



BIROn - Birkbeck Institutional Research Online

Gliga, Teodora and Dehaene-Lambertz, G. (2006) Structural encoding of body and face in human infants and adults. *Journal of Cognitive Neuroscience* 17 (8), pp. 1328-1340. ISSN 0898-929X.

Downloaded from: <http://eprints.bbk.ac.uk/31605/>

Usage Guidelines:

Please refer to usage guidelines at <http://eprints.bbk.ac.uk/policies.html> or alternatively contact lib-eprints@bbk.ac.uk.

Structural Encoding of Body and Face in Human Infants and Adults

Teodora Gliga¹ and Ghislaine Dehaene-Lambertz^{1,2}

Abstract

■ Most studies on visual perception of human beings have focused on perception of faces. However, bodies are another important visual element, which help us to identify a member of our species in the visual scene. In order to study whether similar configural information processing is used in body and face perception, we recorded high-density event-related potentials (ERPs) to normal and distorted faces and bodies in adults and 3-month-old infants. In adults, the N1 responses evoked by bodies and faces were similar in amplitude but differed slightly in latency. The voltage topography of N1 also differed in concordance with fMRI data showing that two distinct areas are involved in face and body perception.

Distortion affected ERPs to faces and bodies similarly from N1 on, although the effect was significant earlier for bodies than for faces. These results suggest that fast processing of configural information is not specific to faces but it also occurs for bodies. In 3-month-old infants, distortion decreased the amplitude of P400 around 450 msec, showing no interaction with image category. This result demonstrates that infants are not only able to recognize the normal configuration of faces, but also that of bodies. This could either be related to an innate knowledge of this particular type of biological object, or to fast learning through intense exposure during the first months of life. ■

INTRODUCTION

Faces are highly salient and biologically important visual stimuli. Nevertheless, at a distance at which face elements are difficult to detect, or with some particular body orientations where the face is not visible, the body scheme (number of body parts, their relative position, and bipedalism) becomes a much more informative cue of the presence of a human being. Bodies also convey emotional as well as identity information (Hadjikhani & deGelder, 2003; Cutting & Kozlowski, 1977). Despite these similar roles of faces and bodies, the study of the processes involved in the perception of human beings has focused more on face perception than on body perception, and we are now in possession of a rich knowledge base concerning face-processing mechanisms. Most of these data suggest that faces are analyzed differently from other visual objects. It has been shown that face processing relies not only on the identification of the component elements (termed local or featural information), but also on the analysis of the relative positioning of internal elements (also referred to as global or configural information). Configural information is less important for nonface object analysis whereas for faces, extracting configural information is so impor-

tant that it can even interfere with local analysis (Farah, 1996). Configural information can be described at two levels: A first level refers to the relative position of internal elements (a face is a face because it has two eyes above the nose, which is above a mouth) and is also designated as first-order information, whereas second-order information deals with the exact distances between these elements and is thought to be important for individual recognition (Maurer, 2002).

When first-order information is disrupted (such as when subjects are asked to recognize inverted faces), recognition of people or expressions is hindered (Valentine, 1988). These observations led Bruce and Young (1986) to propose a model of face processing which contains a compulsory first step of “structural encoding.” The success of this encoding is a prerequisite for subsequent processing, such as individual recognition and emotional expression recognition.

Of course, first-order information must be encoded for all types of objects in order to determine if an object is right side up or not. Nevertheless, nonface object processing suffers less from 180° rotations (Eimer, 2000a) or from the displacement of components. Farah (1996) shows that breaking the normal configuration of a face facilitates the recognition of one component part, whereas no difference is observed for houses. An exception has been observed for those objects for which an expertise has been developed. Car, dog, or bird experts can process these objects in a manner that is

¹CNRS, Unité INSERM 562, Service Hospitalier Frédéric Joliot, CEA/DRM/DSV, Orsay, France, ²Hôpital de Bicêtre, AP-HP, France

similar to face processing (Gauthier & Tarr, 2002; Diarmod & Carey, 1986). On the basis of these results, expertise was proposed as an explanation for face special perception as well (Gauthier, Curran, Curby, & Collins, 2003).

Human body is not a less frequent object in the environment. Despite this, few studies have focused on characterizing the structural processing of bodies. Bodies have been frequently studied in relation to motion (Astafiev, 2004). Body movement, for example, conveys detailed information on the emotion, age, or sex of a certain person (Vaina, Solomon, Chowdhury, Sinha, & Delliveau, 2001; Cutting & Kozlowski, 1977). Nevertheless, inversion impairs recognition of bodies (Reed, Stone, Bozova, & Tanaka, 2003) and delays body-evoked responses (Stekelenburg & deGelder, 2004). If one were to take these results and follow the same line of reasoning used for faces, the “structural encoding stage” should be one of the first steps in body processing as well. This stage would be concerned with the detection of a normal body scheme (a trunk with two arms attached to the superior extremity and two legs attached to the lower extremity and a head above) and should be affected by the distortion of this body scheme.

Human body is as common as human face in visual environment from very early on. Therefore, we might expect similar processing for face and body not only in adulthood, but also along development. On the contrary, because face perception is so important for social communication, to follow gaze direction, to understand speech, or to decipher emotion, it is possible that, early in life, face processing is facilitated by special mechanisms that were selected during the evolution of our species. Parallel studies on face and body perception in infancy are the best way to clarify this issue. To our knowledge, infant studies on body perception are far less numerous than studies on face perception. A rich literature shows that 3-month-olds are already familiar with the structural aspects of faces. For instance, at this age, infants prefer typical configured face stimuli to distorted facial stimuli (for whom the internal elements have been displaced), showing sensitivity to first-order structural information (Maurer & Barrera, 1981). Finer-grained structural information appears to be extracted by infants because they are able to recognize familiar faces (de Haan, Johnson, Maurer, & Perett, 2001). At 4 months, face recognition is based not only on the individual internal elements, but also on their particular association within a face. For example, infants treat a hybrid face that has been created from elements of two familiar faces as unfamiliar (Cashon & Cohen, 2003).

At the same age, infants discriminate point-light displays of human movement from random patterns (Bertenthal, Proffitt, Kramer, & Spenter, 1987). It has been proposed that multiple processing constraints, including stored knowledge of human movement and

body form, contribute to the interpretation of these point-light displays, especially because infants do not discriminate point-light displays depicting unperturbed unfamiliar objects (e.g., a spider) from perturbed versions of the same objects (Bertenthal, 1996). Data on face and body movement imitation (e.g., mouth opening, tongue protrusion, finger movement) are seen as additional proof that infants possess an internal frame of their own face and body (Meltzoff & Moore, 1977). Nevertheless, when using static displays, 12-month-old children seem unable to choose a normal body configuration among distorted configurations where the limbs were not correctly placed on the body. It is not until 18 months that children start showing a preference for the typical configuration of bodies (Slaughter, Heron, & Sim, 2002). Are only the more salient aspect of bodies (such as their movement) processed in the first months of life? An alternative interpretation of infants' behavior in the aforementioned study by Slaughter et al. is possible. We know that faces, especially with eyes open, are extremely effective in capturing infants' attention (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000). It is thus possible that, in these experiments, babies were not able to disengage from looking at the faces and therefore ignored the bodies.

In order to study the difference between the encoding of faces and bodies at a mature stage and at an immature stage, we presented adults (Experiment 1) and infants (Experiment 2) with intact and distorted face and body images. Intact faces were front views of faces with directed gaze. Because we were interested in characterizing not only adult but also infant face perception, the face distortion was created by destroying the T-shaped configuration of the facial elements, a key element in face detection from the very first months of life (Johnson, Dziurawiec, Ellis, & Morton, 1991).

For the intact body condition, we used stimuli that depicted bodies without heads, with arms and legs fully visible. Faces were not presented with bodies in order to avoid cross-category interference. Face cues, even if very schematic (an opaque circle above the neck), have been shown to elicit face-specific brain activity (Cox, Meyers, & Sinha, 2004). Another reason for eliminating the faces was infants' failure to detect important limb displacements when the face was presented at its normal position on the body (Slaughter et al., 2002). Because we do not know what the main cues used by infants or adults for body detection are, both limb position and symmetry were altered for the distorted body condition. This was achieved by displacing an arm or a leg onto the neck. For both faces and bodies, the distortion procedure left the individual parts intact (the local information) while affecting their relative position (the configural information).

The neural bases of face and body structural information processing were characterized by comparing the evoked event-related potentials (ERPs) for these image

categories. Using similar recording systems and the same paradigm for both ages, adults (Experiment 1) and 3-month-old infants (Experiment 2) offer the opportunity to compare the time-course of the infant electrophysiological responses to those of adults, and to determine whether the mechanisms described for face perception can be extended to other human visual aspects.

EXPERIMENT 1: ADULTS

Several different components of ERPs have been studied with respect to object perception (see Figure 1). The P1 component, having a latency of 100–150 msec, is considered to be of occipital origin (Arroyo, 1997) and to reflect the encoding of low-level visual properties (Allison, 1999). Because our interest lies in object-specific knowledge, we focus on subsequent components, in particular, N1 and P2. The N1 component is a negative deflection observed around 170 msec that is more prominent over the occipito-temporal scalp. All objects evoke an N1 response (Itier & Taylor, 2004a), whose amplitude can be increased by attention (Carmel & Bentin, 2002), or by the degree of experience with that class of objects (Rossion, Gauthier, Goffaux, Tarr, & Crommelnick, 2002). Nevertheless, the face N1 is consistently larger in amplitude, shorter in duration, and more right lateralized than that evoked by other complex objects (Itier & Taylor, 2004b; Bentin, Allison, Puce, Perez, & McCarthy, 1996). In the face literature, this component has been termed “N170” and is considered a signature of individual face processing (Carmel & Bentin, 2002). N170 is followed by a parietal–occipital positivity that develops in the 300–400 msec interval, termed P2. This component is considered to reflect memory encoding and recognition. At this latency, ERPs are modulated by the repetition of identical faces (Itier & Taylor, 2004a; Halit, deHaan, & Johnson, 2000; Paller, Gonsalves, Grabowecky, Bozic, & Yamada, 2000), thus reflecting familiarity or recognition mechanisms. Halit et al. (2000) propose that P2 reflects processing related to encoding facial identity and distance from an individual prototypical face. Recordings in the temporal lobe show that intracranial ERPs at the same latency are modulated by learning of face–name association (Puce, Allison, & McCarthy, 1999).

Most studies investigating the structural analysis of faces have focused on the N170 component. The data from these studies suggest that N170 reflects the “structural encoding stage” proposed by the Bruce and Young model. For instance, N170 is delayed by the inversion of faces (Rousselet, Mace, & Fabre-Thorpe, 2004; Bentin et al., 1996), a transformation considered to interfere with configural processing. Its magnetic counterpart (M170) has been demonstrated to be sensitive to the relative position of the internal face elements, showing a decrease in amplitude for distorted faces (Liu, Harris, &

Kanwisher, 2002). Electrophysiological recordings in the monkey temporal lobe led to the discovery of face-selective neurons, which respond only when the internal elements of the face respected a normal configuration (Rolls, 2000), indicating a possible neural mechanism for the encoding of the facial configural information at the level of the N170.

Fewer studies have focused on characterizing the neuronal encoding of body configural information. Stekelenburg and deGelder (2004) showed that, in humans, inversion modulates the body-evoked N1 in the same way as it modulates the face-evoked N1, whereas the inversion of another object (shoes) has no effect on the respective N1. This result shows that although both types of objects, bodies and shoes, are experienced more often in an upright position, only body processing has been affected, which may be due to the supplementary biological importance of the upright position for this class of objects.

The goal of the current experiment was to study body first-order information encoding, and to measure how fast the information on the relative positioning of component elements is extracted, without manipulating their orientation. We hypothesized that N1 might reflect encoding of body first-order information, as it reflects face first-order information encoding. We thus expect that the N1 response should be affected by the distortion of the body scheme.

ERPs from 10 adults were recorded with a high-density array (129 electrodes) while subjects passively viewed images belonging to six categories: intact faces, intact bodies, distorted faces, distorted bodies, scrambled faces, and scrambled bodies. The scrambled versions of face and body images were used in order to control for low-level visual differences between the face and body images (contrast and luminosity). We analyzed how category (faces, bodies) and condition (intact, distorted and scrambled) modulated the latency, amplitude, and topography of the two components of visually evoked potentials that are linked to object-specific processing: N1 and P2. For each component, we performed hierarchical comparisons: We first compared the scrambled and intact image conditions to investigate whether the component showed a sensitivity to object perception. Next we examined the differences between ERPs to faces and bodies in the intact condition to see how specific the component was. Finally, we looked at how distortion affected the component by comparing the intact and distorted conditions.

Results

N1

N1 analyses were done over two symmetrical packs of 10 electrodes corresponding to the T5/T6 position. Peak latency and amplitude, as well as mean amplitude over

the time window of interest, were used as dependent measures. Because peak amplitude and mean amplitude results were similar, only mean amplitude analyses are presented in the text; however, peak amplitude analysis can be found in notes.

Effect of scrambling. The N1 was larger for intact faces and bodies compared to the scrambled images (Figure 1). A main effect of condition (Intact vs. Scrambled) was found in mean amplitude over the N1 interval (180–260 msec) [$F(1,9) = 26.71, p < .001$].¹ Post hoc analyses showed that this effect was present in each category [Intact vs. Scrambled faces: $F(1,9) = 36.73, p < .001$; Intact vs. Scrambled bodies: $F(1,9) = 12.44, p = .006$]. This effect did not interact with hemisphere.

Effect of category in intact images. The N1 response developed significantly earlier for faces than for bodies [$F(1,9) = 36.00, p < .001$]. The grand average peaked at 204 msec for faces versus 228 msec for bodies. This latency difference was observed for both hemispheres [right hemisphere: $F(1,9) = 17.66, p = .02$; left hemisphere: $F(1,9) = 52.32, p < .001$] with no Category \times Hemisphere interaction. These differences in latency were not due to low-level properties of the image categories, because no difference was observed between the latency of ERPs to scrambled faces and bodies ($p = .3$), which had the same luminance and contrast differences as the intact faces and bodies. Mean amplitude analysis revealed no significant differences between the two categories [$F(1,9) < 1$]. Both categories induced stronger responses over the right hemisphere [effect of hemisphere restricted to bodies: $F(1,9) = 12.86, p = .006$; effect of hemisphere restricted to faces: $F(1,9) = 20.05, p = .001$] with no significant Category \times Hemisphere interaction [$F(1,9) < 1$].²

Our dense coverage of the scalp allows recording a refined image of the scalp voltage for each category. As seen in Figure 1, the scalp distributions of the N1 responses to faces and bodies were different. Bodies evoked an occipital–temporal negativity that reversed over the frontal electrodes while the two poles of the N1 component evoked by faces were more outspread, with the negative pole being stronger on the temporal electrodes and the positive pole covering the anterior frontal and frontal electrodes. To better analyze these topographical differences and cover both poles of the N1 responses, four groups of electrodes were chosen: two covering the positive maxima of both categories (frontal anterior and frontal), and two others covering the negative maxima (occipital and temporal). An ANOVA was computed with Electrodes (4 groups of 8 electrodes), Hemisphere, and Category (faces, bodies) as within-subject factors. Over the N1 interval, this showed a significant Electrodes \times Category interaction [$F(3,27) = 4.36, p = .01$].

In order to eliminate the possibility that the observed differences were due to amplitude differences at the level of a common underlying neuronal generator rather than to a difference in the generators (McCarthy & Wood, 1985), the same ANOVA was performed after a vectorial scaling of the two conditions. This alternative hypothesis should be accepted if after scaling, the Electrodes \times Category interaction is no longer significant. However, despite data scaling, this interaction remained significant in our data [$F(3,27) = 6.55, p = .001$].

Effect of distortion. As seen in Figure 2A, distortion of faces and bodies did not affect N1 latency but did affect N1 amplitude, mainly during the first half of the N1 interval. Body distortion was more effective in diminishing N1 amplitude. ANOVAs computed for intact versus distorted images confirmed the absence of a latency effect in each category [intact vs. distorted faces: $F(1,9) = 1.81, p = .2$; intact vs. distorted bodies: $F(1,9) < 1$]. To analyze a main effect of distortion on N1 amplitude, across both categories, despite the difference in N1 latency, we calculated the mean voltage over the 160–200 msec time interval for faces and the 200–240 msec interval for bodies. Normal images evoked larger N1 than distorted ones as revealed by a marginal effect of condition [$F(1,9) = 4.78, p = .05$]. No main interaction with hemisphere was found [$F(1,9) = 2.90, p = .1$]. The Condition \times Category interaction was not significant [$F(1,9) < 1$]. However, in post hoc analyses, a distortion effect was significant only for body category [bodies: $F(1,9) = 6.11, p = .03$; faces: $F(1,9) = 2.75, p = .1$]. The body distortion effect was asymmetric, being more pronounced over the right hemisphere [Condition \times Hemisphere: $F(1,9) = 14.79, p = .004$].

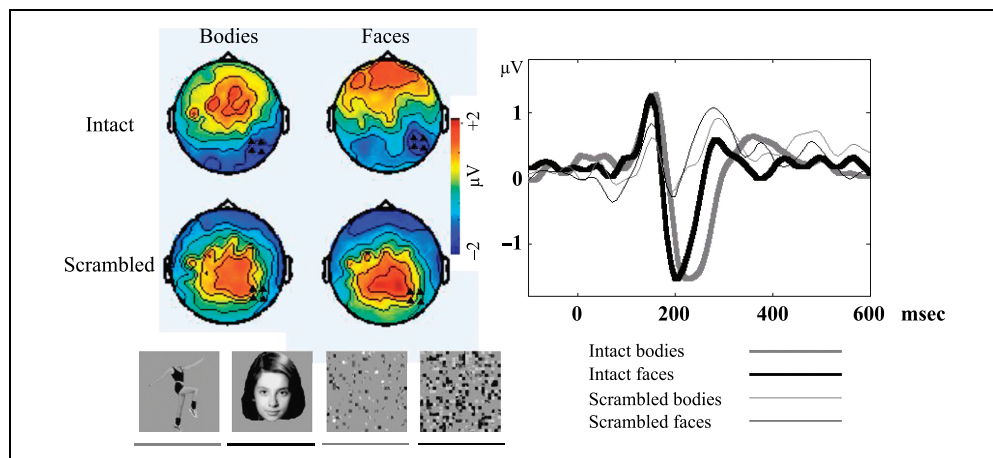
To summarize, body-evoked N1 response was as strong as face-evoked N1 response, but slightly delayed. Its amplitude was reduced by distortion.

P2

The second component of interest developed over the scalp in the 300–400 msec interval, as a parieto-occipital positivity, reversing over the antero-frontal electrodes (Figure 2). Being a slow component, it was less easy to determine the latency of the P2 peak, therefore only mean amplitude values are reported. Two groups of 10 symmetrical electrodes were chosen, one group at the maximum of the positivity and the second at the maximum of the negativity of the P2 response.

No effect of category was present at this latency (scrambled vs. intact: $p = .08$; faces vs. bodies: $p = .3$). There was a main effect of condition: distortion of images reduced the mean amplitude of P2 [Condition \times Electrodes: $F(3,27) = 6.02, p = .003$]. The triple interaction Condition \times Electrodes \times Category was

Figure 1. Voltage maps of adults' ERPs to bodies and faces (first row) and scrambled bodies and faces (second row) at the maximum of N1 (228 msec for bodies, 204 msec for faces). The last column shows the grand-averaged waveform summed over a group of temporo-occipital electrodes (marked by triangles on the voltage maps). An example from each category of stimuli is given on the bottom row. The color and thickness of the lines above the pictures indicate the condition of the waveforms.



not significant [$F(3,27) = 1.47, p = .24$]. Nevertheless, although the effect was strongly significant for faces [Condition \times Electrodes: $F(3,27) = 5.26, p = .005$], the body distortion effect was not significant at this latency [Condition \times Electrodes: $F(3,27) = 2.08, p = .12$]. Thus, distortion reduced the amplitude of P2, especially in the case of face distortion.

Discussion

We investigated the neural correlates of human body and face perception, focusing on the encoding of first-order structural information. It should be noted that the N1 response to faces was slow in our study relative to the latencies reported in other studies (204 msec for faces in our study, whereas latencies of 160–180 msec are generally found). An important difference between

our experiment and others should be underscored: here we used a continuous presentation where in most, if not all, studies a blank screen is inserted between two successive images. Unpublished results from our laboratory confirm that when similar stimuli and experimental conditions are used, but a fixation cross precedes each face image, N1 is recorded with a latency of 170 msec, whereas N1 is delayed as here in response to a second stimulus presented in succession with the first (Gliga, unpublished results).

We observed in Experiment 1 that human bodies and front-view faces evoked equally strong occipito-temporal right-lateralized negativities. Few studies have found categories of stimuli that evoke N1 responses that are comparable in amplitude and laterality to N1 evoked by faces. Therefore, it has been proposed that face processing depends on specific brain mechanisms that generate

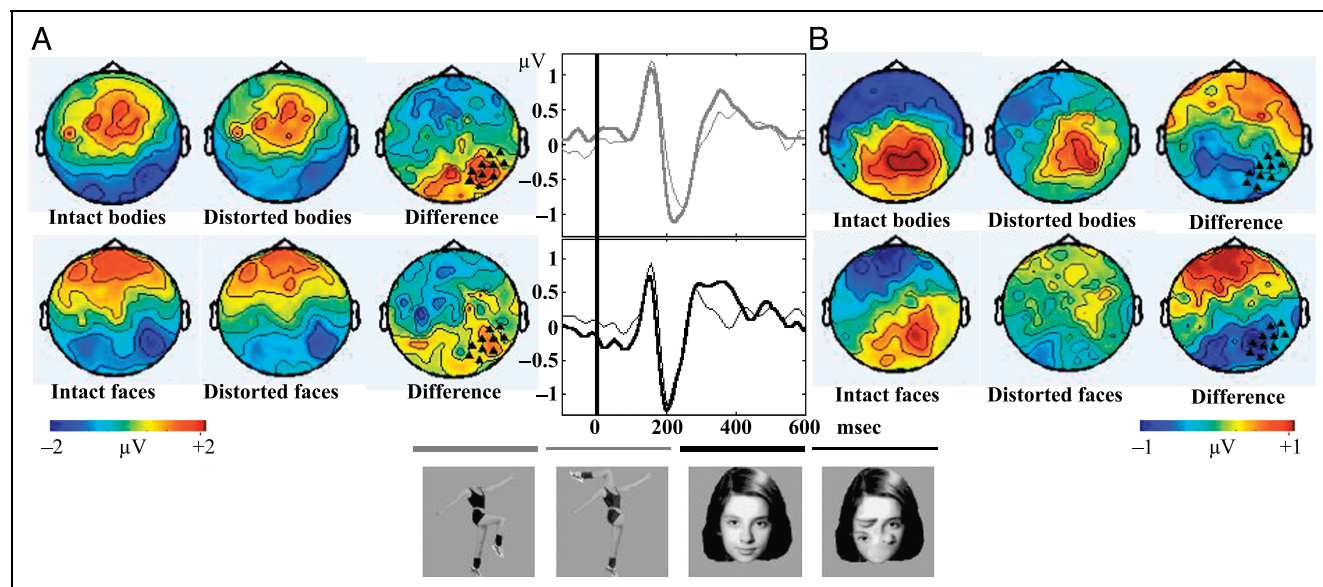


Figure 2. Image distortion effects observed in adults at the level of N1 (A) and P2 (B). For both components, voltage maps of ERPs evoked by intact and distorted images and their difference are shown (first row bodies and second row faces). The central column displays grand-averaged waveforms summed over the group of electrodes that is marked by triangles on the difference maps. An image sample from each of the categories is shown on the bottom row. The color and thickness of the lines above the pictures indicate the condition of the waveforms.

N170 (Bentin et al., 1996). In our experiment, as can be seen in Figure 1, the differences in topography and latency between N1 evoked by the faces and the bodies suggest that they originate in distinct brain areas, with different temporal dynamics. The more overspread N1 topography for faces is compatible with a more distant neuronal source from the surface. Indeed, face perception is associated with ventral temporal activity (fusiform gyrus), in what has been called the “fusiform face area” (FFA) (Kanwisher, McDermott, & Chun, 1997) and also with right superior temporal sulcus activity (Henson et al., 2003; Itier & Taylor, 2004c). Bodies and body parts activate predominantly the right occipito-temporal cortex, in a region closer to the human MT, designated as the “extrastriate body area” (EBA) (Downing, 2001). However, Stekelenburg and deGelder (2004), comparing responses to faces and bodies, recorded identical topographies for both categories and concluded that N170 has a broader functional significance, which could include holistic perception of both faces and bodies. The discrepancy between these two studies could be explained by a sparser electrode coverage of the head in the Stekelenburg and deGelder’s study (49 vs. 129 here) and the fact that a spot replaced the head in their body pictures. It has been shown that this type of stimuli contains enough information about the head to activate the FFA (Cox et al., 2004). It is thus possible that the N1 response to body has contaminated an additional N1 face response triggered by the schematic “head” in this study.

Although the underlying sources for the face and body N1 response were different, suggesting different networks underlying face and body processing, in the present experiment we cannot disentangle a specific body response from a response generated by more general object-processing mechanisms. Nevertheless, we can underline differences between the body N1 observed here and the object N1 described in the literature: the body N1 response was right-lateralized and as large as face N1 response, whereas other objects evoke bilateral and smaller N1 (Itier & Taylor, 2004b). Moreover, although it has been shown that the experimental status of a stimulus category, for example, target versus nontarget, can amplify the N1 response to non-face objects (e.g., cars in Carmel & Bentin, 2002), this explanation cannot account for the strong N1 response to body, because no active task was performed for the body category. Expertise also induces strong N1 responses to the object of expertise (Rossion, Gauthier, et al., 2002). This observation has generated a fierce debate between supporters of face-specific mechanisms and supporters of face processing as only one example of expert processing of objects. The same debate is relevant here. On one hand, bodies are an important biological object, for which specific brain mechanisms might have evolved. On the other hand, bodies are very common objects in our environment for which we need

to develop expertise. Without providing proof for one or the other hypothesis, we show that this debate should not be limited to faces or to a single brain area (i.e. the FFA), but should be enlarged to include other biological objects.

To further compare body and face processing, we were interested in studying how first-order structural information is extracted in the first steps of the perceptual analysis in both cases. More precisely, we asked whether the N1 response, considered to reflect structural processing, is modulated by a distortion of the body scheme as has been observed for face internal configuration distortion. We show here that distorting the body scheme significantly reduces the N1 amplitude. The right hemisphere was more affected by the distortion, allowing us to extend the hypothesis of right-lateralized networks for configural processing (Rossion, Dricot, et al., 2000) to the case of body configural processing.

From this experiment, we cannot determine which aspect of body distortion was critical in producing this effect. Was it the lack of one of the limbs in its normal position or was it the placement of this limb on the neck in the normal position of the head? In other words, was the effect due to the modification of the body scheme or instead to the conflicting information coming from the head region? There is evidence that, even if analyzed in distinct brain regions, information concerning the body and the head is integrated. When viewing a body, expectations about the location of the head relative to the body are formed (Cox et al., 2004; Gliga & Dehaene-Lambertz, submitted). Lewis and Edmonds (2003) observed that face detection is faster when the face is placed at the correct location on the body. It is thus possible that our distorted stimuli generate a conflict between expectation of a head and the presence of an impossible object at this location. More experiments are necessary to determine whether body structural information is encoded as a preattentive template, in order to subserve a fast detection of the face, as suggested by Cox et al. (2004).

The N1 response to faces was less affected by their distortion. In our experiment, faces were randomly presented among other categories of stimuli (bodies and scrambled images). This could have biased subjects’ strategies towards the identification of the image category (“faceness” vs. “bodyness”) to the detriment of other types of analysis (distortion detection, image recognition). Halit et al. (2000) show such effects of stimulus neighborhood on the perceptual strategy, in that case, recognition versus aesthetic judgment. Because face detection relies on multiple structural cues: configuration, external contour, internal elements, the correct external contour of our distorted faces pictures could have been used for a first fast categorization. Indeed, it has been shown that the external facial contour or very schematic faces can generate strong N1 responses (Sagiv & Bentin, 2001; Eimer, 2000b).

Bentin et al. (1996) and George et al. (1996) reported that distorted faces evoke N1 responses with higher amplitude and delayed latency than normal faces (Bentin et al., 1996; George, Evans, Fiori, Davidoff, & Renault, 1996). Isolated eyes are also known to evoke N1 responses with higher amplitude and delayed latency (Bentin et al., 1996). Consequently, it has been proposed that in intact faces, the normal configuration of face elements triggers a global analysis and inhibits part (eye) analysis. The relative weight of these processes changes for the distorted faces where the recorded response would be due to parts (principally eye) analysis. Contrary to the face stimuli in the above studies, in our face stimuli the two eyes were displaced independently (they were never aligned). Very little difference between the intact and distorted faces was noticed in our experiment, suggesting that eye alignment is necessary for triggering the higher amplitude responses in the George et al. study.

The effects of face and body distortion were not restricted to N1. The occipital–parietal positivity that developed over the scalp in the 300–400 msec interval for intact faces and bodies was reduced in amplitude for the distorted images. The strong effect obtained for faces shows that first-order structural information, even if not crucial for face detection, is needed for further processing. Our data do not allow us to point out precisely the nature of this further processing, but the literature offers some hypotheses. At this latency, ERPs are modulated by body and face emotional expressions (Stekelenburg & deGelder, 2004), by comparisons and by encoding of faces with respect to stored individual prototypes (Halit et al., 2000), or by learning semantic associations (Puce et al., 1999). No active learning of faces was required in our experiment, but it is possible that automatic memorization was triggered for intact faces and failed for distorted ones. Evidence for automatic encoding exists; an increased parietal positivity is seen for previously viewed faces (Paller et al., 2000) under explicit, but also under implicit, conditions (Müntz et al., 1998).

EXPERIMENT 2: INFANTS

Having found evidence to suggest that not only face, but also body, structural information is encoded at a perceptual level in the adult brain, and knowing that infants show knowledge of face structure early in life, we wanted to investigate whether body processing capacities develop in parallel.

In order to assess early knowledge of human face and body appearance, the same six categories of images (intact faces and bodies; distorted faces and bodies; and scrambled faces and bodies) were presented to 3-month-old infants while recording ERPs. We expected that, because of the absence of a face in body images, infants would be able to focus their attention on the

structural aspects of bodies, and that the aspects of body scheme modified in our pictures are as important in infants as they were in adults, as in the Experiment 1.

Two ERP components have been linked to object perception in infants: a negative-going deflection over midline and lateral, posterior electrodes whose peak latency decreases from approximately 350 to 290 msec between 3 and 12 months of age, termed N290 in the infant literature, and a positive component most prominent over posterior, lateral electrodes whose peak latency decreases from approximately 450 to 390 msec between 3 and 12 months of age, termed P400 (Halit et al., 2003).

N290/P400 modulation by different experimental factors changes in the course of the first year of life, suggesting a fast visual development during this period (reviewed by de Haan, 2003). In the first months of life, human and monkey faces (de Haan, Pascalis, & Johnson, 2002) or direct versus averted gaze (Farroni, Csibra, Simion, & Johnson, 2002) are discriminated at the level of the infant N290. Whereas face orientation affects the adult N1, it has no effect on N290 until 12 months of age (de Haan et al., 2002). Both these results, and the fact that in the course of development eyes evoke much stronger and faster N1 responses than faces (Gliga & Dehaene-Lambertz, submitted; Taylor & Baldeweg, 2002), suggest that the infant component is not equivalent to the adult N1, and may be initially linked to eye and gaze analysis. The existence of a fast “eye detector” during the first months of life is supported by behavioral studies, which show that the eyes are the most salient feature of the face (Maurer, 1985).

The second component (P400) is faster for faces than for other familiar objects (toys) (de Haan & Nelson, 1999) and sensitive to face orientation. Around 3 months of age, human and monkey faces evoke stronger P400 when viewed right side up versus upside down (Halit, deHaan, & Johnson, 2003; de Haan et al., 2002), suggesting that knowledge of different visual angles of the head is encoded at this latency.

In accordance with these results, we expected that the N1 component would be sensitive to the image category, discriminating especially between presence and absence of eyes in the images, whereas P400 would be sensitive to the distortion of normal structural information. As in the previous experiment, we used scrambled images to help us control for low-level visual differences.

Results

Sixteen 3-month-old infants were included for the final analyses. As has been described previously (de Haan et al., 2002), presentation of the images evoked a succession of three components over the posterior part of the scalp: a medial occipital positivity, peaking around 180 msec (P1), followed by a negativity similar in topog-

raphy that peaked at 300 msec (the infant N1 or N290), and then by a posterior positivity extending over the medial and lateral electrodes from 400 to 600 msec (the infant P2 or P400) (see Figure 3). A first comparison between the responses evoked by the two object categories (faces and bodies) and their scrambled counterparts showed a rather different pattern than that observed in adults. The amplitude of P1 and N290 responses was significantly larger for the scrambled stimuli. At this latency, no difference was observed between the face and body ERPs. This unexpected result led us to conduct a more detailed analysis encompassing the three components: P1 (120–220 msec), N290 (240–340 msec), and P400 (440–530 msec). Three ANOVAs were thus performed over two symmetrical occipital–temporal groups of electrodes, with Hemisphere, Condition (Intact, Distorted and Scrambled), and Category (Faces, Bodies) as factors. For infants, we used the same hierarchical approach of comparisons between conditions as for adults.

P1/N290

The P1 and N290 components showed similar patterns and are thus presented together. The same variables as in adults, latency and mean amplitude, are presented (peak amplitudes in notes).

Effect of scrambling. P1 and N290 mean amplitudes were significantly larger for scrambled images than for intact images³ [Scrambled vs. Intact P1: $F(1,15) = 14.12$, $p = .002$ and N290: $F(1,15) = 5.25$, $p = .03$], whereas the latency of these components was unaffected ($ps > 0.2$). There was no difference between the two scrambled categories [Scrambled faces vs. Scrambled bodies P1: $F(1,15) < 1$ and N290: $F(1,15) = 1.71$, $p = .2$].

Effect of category in intact images. The only significant difference between the two categories was a larger N290 amplitude for faces than for bodies [$F(1,15) = 3.58$, $p = .07$ on the N290 mean amplitude; $F(1,15) = 7.96$, $p = .013$ on the N290 peak amplitude]. This difference did not interact with the factor Hemisphere.

Effect of distortion. No significant effect of distortion on P1, N290 latencies and amplitude was observed.

To conclude, the early components of infant visual ERPs (P1 and N290), contrary to what we observed in adults, are sensitive to low-level visual properties.

P400

No significant difference between scrambled and intact images [Scrambled vs. Intact: $F(1,15) = 1.07$, $p > .1$] or between the face and body categories

[$F(1,15) = 1.19$, $p > .2$] was recorded during this time window (440–530 msec). Analysis of the distortion effects revealed that the distorted configurations induced a decrease in amplitude of the posterior positivity [$F(1,15) = 5.11$, $p = .04$] over both hemispheres [Condition \times Hemisphere: $F(1,15) < 1$] (see Figure 4). The effect was larger for bodies [$F(1,15) = 4.22$, $p = .058$] than for faces [$F(1,15) = 1.91$, $p = .18$] but the Condition \times Category interaction was not significant ($p = .83$).

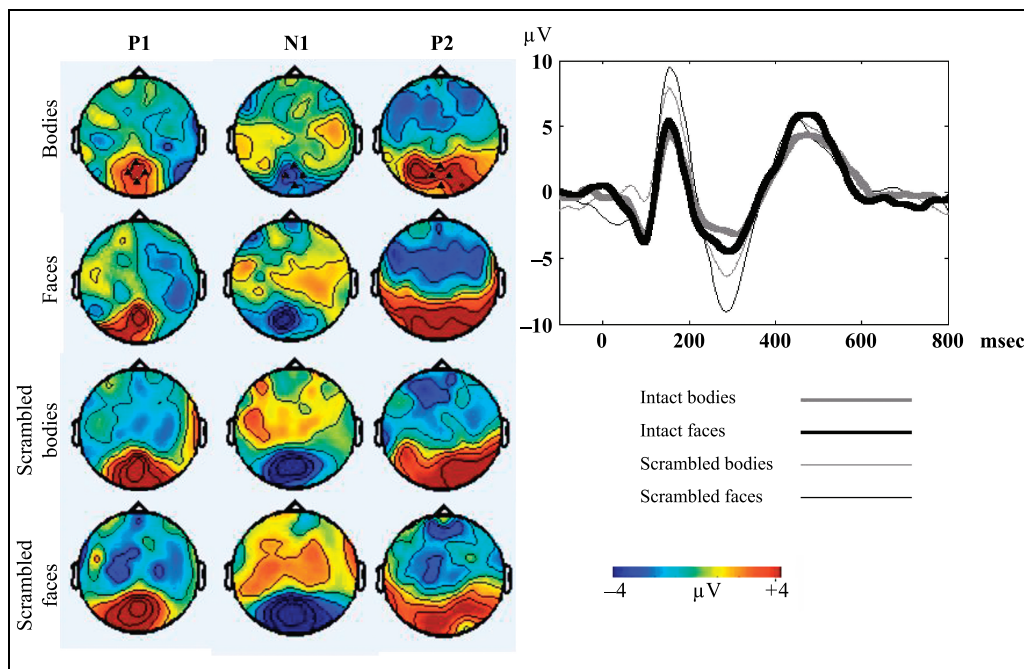
Discussion

Although in adults scrambled objects elicited no N1, confirming that this component reflects the activity of object-specific areas, we obtained different results in infants. At 3 months of age, the N290 was not only present but was even larger for scrambled images than for images of intact objects. The scrambled stimuli were similar in mean luminance and hues to the intact images. Nevertheless, the scrambling procedure used here introduced differences in the quantity of the high-contrast borders (more in the scrambled images) and in total area occupied by the image (the scrambled images occupy the entire space of the stimulus). N290 thus appears more sensitive to these low-level visual factors. The infant and adult N1s also differ in their topography: N290 is very focal and develops over the medial occipital electrodes, whereas the adult N1, for both faces and bodies, extends laterally over the occipital and temporal electrodes.

In adults, retinotopic areas (such as V1 and V2) show increased responses to scrambled objects (Allison, 1999; Grill-Spector, Kushnir, Avidan, Itzchak, & Malach, 1999). Even object-related areas, in the lateral occipital cortex (LOC), manifest a gradual sensitivity to scrambling, along an anterior–posterior axis (Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001). Although anterior areas respond preferentially to intact objects, more posterior regions are still relatively active for moderately scrambled images, suggesting that these later areas are thus more sensitive to local aspects of objects. The differences between N1s to scrambled and intact objects in adults and 3-month-old infants suggest either a different origin of N1 at the two ages (LOC in adults and V1–V2 in infants), or that more extended parts of the visual cortex in infants are still responsive to local features and might not have converged to more global object representations.

Functional parallels have been drawn between the processes underlying N1 components in adults and 3-month-old infants because at both ages species of the stimulus (human vs. monkey) affects the amplitude of this component (de Haan et al., 2002). However, differences in local structural properties might also explain these effects. Different texture in human and monkey faces or a higher contrast between the white sclerotic

Figure 3. Voltage maps of infants' ERPs evoked by bodies, faces, scrambled bodies, and scrambled faces, at the maximum of the P1 (180 msec), N1 (300 msec), and P400 (460 msec) components. The last column shows the grand-averaged waveform summed over a group of occipital electrodes (marked by triangles on the voltage maps).

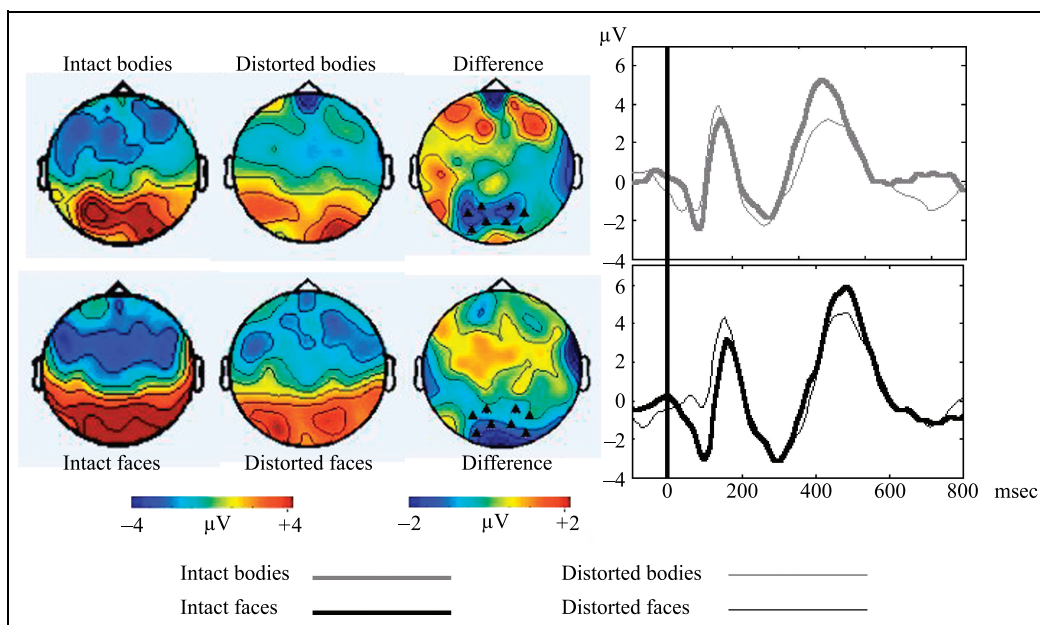


and the iris of human eyes that is not present in monkey eyes (Kobayashi & Kohshima, 1997) could explain the larger response to human faces than to monkey faces. The same mechanism can also explain the difference observed at this latency between directed and averted gaze in 4-month-olds (Farroni et al., 2002): More white sclerotic is seen in the averted gaze than in the directed gaze. When only the orientation of the faces is varied, but not other low-level properties, N1 is not affected, at least not at 3 months (Halit et al., 2003). Thus, we need more evidence to understand precisely what kind of visual processing underlies N1 in young infants and what

drives the rapid development observed over the first year of life (Halit et al., 2003).

The second component of interest in infants, the P400 response, was different in topography from the preceding component, extending over the medial and lateral posterior electrodes. This component was modulated by the configuration of both faces and bodies. Starting around 450 msec after onset, distorted faces and bodies evoked weaker responses than the intact original images. Our results indicate that 3-month-old infants' knowledge of human visual features does not only concern faces but also the human body. It is

Figure 4. Image distortion effects observed in infants at the level of P400. Voltage maps of ERPs evoked by intact and distorted images and their difference are shown (first row: bodies; second row: faces). The last column displays the grand-averaged waveforms summed over the group of electrodes that is marked by triangles on the difference maps.



remarkable that infants do not pay attention only to the dynamic aspects of the body but also to the static, structural aspects as well.

Face and body distortion effects were similar in latency, proving comparable facility in detecting the abnormalities. At the same latency, Halit et al. (2003) found a reduction in P2 amplitude when 3-month-old infants viewed another type of face with an unusual configuration, in this case, inverted faces. Two possible interpretations of the effect that we observe can be given. The first one would be that infants detect a mismatch between the image presented and a global face and body template. In the case of bodies, as was discussed in the previous section, we cannot determine which aspect of the body distortion was crucial. Did infants rely on a global body template or on the presence of an unusual body part at the place of the head? In the first months of life, infants frequently see upper body parts and could thus construct expectancies about the head's position in relation to the torso. Alternatively, the effects that we observe could be novelty effects. Infants could treat the distorted images not as modified versions of the known "face" and "body" categories, but as new objects. Further experiments are thus needed to precisely characterize the initial face and body representations in infancy. Whatever the significance of these responses is, they demonstrate that infants give a special status to the normally configured faces and bodies.

GENERAL DISCUSSION

The main aim of this study was to draw parallels between body perception and face perception mechanisms, with an emphasis on the first-order configural processing.

We show that in adults, configural processing is an early step in the processing of objects other than faces, namely, bodies. Are the strong N1 responses recorded in the adult study due to a special status of faces and bodies? The detection of bodies, along with that of faces, could have been subject to evolutionary pressure for speed and accuracy, resulting in the emergence of specialized neuronal mechanisms. Alternatively, experience with bodies could lead to the development of the body specialization, just as reading experience leads to a word-evoked N1, with a similar latency to that for faces, but originating in more posterior parts of the ventral temporal cortex (Puce, Allison, & McCarthy, 1999). We cannot answer this question as of yet. However, the present experiment allows us to enlarge the group of candidate objects for which specialized processing is seen. Further study of processing of these objects will eventually help us to assess the validity of each of these hypotheses.

In addition, we showed for the first time that knowledge of body configuration is present early in life. Our results support previous studies on human biological

motion (Bertenthal et al., 1987) and show that infants are capable of discovering structural regularities in the environment for other types of common objects than faces. Slaughter et al. (2002) fail to find evidence of such knowledge until 18 months of age when using symmetrical distortions of bodies without altering the head position. The success of our experiment, in which we used distortions that affected body symmetry and interfered with the normal head position, suggests that the former are important parameters used for body perception, both in infancy and in adulthood.

At this point, we need more research to understand the cerebral bases of object recognition and its development. The large P1 and N290 responses obtained for scrambled stimuli suggest that the spatial segregation of the neuronal circuits involved in object perception is not present from the beginning of the development. As has been proposed previously (de Haan, 2003), object-specific areas may initially respond to a wide range of stimuli. Moreover, initial object processing is probably more sensitive to the low-level visual properties of objects. Thus, the origin of infant ERPs is not so clear-cut as it has been believed.

Do our results provide any evidence for the existence of innate knowledge of face and body structure? At three months of age, infants' vision is far from the adult performances but still good enough to allow them to perceive the structural parameters that were manipulated in our experiment (Farroni et al., 2002). Thus, even with poorer vision during the first 3 months of life, we cannot exclude the possibility that infants could have gathered information on face and body structure, at least the information that is needed in order to choose the normal configurations among the face and body images that they saw in the present experiment.

The present data suggest that processing of the human body is similar in many aspects to that of the human face, at least with respect to the structural encoding of these two categories. Experiments in infants have, until now, explored few categories of objects other than human faces. We hope that this first step toward a more extensive study of object perception in infancy will encourage further comparative studies. This should allow a better understanding of those parameters that drive the development of the visual processing areas, leading to the neuronal specialization that is seen in adulthood.

METHODS

Participants

Sixteen full-term infants were tested between 12 and 14 weeks after birth (mean age 12.8 weeks; 5 girls). Seventeen additional infants were rejected for excessive movement or fussiness. Ten adults (mean age 25 years; 7 women) were also tested. The study was

approved by the regional ethical committee for biomedical research. Adult subjects and parents gave their written informed consent.

Stimuli

Stimuli were 16-color gray-scale images, presented on a gray background. Thirty-five different front-view faces and bodies were used. The body images depicted women gymnasts in different positions with arms and legs clearly visible, from which the heads were removed. The abnormal faces were obtained by displacing their internal elements. Abnormal bodies were obtained by displacing a hand or a leg to the place of the head. These transformations left the low-level properties of the images (luminance, contrast, occupancy, local features) intact while destroying its normal configuration. Scrambled images were obtained from the same initial images, by partitioning and scrambling a 25×25 square matrix for each image. Six categories of images were thus used: intact faces, intact bodies, distorted faces, distorted bodies, scrambled faces, and scrambled bodies.

Procedure

Participants viewed the images projected on a screen situated at 120 ± 10 cm (for the adults) or 80 ± 10 cm (for the infants), subtending $18^\circ \times 18^\circ$ of visual angle for the adults and $25^\circ \times 25^\circ$ for the infants. Images were randomly presented, with an ISI of 1500 msec (EXPE6 software: www.lscp.net/expe). No blank screen was present between two successive images, therefore, no offset signature can be seen in the ERPs. Each image was presented twice. The second presentation of the images occurred only after the whole set had been seen once. The entire presentation of images included 70×6 images and lasted 10 min. The presentation of the images was stopped every time the infants looked away and restarted after infants' attention was redirected to the screen. Pauses took place whenever the infant needed comforting. In adults, pauses were fixed every 50 images to avoid blinking artifacts. Adults did not perform any active task; they were instructed to fixate the images and avoid ocular movement.

ERP Recording

Continuous EEG was recorded with a geodesic electrode net (EGI) referenced to the vertex (129 carbon electrodes for the adults, 65 for the infants). The net was applied in anatomical reference to the vertex and the cantho-meatal line. Scalp voltages were recorded during the entire experiment, amplified, filtered between 0.1 and 40 Hz, and digitized at 125 Hz. Then, the EEG was segmented into epochs starting 400 msec before the image onset and ending 1600 msec after it. These epochs were automatically edited to reject trials con-

taminated by eye or body movements. All trials with more than 25 bad channels (infants) or 20 bad channels (adults) were rejected. The electrodes that had been bad all along the experiment have been excluded from the analysis, without interpolation (max 5 in infants). In the infant experiment, every trial that preceded a pause was rejected before averaging. The artifact-free trials were averaged for each subject in each of the six categories: faces, distorted faces, scrambled faces, bodies, distorted bodies, and scrambled bodies. In infants, about 30 trials were kept for each condition (30, 30, 29, 30, 30, and 30, respectively). Averages were baseline corrected, transformed into reference-independent values using the average reference method, and digitally filtered between 0.5 and 20 Hz. Two-dimensional reconstructions of scalp voltage at each time step were computed using a spherical spline interpolation.

Data Analysis

Adults

Two components were examined, N1 and P2.

N1. PEAK MEASURES. Voltages from 10 occipito-temporal sensors, on each hemisphere, around T5 (50 51 52 57 58 59 60 61 63 64 65 66 67 71 72) and T6 (84 85 90 91 92 97 96 98 101 102), were averaged for each condition. N1 peak amplitude and peak latency were determined as the minimum voltage and the time point of the minimum voltage, in the 160–260 msec interval. ANOVAs were performed on these two variables, with Category (bodies and faces), Condition (intact, distorted and scrambled), and Hemisphere (left and right), as within-subject factors.

MEAN AMPLITUDE. Voltage was averaged across the 160–260 msec time window and entered in an ANOVA with the same factors as above.

TOPOGRAPHY ANALYSIS. To analyze the topographical differences between categories, groups of six electrodes, and their symmetrical sensors on the other hemisphere, located over the positive and negative maxima of the dipole response for each category, were chosen (locations equivalent to FP2: 15 9 2 10 3 8; FP1: 18 23 27 19 24 26; F4: 4 124 5 119 113 107; F3: 20 12 25 21 13 7; T6: 85 84 83 90 91 77; T5: 66 71 70 75 65 72; O2: 92 97 96 98 101 93; O1: 59 51 57 64 58 52). Forty-eight electrodes were thus considered for this analysis: 6 electrodes \times 2 maxima \times 2 categories \times 2 hemispheres. The voltage from the six sensors was averaged for each of the four locations and two hemispheres and across a 60-msec time window surrounding the peak of the N1 response to faces (204 msec) and bodies (228 msec). This variable was entered in an ANOVA with Electrodes (4 groups), Hemisphere (left, right), Category (face, body), and Condition (intact, distorted, scrambled) as within-subject

factors. The vector scaling procedure described in McCarthy and Wood (1985) was secondarily applied in order to confirm the electrode site effect.

P2. Two groups of 10 electrodes located at the maximum of the positivity and the second at the maximum of the negativity of the P2 response, and their symmetrical sensors (15 9 2 10 3 4 124 5 119 113; 18 23 27 19 24 20 12 25 21 13; 85 84 83 90 91 92 97 96 98 93; 66 71 70 75 65 59 51 64 58 52), were chosen. The voltage was averaged across the sensors in each location and across the time window surrounding the peak of the P2 (300–400 msec). An ANOVA was computed on this variable with the within-subject factors: Electrodes (2 groups), Hemisphere (left, right), Category (face, body), and Condition (intact, distorted, scrambled).

Infants

Three components were analyzed: P1, N290, and P400.

Two symmetrical groups of four electrodes, over the left (33 32 36 37) and over the right temporo-occipital regions (41 40 44 45), were considered. Voltage was averaged across sensors for each Location (right and left), Category (body and face), and Condition (scrambled, intact and distorted). For each time window, P1 (120–220 msec), N290 (240–360 msec), and P400 (440–530 msec), the peak amplitude and latency, as well as the mean amplitude, were calculated. Each of these dependent measures was entered in an ANOVA, with Hemisphere (left, right), Category (faces, bodies), and Condition (intact, distorted, scrambled) as within-subject factors.

Acknowledgments

This study was supported by a McDonnell Foundation grant.

Reprint requests should be sent to Teodora Gliga, Unité INSERM 562, Service Hospitalier Frédéric Joliot, CEA/DRM/DSV, 4 place du général Leclerc, 91401 Orsay cedex, France, or via e-mail: gliga@clipper.ens.fr.

Notes

1. Peak amplitude measures confirm these results: intact versus scrambled, $F(1,9) = 21.52, p = .001$.
2. Peak amplitudes for faces and bodies were similar [$F(1,9) < 1$] and in both cases the N1 response was right lateralised [Interaction with Hemisphere for bodies $F(1,9) = 12.86, p = .006$ and faces $F(1,9) = 20.05, p = .001$].
3. Peak amplitudes: Scrambled vs. Intact P1: $F(1,15) = 13.50, p = .002$ and N290: $F(1,15) = 10.40, p = .006$; Scrambled faces vs. Scrambled bodies P1: $F(1,15) < 1$ and N290: $F(1,15) = 1.71, p = .2$.

REFERENCES

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face

- perception: I. Potentials generated in the occipitotemporal cortex by face and non-face stimuli. *Cerebral Cortex*, *9*, 415–430.
- Arroyo, S., Lesser, R. P., Poon, W. T., Webber, W. R., & Gordon B. (1997). Neuronal generators of visual evoked potentials in humans: Visual processing in the human cortex. *Epilepsia*, *38*, 600–610.
- Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience*, *7*, 542–548.
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate module? Evidence from human neonates. *Infant Behaviour and Development*, *23*, 223–229.
- 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.
- Bertenthal, B. I. (1996). Origins and early development of perception, action, and representation. *Annual Review of Psychology*, *47*, 431–459.
- Bertenthal, B., Proffitt, D., Kramer, S., & Spenter, N. (1987). Infants' encoding of kinetic displays varying in relative coherence. *Developmental Psychology*, *23*, 213–230.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.
- Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83*, 1–29.
- Cashon, C. H., & Cohen, L. B. (2003). The construction, deconstruction and reconstruction of infant face perception. In O. Pascalis & A. Slater (Eds.), *The development of face processing in infancy and early childhood*. New York: NOVA Science.
- Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, *304*, 115–117.
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognising friend by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, *9*, 353–356.
- de Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: A review. *International Journal of Psychophysiology*, *51*, 45–58.
- de Haan, M., Johnson, M. H., Maurer, D., & Perett, D. (2001). Recognition of individual faces and average face prototype by 1- and 3-month-old infants. *Cognitive Development*, *16*, 659–678.
- de Haan, M., & Nelson, C. (1999). Electrocortical correlates of face and object recognition by 6-month-old infants. *Developmental Psychology*, *35*, 1113–1121.
- de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14*, 199–209.
- Diamod, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473.
- Eimer, M. (2000a). Effects of face inversion on the structural encoding and recognition of faces. Evidence from event-related brain potentials. *Cognitive Brain Research*, *10*, 145–158.
- Eimer, M. (2000b). The face-specific N170 component reflects late stages in the structural encoding of faces. *NeuroReport*, *11*, 2319–2325.

- Farah, M. J. (1996). Is face recognition “special”? Evidence from neuropsychology. *Behavioural Brain Research*, *76*, 181–189.
- Farroni, T., Csibra, G., Simion, F., & Johnson, M. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences, U.S.A.*, *99*, 9602–9605.
- Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, *6*, 428–432.
- Gauthier, I., & Tarr, M. J. (2002). Unraveling mechanisms for expert object recognition: Bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 426–432.
- 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.
- Gliga, T., & Dehaene-Lambertz, G. (submitted). Development of view-invariant representations of faces in humans.
- Grill-Spector, K., Kushnir, S. E., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, *24*, 187–203.
- Hadjikhani, N., & deGelder, B. (2003). Seeing fearful body expressions activates the fusiform cortex and amygdala. *Current Biology*, *13*, 2201–2205.
- Halit, H., deHaan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *NeuroReport*, *11*, 1871–1875.
- Halit, H., deHaan, M., & Johnson, M. H. (2003). Cortical specialization for face processing: Face-sensitive event-related potential components in 3- and 12-month-old infants. *Neuroimage*, *19*, 1180–1193.
- Henson, R., 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.
- Itier, R. J., & Taylor, M. J. (2004a). Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. *Neuroimage*, *21*, 1518–1532.
- Itier, R. J., & Taylor, M. J. (2004b). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex*, *14*, 132–142.
- Itier, R. J., & Taylor, M. J. (2004c). Source analysis of the N170 to faces and objects. *NeuroReport*, *15*.
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns’ preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*, 1–19.
- 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.
- Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature Neuroscience*, *387*, 767–768.
- Lerner, Y., Hendlar, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in human visual cortex. *Cerebral Cortex*, *11*, 287–297.
- Lewis, M. B., & Edmonds, A. J. (2003). Face detection: Mapping human performance. *Perception*, *32*, 903–920.
- Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience*, *5*, 910–916.
- Maurer, D. (1985). Infants’ perception of facedness. In T. M. Field & N. Fox (Eds.), *Social perception in infants* (pp. 73–100). Norwood, NJ: Ablex Publishing.
- Maurer, D., & Barrera, M. (1981). Infants’ perception of natural and distorted arrangements of a schematic face. *Child Development*, *52*, 196–202.
- Maurer, D., Grand R. L., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*, 255–260.
- McCarthy, G., & Wood, C. C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of brain electrical activity. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 74–78.
- Münter, T. F., Brack, M., Grootheer, O., Wieringa, B. M., Matzke, M., & Johannes, S. (1998). Event-related potentials to unfamiliar faces in explicit and implicit memory tasks. *Neuroscience Research*, *28*, 223–233.
- Paller, K. A., Gonsalves, B., Grabowecky, M., Bozic, V. S., & Yamada, S. (2000). Electrophysiological correlates of recollecting faces of known and unknown individuals. *Neuroimage*, *11*, 98–110.
- 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.
- Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, *14*, 302–308.
- Rolls, E. T. (2000). Functions of the primate temporal lobe cortical visual areas in invariant visual object and face recognition. *Neuron*, *27*, 205–218.
- Rossion, B., Gauthier, I., Goffaux, V., Tarr, M. J., & Crommelnick, M. (2002). Expertise training with novel objects leads to left-lateralized facelike electrophysiological responses. *Psychological Science*, *13*, 250–257.
- Rossion, B., Dricot, L., Devolder, A., Bodart, J.-M., deGelder, B., & Zoontjes, R. (2000). Hemispheric asymmetries for whole-based and part-based face processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *12*, 793–802.
- Rousselet, G. A., Mace, M. J. M., & Fabre-Thorpe, M. (2004). Animal and human faces in natural scenes: How specific to human faces is the N170 ERP component? *Journal of Vision*, *4*, 13–21.
- Sagiv, N., & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part-based processes. *Journal of Cognitive Neuroscience*, *13*, 937–951.
- Slaughter, V., Heron, M., & Sim, S. (2002). Development of preferences for human body shape in infancy. *Cognition*, *85*, B71–B81.
- Stekelenburg, J. J., & deGelder, B. (2004). The neural correlates of perceiving human bodies: An ERP study on the body-inversion effect. *NeuroReport*, *15*, 777–780.
- Taylor, M. J., & Baldeweg, T. (2002). Application of EEG, ERP and intracranial recordings to the investigation of cognitive function in children. *Developmental Science*, *5*, 318–334.
- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Delliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 11656–11661.
- Valentine, T. (1988). Upside-down faces: A review of the effects of inversion upon face recognition. *British Journal of Psychology*, *79*, 471–491.