sites were selected a priori for statistical analysis, based on previous studies (e.g. Bentin & Deouell, 2000). Indeed, as in those studies, at these sites the N170 was most conspicuous. The sites were R. Mast (as in Fig. 1B), PO8, P8 and IM2 on the right hemi-scalp, and the corresponding sites on the left (see Fig. 1B).

3. Experiment 1

N170 experiments aimed at exploring the domain specificity of face processing were criticized for comparing stimuli perceived at different levels. Specifically, whereas faces are usually recognized at the subordinate level of the person's identity, non-face stimuli such as houses and animals are typically recognized at the basic (categorical) level (Rossion et al., 2000; Tanaka & Gauthier, 1997). As evidence for the possible confounding effect of categorization level, it is pointed out that prosopagnosic patients are affected by manipulations of levels of categorization more than normal controls (Gauthier et al., 1999). However, this criticism does not take into account findings demonstrating that detection and identification of faces are mediated by separate systems (Bentin & Deouell, 2000; de Gelder & Rouw, 2000; Eimer, 2000). These findings suggest that the N170 is associated with basic level categorization (seeing the shape of a face as a face), whereas the subordinate level categorization probably takes place later, and modulates other ERP components (N400F). Nonetheless, the face specificity assumption and its reflection in N170 would be further reinforced by showing differential effects of task manipulations on the N170 elicited by faces and those elicited by other object categories (cf. Rossion et al., 2000).

Most studies of N170 used an oddball paradigm in which faces were one of several distracter categories. Indeed, in one study in which ERPs to faces were recorded while participants made categorical (animate/inanimate) decisions to faces as well as to other stimulus categories, face specificity apparently disappeared (Johnson, Schmidt, & Barnhardt, 1999). Although a deeper analysis of these data revealed an important methodological difference between Johnson et al.'s study and those in which face-specific N170 potentials were found,[4] it is still possible that the face-specific N170 effects are induced by task-related factors rather than by stimulus category specificity. For example, seen among different distracters in an oddball paradigm, faces could attract reflexively more attention

---

[4] In the Johnson et al. (1999) study ERPs were recorded relative to a linked-mastoids reference. Because the mastoids are sites where lower posterior-temporal activity is most conspicuous, this procedure eliminated the specificity effect of the N170. To avoid the problem most N170 studies used the tip of the nose as a common reference, or re-referenced the data to an average of all recorded sites. Indeed, subtracting the ERPs recorded at the mastoid sites in our own data significantly reduced the N170 categorical specificity.

than other objects and consequently the negativity elicited by faces between 100 and 200 ms would be larger, selectively enhancing the N170 amplitude (e.g. Mangun, Hillyard, & Luck, 1992). However, in a categorical decision task in which every stimulus needs to be deeply processed, such attention-based differences should be less consequential. Therefore, in the present experiment we compared (within participants) the ERPs elicited by faces and by other object categories in the traditional oddball paradigm with those elicited by the same stimuli in a categorical (semantic) decision task.

In the oddball (car monitoring) task the participants were presented with photographs of four equally probable stimulus categories (human faces, birds, cars and furniture items), and instructed to press a button each time they saw a car. The cars' status as targets meant that they received more strategically allocated attention than the distracter categories (which should each have been allocated equal amounts of attention before being ignored). Note that, in principle, this task requires only "low-level", shape-related categorization. In the categorical (animacy) decision task the participants viewed photographs of exemplars of the same four categories,[5] and were required to categorize each stimulus as animate or inanimate. Note that this decision requires deeper (semantic) processing than the oddball discrimination, as within each two-category group (faces + birds and cars + furniture) there is very little similarity of low-level features. Again, all stimuli had to receive attention in order for a decision to be made, but in this task no category was supposed to receive more attention than the others.

If the N170 were face-specific, we would expect it to differ from that elicited by all other categories in the semantic categorization as well as in the oddball task. Non-specific, task-related effects on this component, should they exist, should reduce the difference between faces and all other categories in the semantic categorization relative to the oddball task. The task should also influence the ERPs elicited by categories other than faces. Attention-related effects may enhance the N170 elicited by cars, the target category, in the oddball task. Finally, if depth of processing has an effect on the N170, we should observe its consequences similarly on all stimulus categories.

3.1. Method

3.1.1. Tasks and design

3.1.1.1. Car monitoring This was a standard oddball paradigm in which the participants were instructed to press a button with their right hand every time they detected a photograph of a car, and ignore other stimulus types. The participants were shown a random succession of 280 photographs, 70 from each of four categories: human faces, birds, cars and chairs. These stimuli were presented in two blocks of 140 photographs each, separated by a break, the length of which was determined by the participant.

3.1.1.2. Animacy decision In this (semantic) decision task the participants were instructed to press a certain button if the picture they saw was of a living being, and a different button if the picture was of a non-living object. Both reactions were made with the right hand. The

---

[5] The exemplars used in this task were different from those used in the oddball task.

stimulus categories (with different exemplars) and trial succession were similar to those in the car detection task, except that instead of chairs, cupboards were used. The order in which the two tasks were presented was counterbalanced across participants.

3.1.2. Procedures

The experiment was administrated in a dimly lit, electrically isolated and sound treated chamber. The participants sat on a comfortable chair viewing a computer monitor placed at eye level about 1.2 m in front of them. At the beginning of each task, instructions appeared on the screen. The instructions were also read aloud to the subject by the experimenter. Each trial started with the appearance of a white fixation point at the center of the dark screen. The participants were requested to direct their gaze at the fixation point and refrain from further eye movements. The fixation point was replaced by a stimulus, which was exposed either until a response was delivered, or (if no response was delivered, and in tasks where no response was required) for 350 ms. The fixation point re-appeared on the screen immediately at the offset of the stimulus and remained there during an inter-stimulus interval (ISI) that ranged randomly between 450 and 550 ms. The ISI count began either immediately after the participant's response, or (again, in trials where no reaction was required or received) 1500 ms from stimulus onset.

3.2. Results

Faces elicited a conspicuous N170 with a posterior temporal distribution, larger at the right than the left hemisphere sites; this effect was conspicuous in both the car monitoring and the animacy decision tasks (Fig. 2A,B, respectively). As evident in Fig. 2, the ERPs elicited by other stimulus categories also included a negative (or negative-going) component that peaked later than the face N170; however, the amplitude of this component was considerably smaller for all non-face categories. The negative component elicited by cars (the pre-determined target in the oddball task) was larger than for birds and furniture. Relative to the face N170 its distribution seemed to be more medial and suggested a more posteriorly oriented source (Fig. 2A). Indeed, as demonstrated in Fig. 3, whereas the N170 elicited by faces was larger at the posterior lateral sites, the negative component elicited by cars in the Car Monitoring task, and the difference between that component and the component elicited by cars in the Animacy Decision task were larger more medially, at O1 and O2.

The statistical reliability of this pattern was tested initially by a four-factor ANOVA of the N170 amplitude, with repeated measures. The factors were Task (car monitoring, animacy decision), Stimulus Type (faces, birds, cars, furniture), Hemisphere (left, right) and Site (P7/8, Mastoids, PO7/8, IM1/2). This ANOVA showed that both the task and stimulus type effects were significant ($F(1,11) = 5.1$, $P < 0.05$ and $F(3,33) = 23.7$, $P < 0.001$, respectively) and, more importantly, that these effects significantly interacted ($F(3,33) = 3.4$, $P < 0.5$). In light of this interaction we continued the analysis of the stimulus type effects separately for each task by three-factors ANOVAs with repeated measures, excluding the task factor (Table 1).

The separate ANOVAs showed that in both tasks, the stimulus type effect was significant ($F(3,33) = 8.2$, $P < 0.001$ and $F(3,33) = 24.3$, $P < 0.001$ for the car monitoring

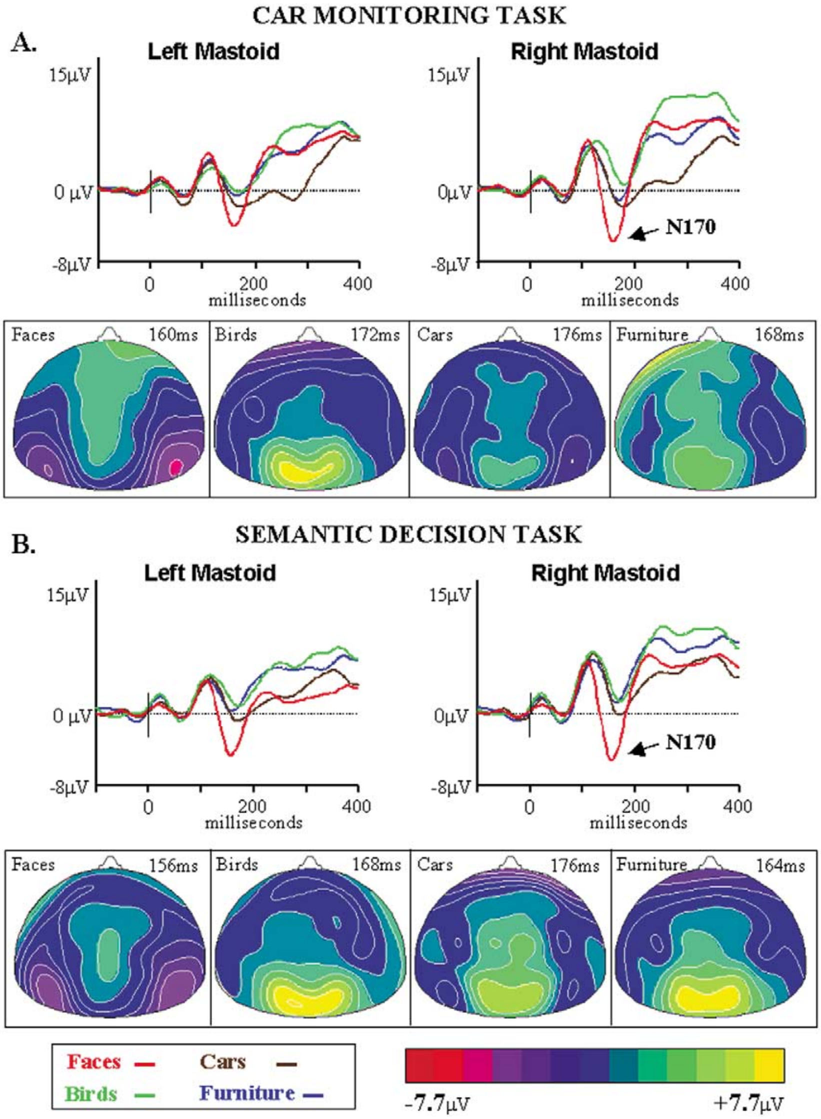

Fig. 2. The ERPs elicited at the left and right mastoids, and their scalp distributions at their respective negative peaks (between 140 and 240 ms) for each of the four stimulus categories in the two tasks of Experiment 1. (A) Car Monitoring task: note the distinct N170 for faces, and the increased negativity of cars relative to birds and furniture. In the scalp distributions, note that the activation for cars differs from that of birds and furniture, though its points of greatest negativity are slightly more medial than those of faces, and do not reach the same magnitude. (B) Animacy Decision task: the difference between faces and all other stimulus categories is evident both in the distinct N170 and in the scalp distributions. Note that the scalp distribution for cars, which were not a target in this task, closely resembles that of birds and furniture.

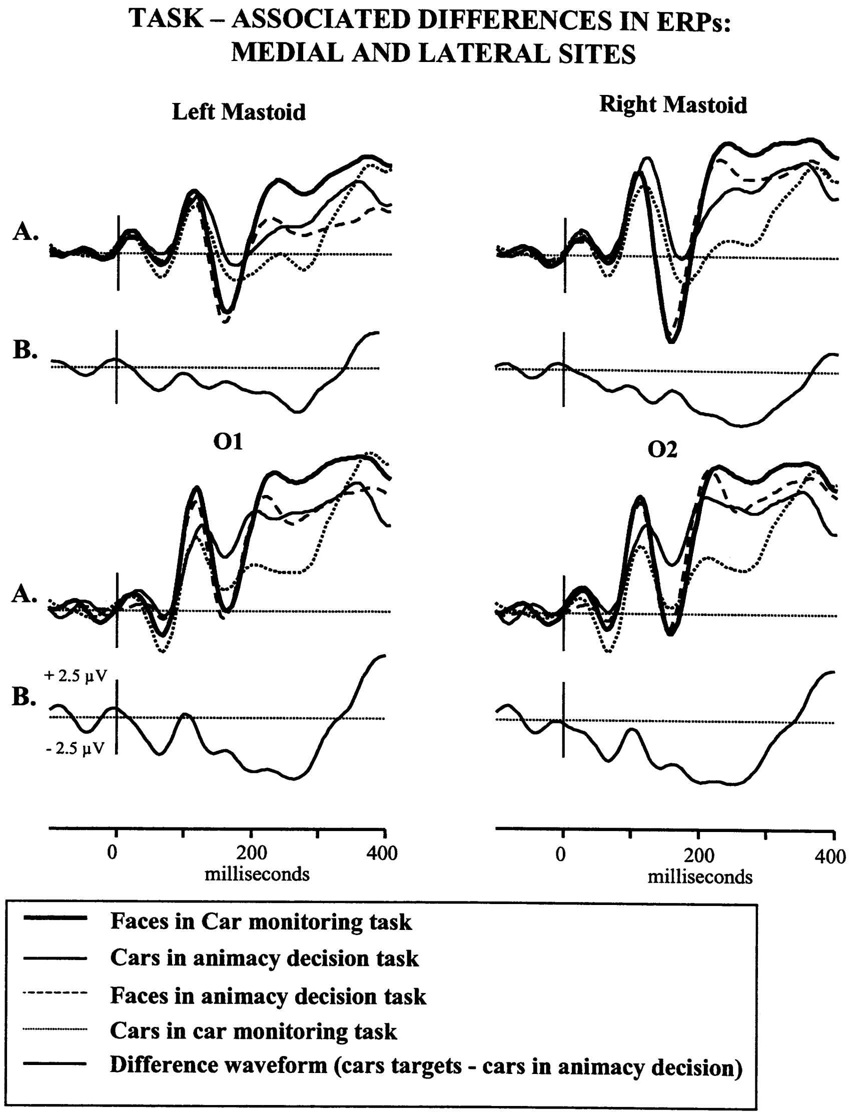

Fig. 3. (A) Comparison of faces and cars in the two tasks of Experiment 1. Note that whereas faces elicited almost identical N170s and generally similar waveforms in both tasks, cars elicited a much more negative ERP in the Car Monitoring task than in the Animacy Decision task. The difference in the activation for cars is greater in medial (O1 and O2) than in lateral sites (Left and Right Mastoids). (B) Difference waveform obtained by subtracting the potential for cars in the Animacy Decision task from the potential for cars in the Car Monitoring task (in which cars were targets). The difference is greater in the medial sites, and peaks at about 270 ms.

and animacy decision tasks, respectively). Post-hoc univariate comparisons revealed, however, that whereas in the animacy decision task the N170 elicited by faces was larger than that elicited by all other stimulus types ($P < 0.001$ for all these comparisons), in the

Table 1
Amplitudes (in μV) of the ERPs elicited by faces, birds, cars and furniture at the posterior temporal scalp sites in the Car Monitoring task and the Animacy Decision task

	Left hemisphere				Right hemisphere			
	L. mast	P7	PO7	IM1	IM2	PO8	P8	R. mast
Car Monitoring task								
Faces	−4.6	−3.9	−3.4	−4.2	−4.6	−4.6	−5.8	−6.5
Birds	−1.4	−1.4	0.6	0.1	0.6	1.0	−0.7	−0.5
Cars	−2.8	−3.7	−2.7	−2.2	−2.4	−2.8	−3.9	−3.0
Furniture	−1.5	−2.2	−1.1	−0.4	−1.4	−2.2	−3.3	−2.2
Animacy Decision task								
Faces	−5.8	−5.9	−4.7	−5.2	−4.6	−4.5	−5.9	−6.4
Birds	−0.4	0.3	2.5	1.7	1.9	2.4	0.2	0.4
Cars	−1.3	−0.6	0.4	0.1	−0.3	0.3	−1.2	−0.8
Furniture	−0.5	−0.7	1.3	0.8	1.6	1.7	0.0	0.3

car monitoring task, a significant stimulus by site interaction ($F(9, 99) = 2.9$, $P < 0.01$) was found, and explained by the fact that the N170 elicited by cars was significantly smaller than that elicited by faces ($P < 0.05$) at the mastoids, but not across all sites ($P = 0.1$). However, even in the car monitoring task, whereas the N170 elicited by faces was significantly larger than that elicited by birds ($P < 0.001$) and furniture ($P < 0.05$), the negative potential elicited by cars was not significantly different than that elicited by furniture ($P = 0.28$), though it was larger than that elicited by birds ($P < 0.05$). Although, at least for faces, the N170 seems to have been larger over the right then the left hemisphere, the inter-hemispheric difference was not significant in either task ($F(1, 11) < 1.00$, in both tasks). In addition to the stimulus type effect, a significant site effect was found on both tasks ($F(3.33) = 3.4$, $P < 0.05$ and $F(3, 33) = 4.5$, $P < 0.01$ in the car monitoring and animacy decision tasks, respectively). Post-hoc analyses revealed that in both tasks this effect reflected larger amplitudes of the negative potential at the P7 and P8 and at the mastoids than at the PO7, PO8, IM1 and IM2. Furthermore, a significant interaction between stimulus type and site in the car monitoring task ($F(9, 99) = 2.9$, $P < 0.01$) and a similar trend (albeit not significant) in the animacy decision task revealed that only faces accounted for the site effect.

Whereas the enhancement of the N1/N170 elicited by cars was expected in the car monitoring task (due to the special status of this category, which might have attracted more attention), it was not expected in the animacy decision task. Therefore, in a post-hoc attempt to explain the apparent differences between the negative component elicited by cars and that elicited by the other two non-face stimulus categories in the animacy decision task, we explored whether this small effect might be accounted for by some residual effects from the car monitoring task. In order to explore this account we analyzed the ERPs elicited by the six participants who were examined with the animacy decision task prior to the car monitoring task separately from those that began the experiment with the car monitoring task. As demonstrated in Fig. 4, the "car-effect" in the animacy decision task was entirely accounted for by the task order: participants who were tested first with the animacy decision

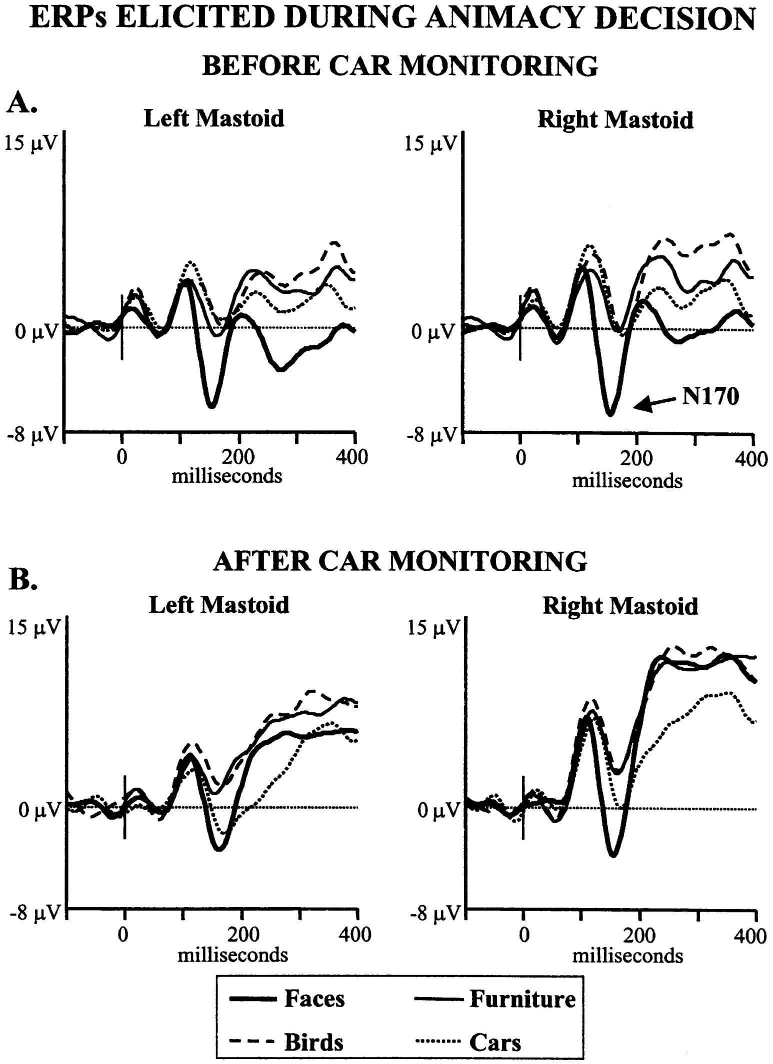

Fig. 4. The effect of task order on the results of the Animacy Decision task. (A) Participants who performed this task first: note that the N1 elicited by cars did not differ from those elicited by birds and furniture. (B) Participants who performed this task after the Car Monitoring task: note that the negativity for cars at the same latency was much greater, almost reaching (in the Left Mastoid) the same magnitude as the face N170. This probably reflects a residual attentional effect, and explains the greater negativity for cars in this task over all subjects (see Fig. 1).

task showed a clear face N170 effect, but cars were not distinct relative to the other non-face categories during the relevant time epoch. In contrast, cars elicited a distinct negative-going component (albeit significantly smaller than that elicited by faces) in the animacy decision task in participants who had processed cars as a distinct (target) category beforehand.

Finally, the effect of task on cars, the target category in the oddball paradigm, was

Table 2
Latencies (in ms) of the ERPs elicited by faces, birds, cars and furniture at the posterior temporal scalp sites in the Car Monitoring task and the Animacy Decision task

	Left hemisphere				Right hemisphere			
	L. mast	P7	PO7	IM1	IM2	PO8	P8	R. mast
Car Monitoring task								
Faces	163	162	162	162	161	163	163	162
Birds	171	167	166	169	177	176	181	182
Cars	182	175	173	178	174	174	180	182
Furniture	170	168	168	168	168	168	172	173
Animacy Decision task								
Faces	161	159	159	159	159	160	161	160
Birds	172	165	162	169	169	168	173	174
Cars	180	175	171	178	177	176	183	184
Furniture	167	161	163	165	167	163	167	169

compared to the task effect on faces by a Task (car monitoring, animacy decision) × Stimulus Type (faces, cars) ANOVA. The dependent variable was the average amplitude across all eight recording sites. This analysis showed a significant interaction between the two factors ($F(1, 11) = 6.4$, $P < 0.05$) demonstrating that the task had different effects on faces and cars. Planned comparisons revealed that whereas faces elicited similar N170 components in both tasks ($t(11) = 0.7$, $P = 0.5$) the negative component elicited by cars during the relevant time epoch was significantly larger in the car monitoring task than in the animacy decision task ($t(11) = -3.0$, $P < 0.025$).

The analysis of the N170 peak latencies was similar to that performed on its amplitudes (Table 2). The four-factor ANOVA showed a significant effect of stimulus type ($F(3, 33) = 7.6$, $P < 0.005$), and a significant effect of recording site ($F(3, 33) = 6.5$, $P < 0.005$). No other main effects were significant. Post-hoc univariate analysis revealed that the main effect of stimulus type was caused by a significantly shorter latency of the N170 elicited by faces (161 ms) than by all other stimulus types (171 ms for birds, 178 ms for cars, and 167 ms for furniture) ($P < 0.005$, $P < 0.01$, and $P < 0.05$ for faces compared with birds, cars, and furniture, respectively).

3.3. Discussion

The main outcome of the present experiment is that ERPs elicited by faces include a negative potential (N170) that is distinct in scalp distribution and immune to strategic manipulation. The ERPs elicited by other stimulus categories may also show an apparently similar deflection during the time range of the N170, between 100 and 200 ms, but this negative (or negative-going) deflection is probably associated with the general perceptual activity reflected in the visual N1 component.[6] This outcome supports the face-specific visual process hypothesis.

---

[6] We will address the distinction between N1 and N170 in Section 5.

The face specificity hypothesis has been recently challenged by findings showing that the N1 may sometimes distinguish between different non-face categories (Rossion et al., 2000; Tovee, 1998a; see also Tovee, 1998b). The present data addressed this challenge by suggesting that factors that occasionally induce differences between the N1 elicited by different non-face stimulus categories are associated with the task rather than with the stimulus type. Two such factors were examined in Experiment 1: the level of processing required for the task completion, and the amount of attention allocated to the various stimulus categories. Whereas the level of processing required for the categorization in the car monitoring task was relatively shallow, based primarily on visual features analysis, animacy decisions required a deeper, semantic analysis. Apparently, the level of processing did not influence the distinctiveness of the negative N1/N170 across categories. The N170 elicited by faces was nearly identical across tasks, and similarly distinct from that elicited by birds and furniture. Furthermore, despite the fact that birds and furniture were assigned to the same response category in the car monitoring task but to different categories in the animacy decision task, the two stimulus types elicited similar ERPs, at least up to about 250 ms from stimulus onset. This similarity supports the idea that semantic meaning and, perhaps, within-category identification of visual stimuli modulate components elicited later than the N1 or N170 (Bentin & Deouell, 2000; see also Johnson et al., 1999).

In contrast to the level of processing, attention can apparently modulate responses associated with visual processing, leading to categorical distinctions. This effect was clearly evident in the negative component elicited by cars, which was distinct from birds and furniture in the car monitoring task, and significantly reduced in amplitude in the animacy decision task. A possible explanation of this task effect is the task relevance requirement for the elicitation of the P300 (Donchin, 1981; Donchin & Coles, 1988). Task relevance, however, should have also influenced the response to birds and furniture (and perhaps even to faces). Whereas in the car monitoring task all non-car categories were, by definition, not relevant, each of these stimulus categories was task-relevant in the animacy decision task. Nonetheless, neither faces nor furniture or birds had larger negative components in the latter than the former task. Indeed, the amplitudes of the N1/N170 for all stimulus categories tended to be smaller (less negative) in the animacy decision than in the car monitoring task.

The other factor that could have been modulated across tasks was attention. In the oddball paradigm, cars were probably more salient than in the animacy decision task because of their special status as task-defined targets. Due to this special salience, cars might have attracted more attention in the former than in the latter task. This idea was strongly supported by the differential distinction between cars and the two other non-face categories observed in the animacy decision task for participants who performed this task after the car monitoring. The absence of such distinction in naive participants (i.e. those who performed the animacy decision prior to car monitoring) suggests that it was induced by a residual influence of the oddball task. It is most likely that some of the salience of cars in the first task was transferred (carried over) to the second.

The putative influence of attention on the negative component elicited by cars supports the face specificity of the N170. A possible account of the N170 specificity for faces is that faces might automatically attract more attention than other stimulus categories, enhancing the amplitude of the N1. Although the present data did not exclude this possibility, it made it unlikely. If attention were the only (or even the major) factor making the N170 elicited by faces conspicuous relative to that elicited by cars, birds and furniture, the implication would be that even in the car monitoring task faces were more conspicuous than cars. Furthermore, although the role of attention in face processing is far from clear, the burden of current evidence does not support the view that faces attract attention reflexively (e.g. Brown, Huey, & Findlay, 1997; but see Cauquil, Edmonds, & Taylor, 2000).

Special processing for faces as a unique characteristic of the visual system has been challenged also by assuming special expertise in processing faces relative to other stimulus categories to which the N170 has been usually compared (e.g. Gauthier et al., 2000; Gauthier, Tarr et al., 1999). The argument is that lifelong experience has made us all experts at processing human faces in comparison with other categories. Therefore, it may be our level of expertise with faces, and not a face-specific system, that is the source of the N170's outstanding reaction to faces. We addressed this challenge in Experiment 2.

4. Experiment 2

The present experiment was designed to address the two interpretations of the distinct ERP manifestation of face processing, i.e. the assumption of expertise for faces and the attentional account. As discussed in Experiment 1, both accounts have been offered as alternatives to the domain specificity hypothesis.

Attempts to support the view that expertise may change the manner in which visual stimuli are processed, and that this change may eliminate the neural processing specificity for faces as reflected in neuroimaging as well as electrophysiological studies, have been focused on exploring the indices of neural activity elicited by non-face stimulus types with which participants became very familiar, compared with those elicited by faces. The rationale of this approach was that if expertise with non-face categories results in the abolishment of the difference between faces and other stimulus types with which the participants are expert, then there is no face-specific system, but rather a system for processing categories for which a certain degree of expertise has been reached. Behavioral evidence in support of this possibility includes the finding that visual expertise with a given dog breed (Diamond & Carey, 1986) or artificial objects (greebles) (Gauthier & Tarr, 1997) can lead to inversion and/or superiority effects similar to those obtained for faces. Neuroimaging evidence supporting this hypothesis was obtained by showing that expertise training with greebles recruited the FFA to nearly the same degree as faces do (Gauthier, Tarr et al., 1999), and that the FFA of car and bird experts was activated in response to these categories in a similar way (Gauthier et al., 2000). Finally, a recent study suggests that N170 can also be influenced by expertise. Tanaka and Curran (2001) reported a significant interaction between the amplitude of the N170 elicited by birds and dogs and the expertise of the participants as dog breeders and bird watchers.

These data are important in showing that brain areas that are naturally involved in processing faces can be invoked to process other visual stimuli. This may happen if, once becoming an expert, the perceiver processes such stimuli similarly to the manner in which he/she processes faces (e.g. using holistic perception strategies). Such results, however, do not preclude the possibility that faces comprise specific characteristic features that trigger the "expert system" automatically, *regardless of expertise*. In the present experiment we examined this possibility by comparing the N170 elicited in humans by human faces to that elicited by faces of apes. Whereas the configuration of ape faces is similar to that of human faces, the shape of their components (particularly the nose) and the addition of hair make them different enough to reduce the possibility that regular humans are expert in processing faces of apes. If the expertise accounts for face specificity, the N170 elicited by human faces should be clearly distinguished from that elicited by ape faces, much as has been found when comparing human faces with faces of other animals (Bentin et al., 1996). However, if face specificity is based on the existence of particular characteristics of the face, which are common to faces of apes and humans, the N170 should be similar for these two stimulus categories. An empirical foundation for this logic has recently been provided by a study in which face inversion effects for human faces, monkey faces and faces of sheep were examined in 5–8-year-old children (Pascalis, Demont, de Haan, & Campbell, 2001). In complete agreement with our rationale, across face orientation, recognition was better for human than for monkey faces (suggesting less expertise for the latter than the former category). More importantly, inversion had a similar detrimental effect on both types of primates, but no effect on the recognition of sheep. This pattern suggests that humans used a similar perceptual (holistic) strategy for processing monkey as well as human faces, but not for processing sheep.

In addition to exploring the expertise assumption, the design of Experiment 2 allowed a further investigation of the influence of task on processing human faces, particularly that of task relevance. In Experiment 1, we found that salience enhanced the N1 elicited by cars and attributed this effect to attentional factors. In contrast, we claimed that attentional factors had less (if any) effect on faces. However, whereas special attention to cars was explicitly required in the car monitoring but not in the animacy decision task, faces had no special status in either of these tasks. To examine the role of strategically directed attention on the N170 elicited by faces, in the present experiment only human faces and ape faces were presented, and the former category was pre-defined as a target. The participants were instructed to press a button each time a human (rather than ape) face was presented. If the neural/perceptual mechanism with which the N170 is associated is the same as that which was modulated by stimulus salience in response to cars, we should observe not only a difference between the N170 elicited by faces and apes in this experiment, but also an enhanced N170 to faces in the present experiment relative to those elicited by faces in Experiment 1.

4.1. Method

4.1.1. Task

4.1.1.1. Human face monitoring A series of 70 new human faces (all unfamiliar to the participant) were presented intermingled with 70 faces of different apes. The participants were instructed to press a button with their right hand whenever they saw a human face, and ignore the apes. The structure of each trial was similar to that of the car detection task of Experiment 1. All participants performed Experiment 2 after Experiment 1.

4.1.2. Procedures

The procedures used in this experiment were identical to those used in Experiment 1 in the Car Detection task.

4.2. Results

The N170 elicited by ape faces in human participants was as conspicuous as that elicited by human faces in the same participants, albeit peaking later (Fig. 5).

The statistical reliability of the observed pattern was examined by ANOVA with repeated measures on the peak amplitudes and latencies of the N170 (Table 3).

In both analyses the factors were Stimulus Category (human faces, ape faces), Hemisphere (right, left) and Site (P7/8, Mastoids, PO7/8, IM). The analysis of the amplitudes showed no effect of stimulus category ($F < 1.0$), no effect of hemisphere ($F(1, 11) = 2.1$, $P = 17.5$), a significant main effect of site ($F(3, 33) = 8.6$, $P < 0.001$), and a significant interaction between stimulus type and site ($F(3, 33) = 10.6$, $P < 0.001$). Post-hoc analysis of the site effect showed that, across tasks, the N170 was larger at the more lateral sites (mastoids and P7/8), than at the more medial sites (PO7/8 and IM1/2) ($F(1, 11) = 14.6$, $P < 0.001$), and larger at the IM1/2 than at PO7/8. No other effects were significant. Post-hoc analysis of the interaction showed that, although the size of the stimulus type effect varied across sites, it was not significant at any site.

Because in the present experiment human faces were targets and ape faces were distracters, we also ran a similar ANOVA comparing the amplitudes of the N170 elicited by apes in the present experiment with the amplitudes elicited by human faces in the Car Monitor-

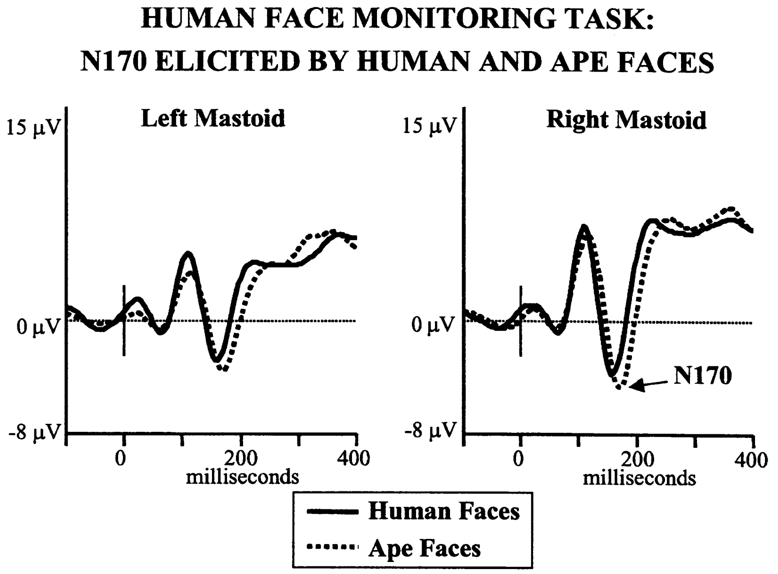

Fig. 5. The N170s elicited by human and ape faces are very similar. The negativity for apes is slightly greater, and is delayed by 10 ms relative to human faces.

Table 3
Amplitudes (in μV) and latencies (in ms) of the N170 elicited by human and ape faces at the posterior temporal scalp sites

Site	Amplitudes				Latencies			
	Left hemisphere		Right hemisphere		Left hemisphere		Right hemisphere	
	Human	Ape	Human	Ape	Human	Ape	Human	Ape
Mastoid	−3.8	−4.3	−5.2	−6.0	161	173	161	174
P8/7	−3.7	−4	−4.1	−4.6	162	169	163	174
PO8/7	−2.2	−1.0	−3.2	−2.5	161	170	161	170
IM2/1	−3.5	−3.4	−3.7	−4.4	160	172	160	171

ing task (Experiment 1). In both these tasks, the faces (human and apes) had identical task relevance. This ANOVA reveled no significant effect of stimulus type ($F(1, 11) < 1.0$), and there was a significant effect of site ($F(3, 33) = 12.4$, $P < 0.001$) and a marginally significant effect of hemisphere ($F(1, 11) = 4.6$, $P = 0.055$), suggesting that the N170 tended to be larger over the right than the left hemispheres. None of the interactions were significant.

The analysis of latencies showed that the N170 elicited by ape faces (172 ms) peaked significantly later than that elicited by human faces (161 ms) ($F(1, 11) = 32.6$, $P < 0.001$). No other main effects were significant (all $F < 1.00$). The stimulus type effect interacted significantly with site ($F(3, 33) = 5.3$, $P < 0.005$), revealing that the difference between apes and humans was larger at the mastoids (13 ms) than at all other sites (9–11 ms). No other effects were significant.

As for amplitudes, also for latencies we repeated the ANOVA design comparing human faces from the Car Monitoring task with ape faces from the Human Face Monitoring task. This ANOVA revealed a marginally significant effect of stimulus ($F(1, 11) = 4.7$, $P = 0.053$), and a significant hemisphere by site interaction ($F(3.33) = 4.2$, $P < 0.05$). No other effects were significant.

Finally, in order to assess possible effects of task relevance on the N170 elicited by faces, we compared its amplitude and latency across the three tasks (Fig. 6). Recall that in the car monitoring tasks faces were distracter stimuli that required no deep analysis; in the semantic decision task faces were processed for meaning, but they had no special significance relative to the other three categories; in the human face detection task human faces were the target stimuli. If strategic factors influenced the brain activity associated with the N170 we should see a modulation of the N170 across tasks.

One-way ANOVA (using the mean amplitude and latency across all recording sites as a dependent variable) found that both the amplitude and the latency of the N170 elicited by faces showed no significant differences across the three tasks (both $F < 1.0$).

4.3. Discussion

The major outcome of the present experiment was that ape faces elicited an N170 component that peaked 10 ms later, but was as large as that elicited by human faces.

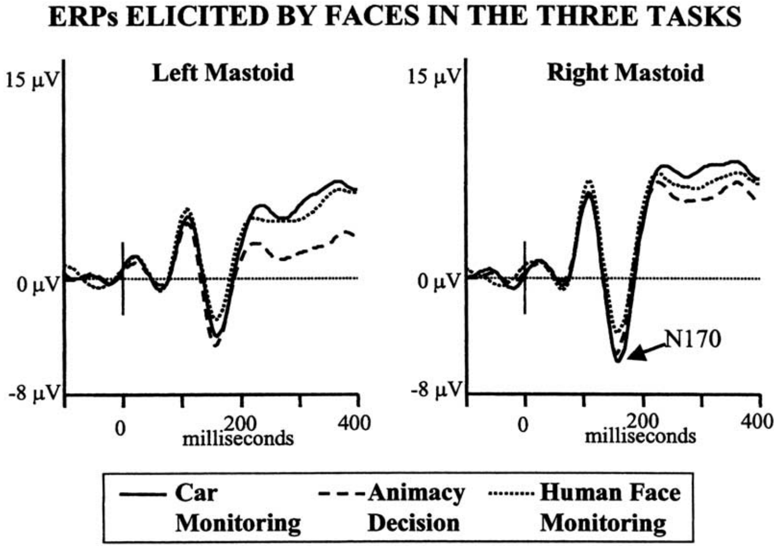

Fig. 6. A comparison of the N170s elicited by human faces across all three tasks demonstrates that the N170 is immune to task manipulations, being similar regardless of the demands the task placed on the system. This constancy could serve as a marker for the system's domain specificity.

Furthermore, compared with the ERPs elicited by non-face categories in the same participants (see Experiment 1), the N170 elicited by ape faces was as distinctive as that elicited by human faces. This outcome contrasts with the absence of a discriminative N170 in response to faces of other animals, such as dogs, cats and birds (Bentin et al., 1996). Because the participants examined in this experiment were not experts in processing faces of apes, the implication of this pattern is that the perceptual mechanism associated with the N170 is discriminatively engaged by human faces and by other stimuli that preserve perceptual characteristics of the human face, even if the particular face category is not frequently encountered. As suggested by the contrast between the similarity of inversion effects for monkey and human faces and the absence of such an effect for sheep faces (Pascalis et al., 2001), this pattern reflects a similarity in the representational template (or processing strategy) for recognizing human and ape faces (Campbell, Pascalis, Coleman, Wallace, & Benson, 1997). A possible implication of this pattern for the specificity/expertise debate is that the expertise effect on the N170 observed in groups of dog breeders and bird watchers (Tanaka & Curran, 2001) indeed reflected the contribution of other mechanisms, for example attentional factors involved in the modulation of the visual N1. We will discuss and extend this hypothesis in Section 5.

The comparison of the N170 evoked by faces across the three tasks revealed no significant differences in amplitude and latency between tasks. This similarity indicates that unlike cars, task-associated strategies or strategically controlled allocation of attention have little, if any, effect on the face N170. The distinction between the task effects on faces and cars points again to the specificity of the visual face processing mechanism and contributes to its characterization. On the one hand, its specificity implies that the stage of structural encoding reflected by N170 is relatively late (low-level feature analysis must be completed, and the realization that the viewed shape has the characteristics of a face must be reached). On the other hand, the absence of a depth of processing effect suggests that this ERP component is associated with the formation of a perceptual category rather than with the elaboration of its meaning and semantic connotations. This is not to say, however, that this perceptual process is immune to top-down influences. For example, recent fMRI (Dolan et al., 1997), and electrophysiological (Bentin et al., in press) data demonstrated that priming could modulate face-specific brain activity.

Finally, an additional implication of the similarity between the processing of human and ape faces is that the system whose activity is reflected by the N170 applies a great deal of generalization: it is face-specific, but not species-specific. Although this generalization does not preclude a developmental approach to face specificity in humans (e.g. Nelson, 2001), it suggests that the tuning of extra-striate mechanisms to processing physiognomic information, though not necessarily innate, is not rigidly constrained by experience. This view can also explain why schematically drawn faces attracted attention in infants who had no opportunity to learn these stimuli before testing (e.g. Johnson & Morton, 1991), and were sufficient to elicit (in adults) an N170, as distinctive as that elicited by photographs of natural faces (Sagiv & Bentin, in press).

The two categories did differ slightly but significantly in latency: the N170 was delayed by about 10 ms for ape relative to human faces. This modulation is similar to that found for inverted human faces (Bentin et al., 1996; Eimer, 2000; Rossion et al., 2000; Sagiv & Bentin, in press). It may indicate that, although the face processing mechanism associated with this component may be tolerant to deviations from the natural aspect of the face characteristics, it requires more time to respond to faces presented in a manner differing from the "optimal configuration" of a face.

5. General discussion

The present study was designed to address a recent debate regarding the domain specificity of the extra-striate visual mechanism implicated in face processing, by exploring its sensitivity to task/strategic manipulations and the observer's expertise. Our approach was to explore the influence of these factors on the N170, a face-specific ERP component. The results of this investigation were clear. Whereas task manipulations did affect a negative component elicited by cars within the same time range as the N170, similar manipulations did not significantly change the N170 elicited by faces. Indeed, across three different tasks, regardless of whether faces had to be ignored (in the Car Monitoring task), were equally task-relevant to other stimulus categories (in the Animacy Decision task), or were the target stimuli (in the Human Face Monitoring task), neither the amplitude nor the latency of the N170 were significantly different. Further, the similarity between the N170 elicited by human faces (with which human observers are experts by necessity) and those elicited by ape faces (with which our participants were not particularly familiar) and even more so, the similar distinctiveness of these components from the ERP activity elicited by other stimulus categories, raise serious caveats regarding the concept that expertise is a major factor modulating the N170, and, by implication, a major factor determining the domain specificity of face processing.

Given the relatively large variability in the manifestations of most (if not all) measures of brain activity, it is conceivable to find various patterns of results across studies even if most (or even all) experimental parameters are constant. These circumstances require clear criteria for accepting or rejecting a claim about the specificity of any brain measure and its link to a particular underlying brain mechanism or process. Several criteria for processing specificity have been suggested in the literature. In single-unit studies the criterion usually adopted is the intensity (frequency) of the neuronal response. For example, some authors suggested that if, in response to a particular stimulus, the response frequency of a cell is at least doubled relative to spontaneous firing or to the response to other (control) stimuli, the cell is assumed to be specific, or at least selective (for such a claim made in the context of face specificity and its actualization in the IT see Rolls & Baylis, 1986). No similar criteria have been suggested for fMRI or ERP components. This is probably because the latter measures of neuronal activity are less accurate, summing the responses of large populations of cells, which most probably represent more than one perceptual mechanism. Hence, the generally accepted criteria rely on statistical evaluation of the differences between responses to stimulus categories, i.e. on the reliability of the observed patterns. An attempt to set additional criteria (at least for ERPs) was made recently (Rossion et al., 2000). These welcome additions include the necessity to show that similar (statistically significant) distinctions cannot be found among other categories (for which the particular component is not supposed to be specific) and that experimental manipulations have distinctive effects on the response to the selected category relative to the response to other (non-selective) categories. Both these criteria were met by the face-elicited N170. No differences were found between the ERPs elicited by birds and furniture within the relevant time range, while both were significantly distinct from faces in Experiment 1. This comparison is particularly pertinent because in both the Car Monitoring and Animacy Decision the task relevance of these three stimulus categories was identical. The same can be said regarding the similarity between the N170 elicited by apes and humans. Recall that in addition to the within-task comparison of these two categories we also found no difference between ape faces in Experiment 2 and human faces in the Car Monitoring task, i.e. when the two categories were equally task-irrelevant. Hence, the criterion requiring that in order to be specific to faces the N170 should distinguish faces from other stimulus categories, but not among other categories (under similar experimental circumstances) is met.

The ERPs elicited by cars, but not by faces, were affected by task requirements. This distinctiveness was evident when comparing task-related effects on the ERPs to cars in the Car Monitoring task (where cars were the target category) versus Animacy Decision (where cars had no special status), with the absence of task-related effects on the N170 elicited by faces in Experiment 1 (in which faces had no special role in either task) versus that elicited by human faces in the Human Face Monitoring task (where human faces were targets). Hence, the second criterion suggested by Rossion et al. (2000) was also met. Furthermore, the consistency of the N170 characteristics across tasks and experiments can, by itself, be taken as an additional criterion: a domain-specific component should be relatively immune to task manipulations, at least those that do not compromise the perceptual quality of the relevant stimulus features. A possible alternative interpretation for the absence of a task effect on the N170 elicited by faces is that faces capture attention exogenously, i.e. regardless of task relevance. However, we consider this interpretation to be unlikely for at least two reasons. One is that, as we mentioned in the discussion of Experiment 1 (Section 3.3), performance data do not support the hypothesis that unfamiliar faces pop out in visual arrays (e.g. Brown et al., 1997). The second reason is that the N170 elicited by photographs of natural faces in which the inner components are mislocated is smaller than (or similar to) that elicited by normally configured faces, despite the fact that such stimuli most probably capture attention at least as much as regular faces do. On the basis of all the criteria above we maintain our initial position (Bentin et al., 1996) that the N170 is a face-specific component associated with a domain-specific (faces) perceptual/neural mechanism.

The negative component elicited by cars, particularly in the Car Monitoring task, could have been a source for claims against the face specificity of the N170. Clearly, as discussed above, this claim is falsified by modulation of this component by task, and most impressively by the comparison of participants who were tested first with the Car Monitoring task with those who were tested first with the Animacy Decision task. However, this negative component raises an important issue, which is the relationship between the N170 and other components elicited during the same time epoch, particularly N1. A comprehensive discussion of the N1 and its sources is beyond the scope of the present study (for a review of N1 and its functional significance see, for example, Naatanen & Picton, 1987; Vogel & Luck, 2000). Briefly, the N1 is a modality-specific sensory component modulated by physical features of the stimulus (such as intensity or color), but not by complex features such as shape or meaning. Its latency is also modulated by sensory factors ranging (in the visual system) from about 100 to 110 ms for checkerboard reversal visual evoked potentials to about 200 ms for weaker stimulation of the primary visual cortex. Most importantly, ample evidence demonstrates that the N1 is modulated by selective attention, mainly showing higher amplitudes in response to attended stimuli than in response to unattended stimuli (for a review and elaboration see, for example, Luck, 1998). Inspection of all N170 data reported in the literature shows no temporal distinction between the face-specific N170 and a preceding N1. Although the distribution of N170 is more lateral than that of the N1 (the N170 peaks at the mastoids and midway between the inion and the mastoids (IM1 and IM2) whereas the N1 peaks at O1, Oz and O2), it is very likely that the N170 also reflects activity generated by N1 sources. This confounding may be the source of much confusion and debates in the literature. Indeed, it is very likely that the negative (or "negative-going") components elicited by other stimulus categories and compared with the N170 in normal experimental practice are in fact manifestations of N1.

The present data support the above view. In particular, comparison of the distribution of N170 elicited by faces and that of the conspicuous negative component elicited by cars in the Car Monitoring task shows that whereas the former reaches its maxima at the mastoids, the latter is largest at the occipital sites (Fig. 2). The assumption that the car-elicited negative component is N1 can also account for the modulation of this component by task. Given the different roles of the car category in the Car Monitoring and the Animacy Decision tasks, it is probable that task-related effects were associated with the greater allocation of attention to cars in the former than in the latter task. It is not surprising, therefore, that the N1 elicited by cars was larger in the Car Monitoring than in the Animacy Decision tasks. Additional support for this hypothesis comes from a recent study by Caldera et al. (2001). In that study the N170s elicited by faces of people from the same race (SR) as the observers were compared with those elicited by faces of people from other races (cross-race; CR). Whereas no CR effects were found on the N170 elicited at lateral sites, such effects were significant at more medial sites. Indeed, inverse solutions (LORETA) applied to these components showed that CR faces modulate a more medial occipital source than SR faces. Furthermore, single-unit recordings (in other studies) showed that neurons in the primary visual cortex respond to pop-out manipulation of attention (e.g. Nothdurft, Gallant, & Van Essen, 1999). It is possible that the low-level (configural) process required to detect oddball targets among visually very different distracters had some "pop-out" quality.

Finally, it is time to return to the specificity/expertise debate. The view that there are neural mechanisms specifically tuned to process faces does not exclude the possibility that these mechanisms can process stimulus categories for which there is expertise. However, the present results suggest that the domain specificity of the face processing mechanism does not depend on the expertise of the viewer. Rather, there are apparently stimulus qualities that are necessary and sufficient to activate the face-specific process. Furthermore, evidently these qualities are not very narrowly constrained, and may even include learning and conceptual factors, as demonstrated by the face-specific N170 elicited by schematic faces. In the present study we did not attempt to specify the relevant triggers. Previous studies, however, suggest that face components (particularly the eyes) seem to be sufficient to trigger face-specific processes (Bentin et al., 1996; Eimer, 2000). Yet, other stimulus categories might induce specific processes in the extra-striate regions (Ishai et al., 1999). Indeed, it is conceivable that specific processing strategies (for example, holistic versus part-based processing) are an intrinsic part of domain specificity and if (following training) viewers use the face-peculiar processes while perceiving other stimuli, the face-specific areas are activated (Gauthier et al., 2000; Gauthier, Tarr et al., 1999). Further research, however, is necessary to disentangle the effects caused by genuine activation of face areas and effects that may reflect the contribution of other mechanisms, such as attention. The latter could be an alternative account for the specificity of the N170 to birds in bird watchers as opposed to dogs in dog breeders (Tanaka & Curran, 2001).

In conclusion, the present data provide additional support for the existence of a domain-specific face perception mechanism in the extra-striate regions of the visual system, and show that, unlike the visual processing of other stimulus categories, the function of this mechanism is immune to strategic or attentional influences across tasks.

Acknowledgements

This study was supported by a grant from the US–Israel Science Foundation (98-00057) to Shlomo Bentin. We are grateful to Anat Prior for valuable comments on a previous version of this manuscript and to Ayelet Landau for skillful assistance.

References

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*, 416–430.

Bentin, S. (in press). The neural basis of face perception. In L. Nadel (Ed.), *Encyclopedia of cognitive science*. London: Macmillan.

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.

Bentin, S., & Deouell, L. Y. (2000). Structural encoding and identification in face processing: ERP evidence for separate mechanisms. *Cognitive Neuropsychology*, *17*, 35–54.

Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition: evidence from developmental prosopagnosia. *NeuroReport*, *10*, 823–827.

Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., & von Cramon, D. Y. (in press). Conceptual priming in visual face-processing: electrophysiological evidence. *Psychological Science*.

Brown, V., Huey, D., & Findlay, J. N. (1997). Face detection in peripheral vision: do faces pop out? *Perception*, *26*, 1555–1570.

Bushnell, I. W. R., Sai, F., & Mullin, J. T. (1989). Neonatal recognition of the mother's face. *British Journal of Developmental Psychology*, *7*, 3–15.

Caldera, R., Thut, G., Servoir, P., Michel, C., Bovet, P., & Renault, B. (2001). Face versus non-face perception and the "other-race" effect: a spatio-temporal ERP study, submitted for publication.

Campbell, R., Pascalis, O., Coleman, M., Wallace, S. B., & Benson, P. J. (1997). Are faces of different species perceived categorically by human observers? *Proceedings of the Royal Society of London, Series B – Biological Sciences*, *264*, 1429–1434.

Carey, S. (1992). Becoming a face expert. In V. Bruce, A. Cowey, A. W. Ellis & D. I. Perrett, *Processing the facial image: proceedings of a Royal Society discussion meeting* (pp. 95–103). Oxford: Calderon Press.

Cauquil, A. S., Edmonds, G. E., & Taylor, M. J. (2000). Is the face-sensitive N170 the only ERP not affected by selective attention? *NeuroReport*, *11*, 2167–2171.

de Gelder, B., & Rouw, R. (2000). Configural face processes in acquired and developmental prosopagnosia: evidence for two separate systems? *NeuroReport*, *11*, 3145–3150.

De Renzi, E. (1997). Prosopagnosia. In T. E. Feinberg & M. J. Farah, *Behavioral neurology and neuropsychology* (pp. 245–256). New York: McGraw-Hill.

de Schonen, S., & Mathivet, E. (1990). Hemispheric asymmetry in a face discrimination task in infants. *Child Development*, *61*, 1192–1205.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S., & Friston, K. J. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, *389*, 596–599.

Donchin, E. (1981). Surprise!…Surprise? *Psychophysiology*, *18*, 493–512.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357–374.

Easterbrook, M. A., Kisilevsky, B. S., Hains, S. J., & Muir, D. W. (1999). Faceness or complexity: evidence from newborn visual tracking of face-like stimuli. *Infant Behavior and Development*, *1*, 17–35.

Easterbrook, M. A., Kisilevsky, B. S., Muir, D. W., & Laplante, D. P. (1999). Newborns discriminate schematic faces from scrambled faces. *Canadian Journal of Experimental Psychology*, *53*, 231–241.

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.

Elman, J. L., Bates, A. E., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: a connectionist perspective on development*, Cambridge, MA: MIT Press.

Farah, M. J., Levinson, K. L., & Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia*, *33*, 661–674.

Farah, M. J., Rabinowitz, C., Quinn, G. E., & Liu, G. T. (2000). Early commitment of neural substrates for face recognition. *Cognitive Neuropsychology*, *17*, 117–123.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychological Review*, *105*, 482–498.

Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition be dissociated from object recognition? *Journal of Cognitive Neuroscience*, *11*, 349–370.

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, *11*, 219–224.

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 mechanisms for face recognition. *Vision Research*, *37*, 1673–1682.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform “face area” increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training ‘greeble’ experts: framework for studying expert object recognition processes. *Vision Research*, *38*, 2401–2428.

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.

Goren, C., Sarty, M., & Wu, P. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*, 544–549.

Haxby, J. V., Horwitz, B., Ungerleider, L. G., Maisog, J. M., Pietrini, P., & Grady, C. L. (1994). The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. *Journal of Neuroscience*, *14*, 6336–6353.

Ishai, A., Ungerleider, L. G., Martin, A., & Haxby, J. V. (1999). Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Science USA*, *96*, 9379–9384.

Jeffreys, D. A., & Tukmachi, E. S. A. (1992). The vertex-positive scalp potential evoked by faces and by objects. *Experimental Brain Research*, *91*, 340–350.

Johnson, M. H. (1997). *Developmental cognitive neuroscience*, Cambridge, MA: Blackwell Press.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborn preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*, 1–19.

Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development. The case of face recognition*, Cambridge, MA: Blackwell Press.

Johnson Jr., R., Schmidt, J. M., Barnhardt, J., Singh, C., & Zhu, J. (1999). *ERP investigation into the neural basis of category-specific knowledge*. Paper presented at the 39th meeting of the Society of Psychophysiological Research, Granada, Spain.

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.

Karmiloff-Smith, A. (1995). Annotation: the extraordinary cognitive journey from foetus through infancy. *Journal of Child Psychology and Child Psychiatry*, *36*, 1293–1313.

Karmiloff-Smith, A. (1997). Crucial differences between developmental cognitive neuroscience and adult neuropsychology. *Developmental Neuropsychology*, *13*, 513–524.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Science*, *2*, 389–399.

Kohn, B., & Dennis, M. (1974). Patterns of hemispheric specialization after hemidecortication for infantile hemiplegia. In M. Kinsborne & W. L. Smith, *Hemispheric disconnection and cerebral function* (pp. 34–47). Charles C. Thomas: Springfield, IL.

Luck, S. J. (1998). Neurophysiology of selective attention. In H. Pashler, *Attention* (pp. 257–295). East Sussex: Psychology Press.

Mangun, G. R., Hillyard, S. A., & Luck, S. J. (1992). Electrocortical substrates of visual selective attention. In D. E. Meyer & S. Kornblum, *Attention and performance* (pp. 219–243). , XIV. Cambridge, MA: MIT Press.

Maurer, D., & Barrera, M. (1981). Infants’ perception of natural and distorted arrangements of a schematic face. *Child Development*, *52*, 196–202.

Maurer, D., & Young, R. (1983). Newborns’ following of natural and distorted arrangements of facial features. *Infant Behavior and Development*, *6*, 127–131.

McCarthy, G. (1999). Physiological studies of face perception in humans. In M. S. Gazzaniga, *The new cognitive neuroscience* (pp. 393–409). Cambridge, MA: MIT Press.

McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception II. Response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral Cortex*, *9*, 431–444.

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

McNeil, J. E., & Warrington, E. K. (1993). Prosopognosia: a face-specific disorder. *Quarterly Journal of Experimental Psychology*, *46A*, 1–10.

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, R., & Kleiner-Gatehercoal, C. A. (1999). Face perception during early infancy. *Psychological Science*, *10*, 419–423.

Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, *9*, 555–604.

Naatanen, R., & Picton, T. W. (1987). The N1 wave of the human electric and magnetic response to sound: a review and analysis of the component structure. *Psychophysiology*, *24*, 375–425.

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, *10*, 3–18.

Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: correlates of "popout" under anesthesia. *Visual Neuroscience*, *16*, 15–34.

Pascalis, O., Demont, E., de Haan, M., & Campbell, R. (2001). Recognition of faces of different species: a developmental study between 5 and 8 years of age. *Infant and Child Development*, *10*, 39–45.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., & Fabre-Grenet, M. (1995). Mother's face recognition by neonates: a replication and an extension. *Infant Behavior and Development*, *18*, 79–85.

Puce, A., Allison, T., & McCarthy, G. (1999). Electrophysiological studies of human face perception III. Effects of top-down processing on face-specific potentials. *Cerebral Cortex*, *9*, 445–458.

Rhodes, G. (1988). Looking at faces: first-order and second-order features as determinants of facial appearance. *Perception*, *17*, 43–63.

Rolls, E. T., & Baylis, G. C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research*, *65*, 38–48.

Rossion, B., Delvenne, J. -F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., & Guerit, J. -M. (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., & Crommelinck, M. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: an electrophysiological account of face-specific processes in the human brain. *NeuroReport*, *11*, 69–74.

Sagiv, N., & Bentin, S. (in press). Structural encoding of human and schematic faces. *Journal of Cognitive Neuroscience*.

Sergent, J. (1988). Face perception and the right hemisphere. In L. Weiskrantz, *Thought without language* (pp. 108–113). New York: Oxford University Press.

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

Simion, F., Cassia, V. M., Turati, C., & Valenza, E. (2001). The origins of face perception: specific versus non-specific mechanisms. *Infant and Child Development*, *10*, 59–65.

Simion, F., Valenza, E., Umilta, C., & Dalla Barba, B. (1998). Preferential orienting to faces in newborns: a temporal-nasal asymmetry. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 1399–1405.

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

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology*, *46A*, 225–245.

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*, 583–592.

Tovee, M. J. (1998). Is face processing special? *Neuron*, *21*, 1239–1242.

Tovee, M. J. (1998). Face processing: getting by with a little help from its friends. *Current Biology*, *8*, 317–320.

Valenza, E., Simion, F., Cassia, V. M., & Umilta, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 892–903.

Vogel, E. K., & Luck, S. J. (2000). The visual N1 component as an index of a discrimination process. *Psychophysiology*, *37*, 190–203.

Walton, G. E., Bower, N. J. A., & Bower, T. G. R. (1992). Recognition of familiar faces by newborns. *Infant Behavior and Development*, *15*, 265–269.