Development of a view-invariant representation of the human head

Teodora Gliga [a,*], Ghislaine Dehaene-Lambertz [a,b]

[a] *CNRS, Unité INSERM 562, Service Hospitalier Frédéric Joliot, CEA/DRM/DSV,*
4 place du général Leclerc, 91401 Orsay Cedex, France
[b] *Service de Neuro-Pédiatrie, CHU Bicêtre, AP-HP, France*

Abstract

Do infants perceive visual cues as diverse as frontal-view faces, profiles or bodies as being different aspects of the same object, a fellow human? If that is the case, visual exposure to one such cue should facilitate the subsequent processing of the others. To verify this hypothesis, we recorded event-related responses in 4-month-old infants and in adults. Pictures of eyes were interleaved amongst images belonging to three human contexts (frontal-view faces, profiles or bodies) or non-human contexts (houses, cars or pliers). In adults, both profile and frontal-face contexts elicited suppression of the N170 response to eye pictures, indicating an access to a view-invariant representation of faces. In infants, a response suppression of the N290 component was recorded only in the context of frontal faces, while profile context induces a different effect (i.e., a P400 enhancement) on eye processing. This dissociation suggests that the view-invariant representation of faces is learned, as it is for other 3-D objects and needs more than 4 months of exposure to be established. In a follow-up study, where infants were exposed to a short movie showing people rotating their heads, the profile-induced P400 effect was speeded up, indicating that exposure to successive views of the same object is probably a way to build up adult-like face representations.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Infant face perception; View invariance; Response suppression; ERPs; N170

---

* Corresponding author.
  E-mail address: gliga@clipper.ens.fr (T. Gliga).

1. Introduction

Knowledge acquisition in different domains (language, individual recognition, and theory of mind) depends on participating in social interactions, from the very first moments of life. In order to benefit from social contact infants have to be able to easily detect another human being's presence in the visual scene despite the variable conditions of distance, luminosity, view-point, and partial occlusion. Faced with this task, adult subjects only need 350 ms to answer correctly, even when presented with very complex and variable pictures (close-ups but also wider scenes where the human silhouette was only partially seen) (Rousselet, Mace, & Fabre-Thorpe, 2003). Years of experience made humans very proficient at this task and supplementary practice does not improve the performance (Fabre-Thorpe, Delorme, Marlot, & Thorpe, 2001).

The neuronal properties of the face-selective neurons found in the infero-temporal cortex of monkeys give us hints about how a system that does a rapid categorization of an object as a human being, whatever the point of view, might be organized. Numerous electrophysiological studies have shown that faces activate a distributed neuronal network (Young & Yamane, 1992). The face-specific neurons generally fire more strongly for a small subset of the studied faces. Nevertheless, in the first part of their response, the tuning is more general, discriminating only between the face category and an object category, for example (Sugase, Yamane, Ueno, & Kawano, 1999). On the other hand, different head views activate adjacent, superposed, cortical columns (Tanifuji, Tsunoda, & Yamane, 2002). While most of these neurons are view-specific, some of them are activated by all head views, being thus invariant with respect to the in-depth rotation (Perrett, 1985). Combining the coarse coding of faces at a category level with the strongly interconnected representations of the different head poses could allow the rapid view-invariant detection of another human being in a visual scene (Thorpe, Fize, & Marlot, 1996).

It was shown nevertheless that view-invariant representations of complex objects are found in the infero-temporal cortex of monkeys only after extensive training with different views of the object or after manipulation of the object during several months (Logothetis, Pauls, Bulthoff, & Poggio, 1994; Lueschow, Miller, & Desimone, 1994). There is considerable controversy, however, as to whether face recognition requires similar intense learning, during infant development. Sensorial or attentional biases give a special status to humans among other objects. Humans are not only the most frequent objects in infants' visual environment, but are also associated with multimodality cues, such as speech, spontaneous motion, smells, hunger, and discomfort appeasement. In a violation of expectation study, Bonatti, Frot, Zangl, & Mehler (2002) show that while 10-month-old infants would accept that a duck-toy becomes a truck, after passing behind a screen, they are surprised if one of the transforming objects is a doll (a human being). The authors interpret their results as a special status given, early in life, to the "human beings" category, which is distinguished from the multitude of other object categories.

Indeed, young infants manifest a lot of knowledge on the visual aspects of their conspecifics. They show preference for face-like stimuli from birth (Johnson, Dziurawiec, Ellis, & Morton, 1991). By 2 months of age, they can discriminate between normally configured and distorted faces (Maurer & Barrera, 1981) or between upright and inverted faces (Halit, deHaan, & Johnson, 2003) and are able to recognize a particular face after only a few minutes of exposure (Blass & Camp, 2003). At 6 months, infants recognize an individual that is facing them after seeing a 3/4 and a profile view (Pascalis, de Haan, Nelson, & de Schonen, 1998). To succeed in this task they should, by implication, be familiar with the transformation that a human face undergoes when rotating. Proprioceptive information supplements visual information when learning about the human body. A number of studies showed that newborns are capable of matching their own movement (mouth opening, tong protrusion, head turning, and finger movement) with that of an adult actor (Meltzoff & Moore, 1977, 1989). While the underlying mechanism allowing this matching is still unknown, such an understanding of another person's human body through their own body properties can only help infants integrate the visual transformations of a conspecific in motion.

This special status given to conspecifics might accelerate infants' understanding of the transformations that a human head and body undergo when moving and the building of strongly interconnected representations for the various poses. In order to study the existence of such representations in infants and adults we used a paradigm based on context-induced effects. The rationale behind this paradigm is that the processing of a target image should be modulated by the presence of images belonging to a related context, relative to a neutral context.

We recorded and compared ERPs to a given set of 30 eye pictures, randomly presented amongst 70 context pictures, belonging either to a human or a non-human category (Fig. 1). We hypothesized that if a human context facilitates the processing of eyes, due to shared neuronal representations, ERPs signals triggered by the presentation of eye pictures in a human context should diverge from those recorded in the control context. In the first experiment, frontal-view faces were presented with houses as contexts. Because we were confident that infants will easily notice the direct similarity between the human faces and the human eyes, this first experiment was performed to validate the new experimental paradigm. The second experiment contains the critical contrast: a human profile vs a car context. If a modulator effect is found in this case it means that profiles and eyes are processed as associated features. As we will further see, different results are found for the infant and the adult groups. The third experiment, opposing human bodies and pliers, allows us to verify whether the context-induced effect in the second study is due to more general associations between the distinct human body parts or specific to the processing of the human head. The last two experiments presented are follow-up infant studies whose goal is to understand what the factors are that lead to the aforementioned developmental change in head perception.
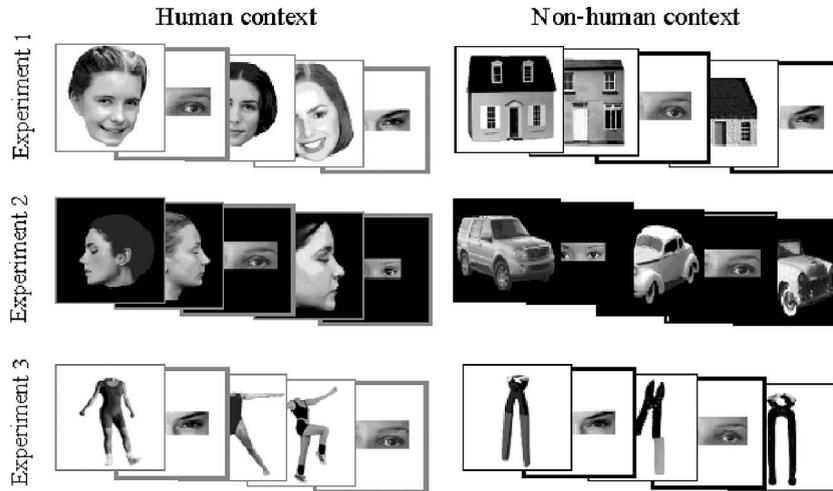
Fig. 1. Experimental design. Identical pictures of eyes were pseudorandomly presented amongst context pictures belonging to a human or to a non-human category. The human corporal context consisted of frontal-view faces in Experiment 1, profile faces in Experiment 2 and bodies (without heads) in Experiment 3 and contrasted with houses, cars and pairs of pliers. The block order was counterbalanced across subjects. Three groups of eight adults and 16 4-month-old infants were tested. By comparing responses to identical images (the eyes), appearing with the same probability in human and non-human contexts, we ensure that the observed differences can only be due to a different modulation of eye processing by the context images. Since eyes are not visible in the profile or whole body views, any observed differences would indicate an access to a more abstract representation.

2. General methods

2.1. Stimuli

The same 10 black and white images of eyes were used in all experiments and were taken from a set of faces, distinct from the frontal-view faces presented as human context in Experiment 1. In each experiment, 10 different human context pictures were paired with 10 structurally similar non-human pictures. There was no difference in mean luminance between the paired contexts ($ps > .1$).

2.2. Subjects

Different subjects were tested for each infant and adult experiment, in order to prevent long-term context effects (Cave, 1997). The study was approved by the regional ethical committee for biomedical research. Parents of the infants tested gave their written informed consent.

2.3. Procedure

All participants passively viewed images projected on a large screen located 120 ± 10 cm (adults) and 80 ± 10 cm (infants) away, spanning a visual angle of $18° \times 18°$ (adults) or $25° \times 25°$ (infants). Distance to screen was bigger in adults in order to decrease any discomfort caused by prolonged continuous visual stimulation. Each participant viewed two blocks, a human context and a non-human context block, each containing 70 context and 30 eye images. Block order was counterbalanced across subjects. In each block, images were pseudorandomly presented, with the only constraint that images of eyes were not consecutive. Each image remained on screen for 1500 ms. No blank screen was inserted between two successive images. Breaks were inserted whenever infants needed comforting and every other 50 images for adult experiments. Stimuli were presented using the EXPE software package (Pallier, Dupoux, & Jeannin, 1997).

2.4. ERP recording and data analysis

EEG was digitized continuously at 250 Hz using a high-density geodesic electrode net referenced to the vertex (65 electrodes in infants and 129 in adults). EEG was segmented into epochs starting 400 ms prior to image onset and ending 1600 ms after onset. Channels contaminated by eye or motion artifacts were automatically rejected. Trials with more than 25 contaminated channels or trials where the infant was looking away were rejected. The artifact-free trials were averaged for each participant in each of four trial types: human context, non-human contexts, target eyes in human context, and target eyes in non-human contexts. 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.

Analyses were performed on ERPs to the *same* eye images presented in a human or a non-human context. We focused on the ERP components known to be sensitive to faces/eyes. In adults, faces classically elicit a deep negativity (N170) on temporal electrodes with a reverse of polarity over parietal areas around 170 ms (Jeffreys, 1989; Bentin, Allison, Puce, Perez, & McCarthy, 1996; George, Evans, Fiori, Davidoff, & Renault, 1996). The N170 was studied on a group of seven electrodes centered on T5 and T6, as it is classically done. The peak value was entered in an ANOVA test with Condition (Human vs. Non-human context) and Hemisphere (Right vs. Left) as within-subject factors. A similar analysis was computed on the latency of the peak. To confirm the statistical significance of the results, a non-parametric test (the Wilcoxon signed rank test) was performed as well.

In infants, two components have been related to face perception N240 and P400 (de Haan, Johnson, & Halit, 2003). Thus, two time windows were analyzed: 200–300 ms (N1 or N240) and 540–600 ms (end of the P400). Voltage was averaged on two groups of three occipital electrodes (medial and lateral) in each hemisphere during these time windows. ANOVAs were computed with the same factors as in adults:

Condition and Hemisphere, and the additional factor Location (Medial vs. Lateral). The groups of electrodes analyzed in infants and adults are equivalent to those used by de Haan, Pascalis, & Johnson (2002).

3. Experiment 1: Eye exemplar-invariant representations

The actual imaging techniques measure mean activity over large patches of cortex, and therefore have no access to representations that span only limited cortical areas, like for example the face-processing units described in monkeys, which only have half a millimeter diameter (Tanaka, 2000). Repetition or priming paradigms have been proposed as a tool to circumvent this intrinsic spatial limitation (Naccache & Dehaene, 2001). These paradigms are based on the observation that activity, measured with functional MRI or ERPs, decreases when the same stimulus is repeatedly presented (see Dehaene-Lambertz & Dehaene (2004) for an example in ERP domain and Grill-Spector et al. (1998) in fMRI). This observation has been brought together with two other facts also observed when a stimulus is repeated, on one hand facilitation of behavioral responses and on the other hand, at the cellular level, decrease in firing of neuronal units (Desimone, Albright, Gross, & Bruce, 1984; Miller, Gochin, & Gross, 1993). More importantly, the same observations are made when the repeated stimuli are not exactly similar but only share a common property, suggesting that it is possible to target different levels of representation by varying the property that is repeated. Repetition paradigms have been used in this way to characterize the nature of the processing steps in visual object perception (Grill-Spector & Malach, 2001), word perception (Dehaene, Cohen, Sigman, & Vinckier, 2005) or phoneme perception (Dehaene-Lambertz & Gliga, 2004).

In this approach, the first step consists of isolating the network involved in the object representation by studying which components in ERP, or which brain regions in fMRI, display repetition suppression when the same object is repeated. Then in successive experiments, shared properties between the context and the target objects are varied in order to identify which coding variations are relevant and which are not relevant to obtain a repetition suppression effect on the same component or in the same brain region. In the first experiment, we thus presented eyes amongst frontal-view faces and studied two populations, adults and 3- to 4-month-old infants. Eyes were chosen as targets because when scanning a face, infants spend most of the time looking at eyes (Haith, Bergman, & Moore, 1977). Moreover, the ERPs to eyes are stronger than those to faces, from early childhood (Taylor, Edmonds, McCarthy, & Allison, 2001) until adulthood (Bentin et al., 1996).

In adulthood, face perception is associated with a prominent temporal–occipital negative wave with a reversal of polarity over the vertex, with a latency of 160–200 ms, the ''N170'' (Bentin et al., 1996; George et al., 1996; Jeffreys & Tukmachi, 1992). This ERP component is larger for faces (or face components, like eyes) than for other objects (Allison, Puce, Spencer, & McCarthy, 1999; Itier & Taylor, 2004). The N170 is triggered even in the absence of attention, suggesting that part of face processing is automatic (Carmel & Bentin, 2002). The underlying mechanisms are

mainly related to structural aspects of face perception because face inversion (Eimer, 2000) or distortion (Bentin et al., 1996; George et al., 1996) but not familiarity (Bentin & Deouell, 2000) modulate this component.

In infancy, a succession of two ERP components has been described, when perceiving faces – an occipital negativity with a latency of 250–300 ms (the ''N290'') and a more spread posterior positivity, which develops over a wider time interval (400–600 ms, the ''P400'') (de Haan et al., 2003). Due to its similarity in sign over the posterior electrodes the N290 was proposed to be the precursor of the adult N170 (de Haan et al., 2002). Nevertheless, only few infant studies compared face ERPs to other objects or to other human attributes ERPs. When compared to images that had a face-like contour, but whose interior was noise-like (obtained through phase scrambling), intact faces evoke a stronger N290 (Halit, Csibra, Volein, & Johnson, 2004) but a smaller N290 when compared to checkerboards (Gliga & Dehaene-Lambertz, 2005). Thus, the infant N290 component seems very sensible to low-level visual properties (to high contrast-borders or spatial frequencies) and not to a certain category, as in adulthood. When low-level properties are equated, within the face category (comparing upright and inverted faces (de Haan et al., 2002) or distorted and intact faces (Gliga & Dehaene-Lambertz, 2005)) the N290 shows no difference, while the later component, the P400, is enhanced for the intact, canonical faces. Thus, these two components reflect different steps in object processing in infancy.

While fMRI studies made extensive use of repetition paradigms, this approach is only starting to be employed with ERPs. Using a repetition paradigm, Jacques & Rossion (2004) observed that the face-evoked N170 was reduced in amplitude when a concurrent, distinct face was present on the screen. A general, exemplar-invariant, face representation is thus accessed at this stage. Therefore, we expect to see a similar suppression of the eye-evoked N170 in the context of front-view faces that would replicate Jacques & Rossion's (2004) findings. We are interested as well in seeing whether image repetition induces response suppression in infancy. Finding a similar behavior in both populations will encourage the use of this paradigm with young infants.

3.1. Methods

3.1.1. Stimuli

Ten different grayscale images representing front-view faces and houses, in two different sizes, were used for each context. The image background was white.

3.1.2. Subjects

Eight right-handed adults and 16 infants were tested between 12 and 15 weeks after birth (mean age 14 weeks). The data from 23 additional babies were rejected because of excessive movement or fussiness.

3.1.3. ERP recording and data analysis

In adults, we recorded a mean of 171 artifact-free trials (63 face trials, 62 house trials, 23 eyes amongst faces, and 23 eyes amongst houses trials). In infants, a mean

of 176 artifact-free trials was recorded (65 face trials, 62 house trials, 25 eyes amongst faces trials, and 24 eyes amongst houses trials).

3.2. Results

3.2.1. Adult results

The sharp negative signal classically induced by eyes and faces in adults – the N170 – was recorded in our experiment over posterior temporal regions, with an inversion of polarity above the parieto-central areas. The amplitude of the N170 to eyes was reduced. A mean amplitude of $-0.61\ \mu V$ was recorded in the face contexts vs $-2.65\ \mu V$ in the house context inducing a main effect of Condition ($F(1,7) = 20.55$, $p = .003$; Wilcoxon $= 36$, $p = .007$). The N170 latency was also accelerated when presented in a front-view context relatively to the house context (208 ms vs. 216 ms, $F(1,7) = 5.73$, $p = .048$; Wilcoxon $= -27$, $p = .06$; Fig. 2). No significant interaction between Condition and Hemisphere was observed for amplitude nor for latency.

3.2.2. Infant results

In infants, the typical deflections induced by visual stimulation were recorded: an occipital negative signal, peaking between 200 and 300 ms (the "N290"), followed by a sustained and wide-spread positivity that develops over temporal-occipital areas, between 350 and 550 ms (the "P400"). The N290 to eyes was reduced in the context of front faces relative to the control context (main effect of Condition: $F(1,15) = 8.91$, $p = .009$; Fig. 3). No significant interaction between Condition and Hemisphere nor between Condition and Location (medial vs. central electrodes) was observed.

3.3. Discussion

As expected, in adults, the frontal face context induced a reduction in amplitude of the N170 component. The amplitude reduction was accompanied by a significant latency shift. A faster response in the related, human context is compatible with the behavioral facilitation that accompanies response suppression effects (Henson & Rugg, 2003). The relatively late latency of the eye N170 in the neutral context (the houses) is probably due to the continuous presentation that we used with no blank screen between pictures. When the same images are presented interleaved with a screen displaying only a fixation cross, latencies of 170–180 ms are observed (unpublished data).

As in adults, a front-view face context induced a strong suppression of the early negativity in infants – the N290. Eyes size and retinal position being slightly different between context and target images and belonging to different individuals, we can infer that the neuronal code possesses a certain degree of invariance even at 4 months of age. Our data thus bring the first proof for a human eye-category-specific neuronal representation in infancy.
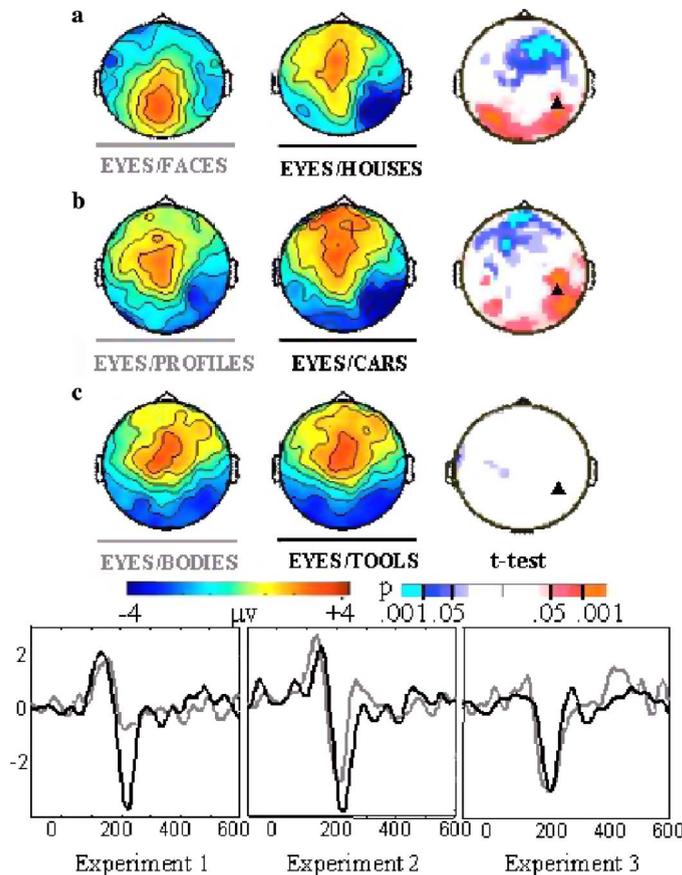
Fig. 2. N170 response suppression in adults observed in Experiment 1 (a) and 2 (b) but not 3 (c). The first and the second columns show the voltage maps of ERPs to the same eyes pictures in a human context and a non-human context at the maximum of the N1 and the third column displays the topography of the significant differences between these two conditions (*t*-test maps). The last row shows the grand-averaged waveform, recorded at a right temporal-occipital electrode (marked by a triangle on *t*-test maps), for each of the experiments.

The similarity between the infant and adult effects could be seen as the signature of an early active "eye" detector (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000; Farroni, Csibra, Simion, & Johnson, 2002). It was repeatedly shown that eye gaze modulates the N290 in infants. Directed gaze evokes a stronger N290 response (Farroni et al., 2002) than averted gaze. This effect might be related to the higher contrast between the white sclerotic and the iris present in direct gaze and absent in averted gaze. However, when all these differences are equated by using 3/4 views, this effect is still present, suggesting a genuine modulation of the N290 by gaze direction (Farroni, Johnson, & Csibra, 2004). While across development, eyes evoke a stronger N290 or N170 response than faces (Taylor et al., 2001 and in our own
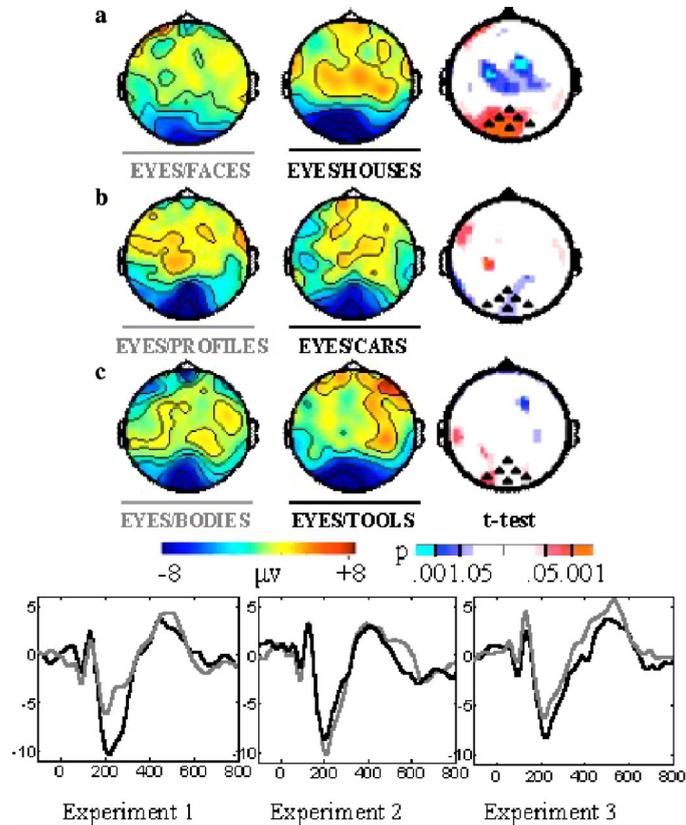
Fig. 3. N290 response suppression in infants observed in Experiment 1 (a) but not 2 (b) and 3 (c) (same format as Fig. 2). The last row shows the grand-averaged waveforms, summed over a group of occipital electrodes (marked by triangles on *t*-test maps), for each of the experiments.

study), it is no more the case in adults. More data are needed to clarify the relationship between the eye-processing mechanisms taking part in the N170 and the N290.

4. Experiment 2: View-independent head representations

Having found that the N1 in infants (N290) and in adults (N170) is sensitive to repetition, we can vary the context images in order to characterize the properties of the representation that is computed at both ages at the latency of the N1 component. Our second experiment questioned whether two different head poses share a common representation. If a response suppression of the eye responses is observed in a profile context, with respect to a non-related context, we can conclude that it was not only an eye representation but also a face representation, independent of the view, that is accessed in the first steps of visual object processing. Another possibility is that profiles and eyes do not share any common representations, but are

only semantically related. In that case, we expect no repetition suppression of N1, but a later effect when semantics is accessed. In adults, components related to semantic processing of pictures are recorded later around 400 ms (200–500 ms) over centro-parietal areas (Barrett & Rugg, 1990; Kiefer, 2001).

Three-month-old infants are familiar with the different appearances of a human head, as suggested by their capacity to recognize a person based on a frontal view as well as on a profile (Pascalis et al., 1998). Nevertheless, to our knowledge, there is as yet no proof that infants process profiles and front-view faces as related objects.

We therefore tested a new group of adults and infants, with the same experimental design that we used in the first study.

4.1. Methods

4.1.1. Stimuli

Ten different female profiles, with eyes closed, and 10 different 3/4 views of cars were used as context images. Half of the profile and car images had a rightward orientation, the other half a leftward orientation. In frontal views, hair clearly delimits the face while the clear contour of Caucasian profiles is less contrasted against a white background. Therefore, the image background was black in this experiment to increase image visibility, especially for infants.

4.1.2. Subjects

Eight right-handed adults and 16 infants (mean age 14.8 weeks) participated in this study. The data from 13 additional babies were rejected for excessive movement or fussiness.

4.1.3. ERP recording and data analysis

In adults, a mean of 159 trials was recorded (55 profile trials, 55 car trials, 24 eyes amongst profiles trials, and 25 eyes amongst cars trials). In infants, we recorded a mean of 130 artifact-free trials (48 profile trials, 47 car trials, 20 eyes amongst profiles trials, and 20 eyes amongst cars trials).

4.2. Results

4.2.1. Adult results

As in the front-view context, in the profile context the eye-evoked N170 was reduced in amplitude (-2 μV in the profile context vs. -3.6 μV in the car contexts, main effect of Condition: $F(1, 7) = 8.44$, $p = .023$; Wilcoxon = 34, $p = .02$) and had a shorter latency (209 ms vs. 224.5 ms, $F(1, 7) = 13,37$ $p = .008$; Wilcoxon = -33, $p = .002$). No significant interaction between the factors Condition and Hemisphere was recorded.

A closer look at the N170 suppression effect in both experiments shows that in frontal-view context it started earlier over the left hemisphere, became bilateral, and finished over the right hemisphere, while in profile context mainly the last part of the response, over the right hemisphere, was observed. These differences in onset

are demonstrated by a significant condition (human vs. non human) × experiment (front-view vs. profile) interaction during the 184–212 ms time window ($F(1,14) = 9,12$, $p = .009$) that became non-significant during the following 212–240 ms time window ($F(1,14) = 1,52$, $p > 0.2$).

4.2.2. Infant results

The infant N290 was *not* modulated by the profile context ($F(1,15) < 1$; Fig. 3) yielding a significant interaction of Condition (human vs. non-human) by Experiment (Experiment 1 vs. Experiment 2) ($F(1,30) = 7.18$, $p = .01$).

The profile context modulated response to eyes later on by extending the duration of the P400 signal (main effect of Condition: $F(1,15) = 4.96$, $p = .04$, Fig. 4). No significant interaction between Condition and Hemisphere nor Condition and Location was observed. Post hoc analysis show nevertheless that this effect was present predominantly at medial electrodes (medial electrodes: $F(1,15) = 8.82$, $p = .009$; lateral electrodes: $F(1,15) = 1.07$, $p > 0.3$). As for the first time window, there was a significant Experiment by Condition interaction at this latency ($F(1,30) = 3.82$, $p = .05$).

4.3. Discussion

In adults, the repetition suppression effect observed in a profile context suggests that a common view-invariant representation is primed both by frontal-view and profile faces. However, the delay of the repetition effect observed with profile context suggests two processing stages in adults: a first stage sensitive to the mere presence of eyes or to the frontal orientations of the images, beginning at 180 ms and followed
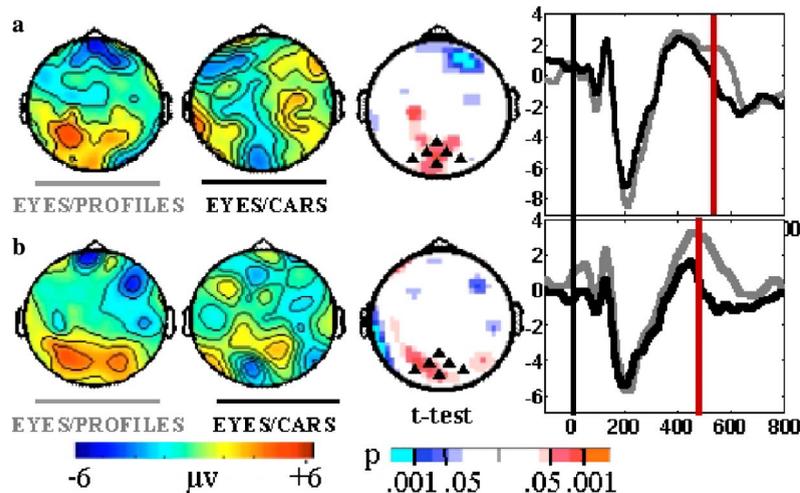

Fig. 4. P400 enhancement in infants in Experiment 2 (a) and after familiarization with rotating heads in Experiment 4 (b). A few minutes of presentation of a movie showing five women turning their head from left to right, with a pause at frontal view, speeds up the P400 enhancement in infants.

30 ms later by second stage sensitive to a context of faces, independently of their orientation.

At first glance, our results appear to contradict those brain imaging studies that found little evidence for view-invariant representations of faces. However, in contrast to our study, these studies did not look at view-invariant representations at a category level but at specific individual representations. Repeated presentations of different points of view of a face were compared to repetitions of faces belonging to different individuals (Grill-Spector, Kushnir, Avidan, Itzchak, & Malach, 1999; Pourtois, Schwartz, Seghier, Lazeyras, & Vuilleumier, 2005). This contrast revealed very little response suppression, and a strong dependency on familiarity (stronger effects for familiar images and located in the frontal lobe, thus corresponding more to a semantic and not to a perceptual effect) and on the angular distance between the repeated images (stronger effects for smaller angles) (Grill-Spector et al., 1999; Pourtois et al., 2005). Recognition of a new individual is thus based on viewer-centered representations, while for general categorization of an object as a human head, view-invariant representations are used, as suggested by our results. This difference might be due to the quantity of information needed for these two tasks; coarse information, based on the general geometry of the human head (Biederman, 1987) might be enough to detect a head, while finer information, which changes with the point of view, is used for individual recognition.

In infants, the profile context induced no repetition suppression of the N290 and there was a significant interaction between experiment and context during this time window, demonstrating that 4-month-old infants are not able to access a view-independent representation of faces. However, a later enhancement of the P400, relative to the control context, was recorded. A significant interaction between experiments and contexts was again present, demonstrating that no common stage was accessed in both experiments. An enhancement rather than a reduction of a component could be interpreted as an attentional effect (Luck & Hillyard, 1999) due to the fact that the presence of eyes in the visual scene may be expected more in a profile context than in a non-related car context. The fact that infants of this age associate two physically dissimilar images, such as profiles (with eyes closed) and eyes, contrasts with previous studies, which showed that in the first 6 months of life, object classification appears to be driven mostly by physical similarity (Behl-Chadha, 1996; Quinn, Eimas, & Rosenkrantz, 1993). Infants would categorize physically similar objects together, like cats (Quinn et al., 1993) or chairs (Behl-Chadha, 1996), but failed when they encountered too much variability (e.g., female lions would be categorized together with the cats but not with the male lions). Only at the end of the first year of life, infants would start associating physically distinct objects or object parts that are presented simultaneously (Cashon & Cohen, 2003; Fiser & Aslin, 2002). Despite these perceptual problems, infants succeed at associating profiles and frontal views of faces (here, eyes) at an age when they fail with equivalent object-processing tasks. The face-processing biases, which we detailed in the introduction, together with the frequent encounters with faces in general, might booster the association of the dissimilar head poses.

5. Experiment 3. Probing the selectivity of the previous N170 suppression effects

The frontal face and profile-induced response suppression on the eye-evoked N170 was interpreted in adults as the result of accessing eye exemplar-invariant and view-invariant head representations. However, this representation might not be limited to face but concern a more general property of "humanness", being thus invariant to the body part that is perceived. Nevertheless, it was shown that body and face processing involve distinct ventral temporal cerebral sites (Peelen & Downing, 2005). Moreover, the human body perception evokes a N170 component which has slightly different topography and latency than the face-evoked N170 (delayed by ∼30 ms), suggesting as well the involvement of different neuronal generators (Gliga & Dehaene-Lambertz, 2005). Based on these studies we do not expect a common neuronal mapping of human eyes and human body representations.

We also cannot exclude that the effect of repetition suppression that we observed for profiles might be related to top-down effects through a semantic matching of the repeated pictures, at higher levels of processing (Bar, 2003). Human body as well as profiles and eyes belong to a "human body parts" semantic category. In order to address this issue, we compared a human body context to a physically matched non-human context (pliers) in the third experiment. The absence of repetition suppression in this case, despite the fact that bodies and eyes are semantically related, would confirm that the N170 suppression is not induced by semantic proximity, being a genuine marker of the access to common perceptual representations.

In infants, no common mapping of the two head views was found in the previous experiment. The profile context enhanced the P400 component. As in adults this third experiment will help us to better define the specificity of this effect. Three-month-old infants are already familiar with the human body structure (Gliga & Dehaene-Lambertz, 2005). Is it the pairing of two familiar images (profiles and eyes or bodies and eyes) that induces the P400 enhancement?

5.1. Methods

5.1.1. Stimuli

The contexts consisted of 10 images depicting gymnasts' bodies, from which the head was removed and 10 different images of pliers. A white background was used for these images.

5.1.2. Subjects

Eight adults and16 healthy infants were tested between 12 and 15 weeks after birth (mean age 14.8 weeks). The data from 19 additional babies were rejected for excessive movement or fussiness.

5.1.3. ERP recording and data analysis

In adults, we recorded a mean of 154 artifact-free trials (53 body trials, 53 tool trials, 24 eyes amongst bodies trials and 24 eyes amongst tools trials). In infants,

we recorded a mean of 153 artifact-free trials (58 body trials, 55 tool trials, 21 eyes amongst bodies trials and 20 eyes amongst tools trials).

5.1.4. Analysis

A later, longer lasting effect was observed in this condition in adults. This effect was recorded over the left parietal electrodes with a reverse of polarity over the right frontal electrodes, in the 360–488 ms interval. To analyze this difference a three-factor ANOVA was computed on the voltage averaged over two groups of 18 electrodes (left parietal and right frontal), with Condition and Electrodes as within-subject factors and Experiment as between-subjects factor.

5.2. Results

5.2.1. Adult results

The body context did not affect the amplitude or the latency of the N170 evoked by eyes (amplitude $F(1,7) = 1.5$, $p > 0.2$, Wilcoxon = 28, $p = .19$; latency $F(1,7) < 1$, Wilcoxon = 13, $p = .5$), yielding a significant condition by experiment interaction (front-view and profile faces vs. bodies: N170 amplitude: $F(1,21) = 17.07$, $p < .001$; N170 latency: $F(1,21) = 5.64$, $p = .027$). A later effect developed in the 360–488 ms interval, as a stronger parietal positivity for the eyes in the non-human context, reversing over the frontal electrodes. This effect was significant, as shown by the interaction Condition by Electrodes ($F(1,7) = 18.15$, $p = .003$), and specific to this experiment (Experiment 1 and Experiment 2 vs. Experiment 3: $F(1,21) = 18.74$, $p = .003$).

5.2.2. Infant results

In infants, a body context effect was observed around the beginning of P400. However, this effect was weak. No main effect of condition was present. However, there was an significant interaction Condition X Location ($F(1,15) = 4.55$, $p = .04$) due to the fact that the effect was significant only over the lateral occipital electrodes, as shown by post hoc analysis (lateral electrodes: $F(1,15) = 4.51$, $p = .048$; medial electrodes: $F(1,15) < 1$). Although the contextual modulation of the P400 component appeared less important in this study compared to the profile-induced effect, the interaction Experiment (Experiment 2 vs. Experiment 3) and Condition was not significant.

5.3. Discussion

The absence of the adult N170 modulation by a semantically related context strengthens our hypothesis that the N170 response suppression in the first two experiments is due to specific access to head representations that are, in adults, both exemplar-invariant and invariant to the point of view. A later effect was nevertheless observed in the body context condition which was not found in the previous two experiments, the front view and the profile contexts at a re-analysis of the data. This effect developed as a stronger central positivity for the eyes in the non-related, tool context. A similar modulation of this late ERP component (also named late positive

component) was found for the presentation of semantically incongruent words and images (Sitnikova, Kuperberg, & Holocomb, 2003) and could thus be related to the semantic analysis of visual scenes. However, this effect was not present in the previous two experiments, where a semantic incongruity occurred as well (between the house or the car images and the eye images). It is therefore possible that in our study this component is not related to general semantic matching (no such judgments were demanded from the subjects), but to specific long-distance associations between the distinct representations of the human body and of the human eyes.

In infants, the body context had a small effect on the eye ERPs at the latency of the P400 component. Increases of the amplitude of the P400 have been observed in the infant literature for familiar relative to less familiar stimuli (e.g., known vs. new faces (Nelson, 2001), upright vs. reversed faces (de Haan et al., 2002) and normal vs. distorted faces and bodies (Gliga & Dehaene-Lambertz, 2005)). Nevertheless, based on the results of this third experiment, general familiarity cannot explain our results in infants. Since not only profiles but also bodies are familiar objects, we would in consequence expect a similar effect for the body vs tool contexts than the one we found in Experiment 2. Yet, the P400 signal enhancement was weaker and more lateral in Experiment 3 than in Experiment 2, suggesting that the link between the representations of the two head views, profile and front view, might be stronger than the association of the two body parts, head and trunk. While profile and eyes are seen in close temporal succession, eyes and bodies are more often associated in the same view. The observation that eyes are more expected after profile than after bodies images suggests that for infants to learn an association, the simultaneous exposure to the two objects is not sufficient and learning is constrained by other parameters. In the following experiment, we will try to manipulate one of these parameters.

6. Experiment 4: Short exposure to human heads in rotation facilitates profile–front-view face associations

A response suppression of the early eye response by profiles in adulthood but not in infancy was observed. Faced with such a striking developmental difference, investigating the possible factors that may lead to an adult-like face representation is inevitable. One of these factors is of course longer experience with human heads in rotation. It has been shown that two distinct views of an object are more easily associated if they are connected by a rotation movement (Kourtzi & Shiffrar, 1999) and that this kind of association leads to view-invariant object representations (Logothetis et al., 1994; Lueschow et al., 1994). Nevertheless, most of the initial caregiver–infant interactions involve a direct eye contact and young infants cannot move around a human being in order to notice its 3D properties. Thus, although frequently seeing profiles and frontal views of faces in close spatial-temporal relationships, they might have seldom had the chance of noticing that a profile can become gradually a frontal-view face. The goal of this fourth infant experiment is to see whether, given this kind of input, infants would be capable of using it to strengthen the link between the profile-view and the frontal-view representations. Alternatively, 4-month-old babies

could be too young to use such cues; a better understanding of 3D object in-depth rotation, through visual and manual exploration, being a prerequisite.

In order to find out whether young infants are sensitive to learning human head 3D properties through rotation, an additional group of infants was tested in the profile vs. car context experiment after 3 min of familiarization with a video containing five women turning slowly their heads from one side to the other.

6.1. Methods

6.1.1. Subjects
Sixteen infants (mean age 14.4 weeks) participated in this study. The data from 16 additional babies were rejected for excessive movement or fussiness.

6.1.2. Stimuli
For this experiment, five different women were video recorded while they were slowly moving their heads from left to right. Pauses were marked at the profile and front views. The resulting movie was 3 min long and contained two different sequences for each of the woman.

6.1.3. Procedure
Infants first watched the familiarization videos (i.e., 3 min) and were subsequently presented with the same stimuli set used in Experiment 2.

6.1.4. ERP recording and data analysis
We recorded a mean of 142 artifact-free trials for this experiment (51 profile trials, 54 car trials, 23 eyes amongst profiles, and 24 eyes amongst cars trials).

6.2. Results

In the familiarized group, after seeing a movie presenting five different women slowly turning their heads from left to right, the P400 enhancement started 60 ms earlier than was observed in Experiment 2 (Fig. 4). This led to a significant context effect during the 480–540 ms interval (main effect of Condition: $F(1, 15) = 7.4$, $p = .015$). In Experiment 2, no significant difference was found in this interval (main effect of Condition: $F(1, 15) = 1.96$, $p = 0.18$). The opposite observation characterizes the 540–600 ms interval during which a main effect of condition was observed ($F(1, 15) = 4.96$, $p = .04$) while this effect fades in the familiarized group ($F(1, 15) = 3.09$, $p = .09$). As in experiment 2, the P400 effect concerned both hemispheres and both electrode locations, as no significant interactions between these factors and Condition were found.

6.3. Discussion

This fourth infant experiment showed that while by 4 months of age infants' experience with human heads in rotation was not rich enough to allow the common map-

ping of different head views, they are nevertheless sensitive to this kind of information. Only 3 min of observing a human profile gradually turning into a frontal-view face were enough to speed up the P400 enhancement. There are certainly other types of learning that will accompany infants in their discovery of the 3D properties of the human body. We mentioned earlier the manual exploration of objects which was shown to lead to view-invariant representations in monkeys (Logothetis et al., 1994) and the proprioceptive perception of body motion. Visual dynamic properties are nevertheless among the most salient and the earliest used by infants to learn about object structure. At an age when they have a hard time putting together dissimilar objects (Quinn et al., 1993), 2-month-old infants are able to perceive a two-part object as unitary if the two parts move together, either doing a translational movement (Kellman & Spelke, 1983) or rotating (Johnson, Cohen, Marks, & Johnson, 2003). Core knowledge about object cohesion in such moving displays could facilitate the perception of the object unity, including the unity of the different views.

Wang, Tanifuji, & Tanaka (1998), using intrinsic optical imaging of monkeys' infero-temporal cortex (area TE), reported that the activation spot triggered by a doll face moved systematically as the face rotated, delimiting a continuous map of the different views of a face. On the other hand, Keysere, Xiao, Földiak, & Perrett (2001) showed that when perceiving a head in rotation, not only the neurons tuned to the actual head pose fire, but also neurons tuned to the neighboring head view, as if ''predicting'' the next pose that is to be seen. It is possible that the concomitant activation of adjacent neuronal columns, tuned to the different views of a face, when seeing this face in rotation, may allow the association of these different but related poses and the convergence, eventually, to the view-invariant representations observed in adults. It remains to be determined whether this type of learning is possible, in early infancy, for objects other than faces or whether, on the contrary, object familiarity is a key feature.

7. Experiment 5: Does the N290 response suppression in infancy reflect low-level feature processing?

Two processing steps were revealed by the previous three infant experiments. A first step is indexed by the N290 and is sensitive to human eye repetition. A second step, developing in the P400 interval, seems to be triggered by the expectancy of eyes in a related context, especially for a profile context (Experiments 2 and 4).

In adults, our experiments suggest that the N170 component corresponds to the successive access to an eye representation that is exemplar-invariant, then to a face representation that is view-invariant. In infants, the N290 component is not modulated by a profile context. It is thus possible that the N290 suppression is simply triggered by the repetition of low-level structural properties of human eyes (i.e., the strong contrast between the cornea and the iris, Kobayashi & Kohshima, 1997), without access to more general properties of the human face. On the contrary, face representations in infants might already reflect universal properties of human frontal faces and thus a certain degree of abstractness.

In order to oppose these hypotheses we presented to a new group of infants the same eyes pictures amongst front-view faces, which this time had their eyes closed. In this way, while still being presented with frontal faces, the physical similarity between the target eyes and the context faces was diminished. The same non-related context as in the Experiment 1 was used (house pictures).

The rationale behind this new experiment is that if infants' face representations reflect the universal properties of human faces (being an indissociable association of two eyes, a nose, a mouth and a head contour), the missing eyes in the faces with eyes closed should not prevent infants from recuperating all the face relevant information. In consequence, we would expect the eye-evoked N290 response to be suppressed in this experiment as in the first experiment, even in the absence of a direct eye repetition.

At the same time, eye expectance being very high in the context of faces with eyes closed, as it was in the context of profiles, we should also replicate the previous P400 modulation. The P400 component should be larger in the front-view condition, compared to the non-related house condition.

7.1. Methods

7.1.1. Subjects
Twelve infants (mean age 14.7 weeks) participated in this study. The data from 21 additional babies were rejected because of excessive movement or fussiness.

7.1.2. Stimuli
The same 10 front-view faces and houses as in Experiment 1 were used. The eyes were removed from the faces with an image editor. The eyelids were re-created so that the eyes looked naturally closed.

7.1.3. ERP recording and data analysis
We recorded a mean of 152 artifact-free trials for this experiment (51 profile trials, 54 car trials, 23 eyes amongst profiles, and 24 eyes amongst cars trials).

7.2. Results

Eye-evoked ERPs in the front-view context diverged from the eye-evoked ERPs in the house context at the same two time points that were already modulated in the previous experiments. The first difference was recorded in the 200–260 ms interval, as a decrease of the N290 amplitude in the front-view context, especially on the left hemisphere (Fig. 5). No main effect of Condition was observed ($F(1, 11) = 2.9$, $p > 0.1$). The effect of condition was only significant over the left hemisphere ($F(1, 11) = 5.22$, $p = .043$) although the interaction Condition × Hemisphere was not significant ($F(1, 11) < 1$).

A second difference was observed over the 520–580 ms interval. As in Experiment 2, the P400 was enhanced for the eyes presented in the front-view context relatively to the house context, however, only over the right hemisphere (Fig. 5) yielding a
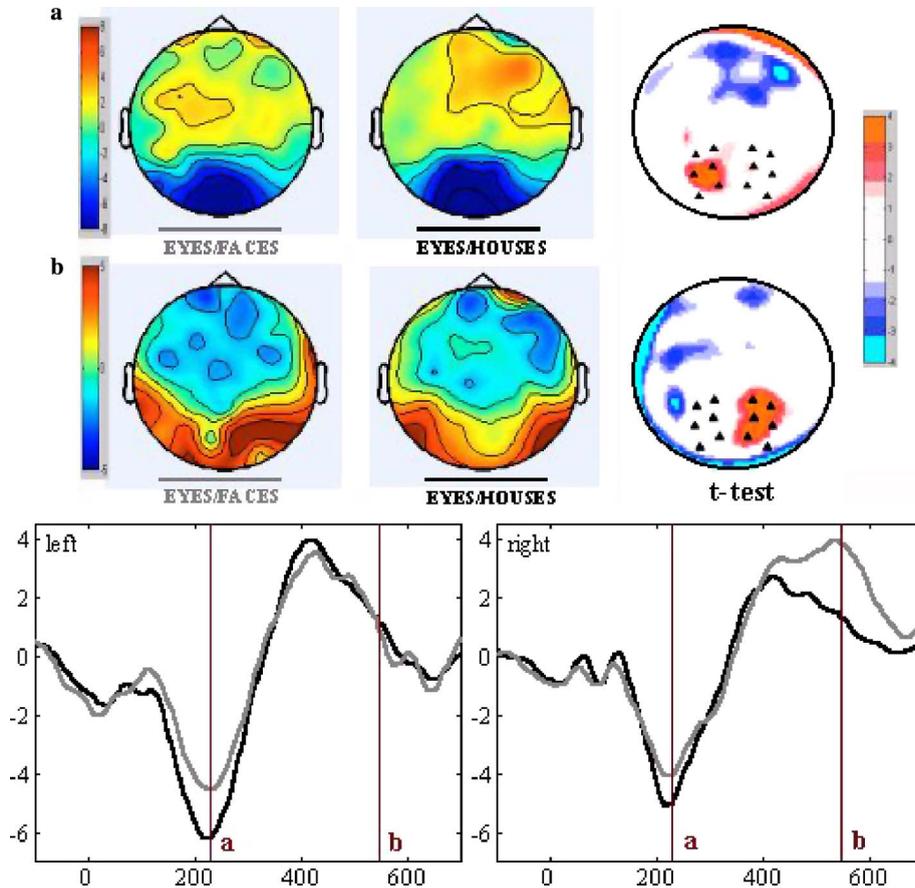
Fig. 5. In infants, eye-evoked N290 is suppressed (a) and the P400 is enhanced (b) in a context made of front-view faces, with the eyes closed. The last row presents the grand-averaged waveforms, summed over the left and right groups of electrodes used in the analysis.

marginal main effect of Condition ($F(1, 11) = 3.52$, $p = .08$) and a significant interaction Condition × Hemisphere ($F(1, 11) = 5.51$, $p = .038$). Post hoc analyses show a significant effect of Condition on the right ($F(1, 11) = 26.32$, $p < .01$) but not the left hemisphere ($F(1, 11) < 1$).

7.3. Discussion

Although weaker, a similar suppression of the eye-evoked N290 was observed when eyes were closed or when eyes were open in the context front-view faces. These results suggest that the repetition of physically similar images is not a necessary condition to induce response suppression in young infants and that by 4 months of age, infants perceive eyes as intrinsic and indissociable components of human faces and are able to fill-in the missing information in an incomplete

face. Similar filling-in of the missing facial information was found in adults. Two meaningless dots are given an eye status if they are presented at eye's place in a schematic face, and evoke subsequently a N170 response (Bentin & Golland, 2002; Bentin, Sagiv, Mecklinger, Friederici, & VonCramon, 2002). However, the electrodes displaying the N290 suppression effect in the present experiment are only a subgroup of the electrodes concerned by this effect in the first experiment, when a direct eye repetition was used. This suggests that a larger group of neurons are sensitive to the structural aspects of eyes than to the general facial properties.

The later effect, the P400 enhancement, confirms our interpretations of this component in terms of expectancies. As in the profile context, there is a strong probability of seeing eyes after having seen a face with the eyes closed. Similar modulation of object-evoked responses presented in a related context was observed in adults, and concern longer latency components (Ganis & Kutas, 2003). Further studies will tell us whether the infant P400 modulation in our studies is specific to the special social status of the eyes (infants look for eyes when scanning a face, Haith et al., 1977) or due to more general knowledge on object associations.

8. Repetition paradigms increase ERP selectivity in infant studies

In the light of the above results, we want to underscore the power of our design, especially for infant studies. By using the same target images in all contexts, we overcame the problem of comparing ERPs to objects having different low-level visual properties (Gliga & Dehaene-Lambertz, 2005). This kind of comparison would be uninformative in the present study: for example there was no significant difference between human and non-human categories if we examined the ERPs to the different context images (Fig. 6), either for the N290 component (human vs. non-human category: $p = 0.1$) or the P400 component ($F(1, 45) < 1$). Post hoc analyses restricted to each experiment showed similar results ($ps > 0.2$) except for profiles vs. cars $F(1, 15) = 6.11$, $p = .02$ at the N290 time window.

While eye-induced face N290 suppression could account for similar N290 response for faces and houses, this interpretation does not hold for the other two contrasts (profiles vs. cars or bodies vs. tools) or for the lack of P400 difference in any of the contrasts. These results could have been interpreted as evidence for undifferentiated representations of the studied objects (e.g., faces and houses). This could mean either that infants treat these categories as equivalent, or that these distinct categories are not yet segregated on the cortical surface (Westermann & Mareschal, 2004). Our adaptation paradigm contradicts the first interpretation as it reveals selective neuronal associations between human corporal contexts and eye picture presentation. Thus, our design overcomes the problem of the low spatial resolution of ERPs, providing a more powerful tool for the study of object representations in infancy.
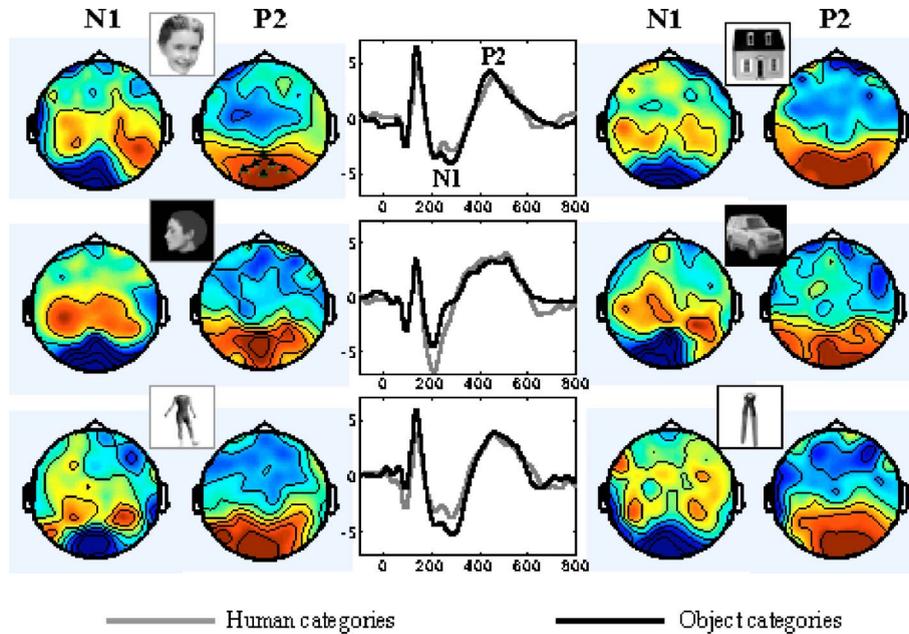
Fig. 6. ERPs to the different image categories in infants. The voltage maps show the topography of the ERPs at the maximum of the N1 and P2 components. The waveforms show the grand average recorded over a group of occipital electrodes (marked by triangles on the voltage maps).

9. General discussion

Different classes of objects benefit from being presented along with an appropriate context (Bar, 2004). Contextual modulation is thought to mediate very fast object recognition in natural scenes (Lewis & Edmonds, 2002; Thorpe et al., 1996). Such contextual modulation was observed for face perception as well. Bentin & Golland (2002) & Bentin et al. (2002) show that a drawing that has no structural similarity with faces can evoke a N170 if presented in a face context (two meaningless points presented as eyes in a face will elicit a N170). In the present study, we extend this finding by showing that both in adulthood and infancy, eye processing is modulated by a human context, even in the absence of any physical similarity between the context and the target images (profiles with eyes closed, and bodies without heads). However, similarities and differences were present between the two studied ages, adults and 3- to 4-month-old infants.

The rationale behind our experimental design was that activity decreases in a network coding a representation (Desimone, 1996) whenever this representation is repeatedly accessed, giving us cues about the common property detected between context and target images. Eye representations (exemplar-invariant) and head representations (view-invariant) were targeted in adults and young infants (3–4 months old). Our main finding is that although infants expect eye presence in a profile con-

text, they need more than 4 months to build the view-invariant representations that we observed in adults.

In adults, both a frontal view and a profile context suppressed the eye-evoked N170. The N170 is associated with face, and face-part structural processing, being modulated by the general face schema (Bentin et al., 1996; Bentin & Deouell, 2000; Gliga, 2003) or orientation (Stekelenburg & deGelder, 2004), or simply by their detection. Face familiarity (Bentin & Deouell, 2000) or emotional expression (Stekelenburg & deGelder, 2004) do not modulate this component. Thus, the N170 is not a unitary component; it reflects multiple processing steps taking place in parallel or successively, but in a short time interval. Our paradigm allows a fine "dissection" of this component. We showed that two types of representations are built successively, an eye representation or frontal-view representation, accessed around 180 ms and a view-invariant face template, 30 ms later. We can imagine further experiments that will better characterize the face-processing steps, separating local and global processing for example or category-general and exemplar-specific processing.

In infants, the two visual-evoked responses, the N290 and the P400, were differently modulated by the human contexts. The eye-evoked N290 is diminished by a frontal-face context, which could contain eyes or not, while the P400 is enhanced by the physically dissimilar but related profile or frontal face with eyes closed contexts. These different results cannot be due to different low-level properties of the non-human categories (houses in Experiment 1 and 5 and cars in Experiment 2 and 4). The N290 to eye picture presentation was similar across the non-human contexts ($ps > 0.2$), suggesting that these contexts were, as expected, neutral with respect to eye processing.

A number of infant studies have proposed that the N290 is related to eye processing. Because it is modulated when eyes with directed and averted gaze (Farroni et al., 2002) or human and monkey eyes (de Haan et al., 2002) are contrasted. However, the presence of open eyes was not necessary to trigger eye N290 suppression in our study. Thus, the N290 does not reflect only low-level feature processing (the strong iris-cornea contrast in the human eye) but more general properties of the human face.

The P400 is modulated when no direct perceptual cue links the context and the target images and is subject to fast learning, as demonstrated by the fourth infant experiment. Such fast perceptual learning and generalization capacities have already been observed in the infant literature. Two minutes of exposure to four different faces are sufficient for 3-month-old babies to build an average face (a prototype) (de Haan, Johnson, Maurer, & Perett, 2001). In another study, Johnson, Amso, & Slemmer (2004) show that 4-month-old infants' capacity to anticipate an object's appearance from behind a screen is improved after 2 min exposure to this object in motion. The authors conclude that an associative mechanism, needing only limited exposure, is at work when learning object properties in infancy.

The similarity between the infant and adult context effects latency in some of the present experiments (the adult N170 and the infant N290 are modulated in the frontal-view context) in opposition to the differences observed in other experiments (the adult N170 but the infant P400 are modulated by the profile context), can also be dis-

cussed from the point of view of the equivalence between the infant and adult visual-evoked potentials. The question of whether only the N290 or both the N290 and the P400 are precursors of the adult N170 was repeatedly raised (de Haan et al., 2003; Halit et al., 2003; Taylor et al., 2001). ERPs do not give a direct access to the underlining neuronal substrate. Nevertheless, when taking into account multiple ERP parameters, like the topography of the components, their latency, and their modulation by certain experimental conditions, it is possible to make inferences about the equivalence of the underlying generators. Based on similar topography and sign – both are posterior, negative-going deflections – and on similar modulation by the face species (human vs. monkey) (de Haan et al., 2002), the N290 had been designated as the "infant N170". The results of our own experiments, showing response suppression of the N170 and of the N290 by the front-view contexts, seem to confirm this hypothesis. On the other hand, while the face orientation affects the N170 in adults it modulates the P400 at 3 and 6 months of age and both the N290 and the P400 at 12 months of age (de Haan et al., 2003). The authors of this study interpreted these results as a gradual development of both N290 and P400 components into the adult N170, due to anatomical and processing speed modifications. Indeed, gradual changes in latency and form of the N170 component are observed along childhood (Taylor et al., 2001). A similar shift is observed in our study, an adult N170 vs an infant P400 modulation of the eye-evoked responses, in a profile context. Contrary to the de Haan et al. (2003) study, in our study opposite sign effects are found at these latencies, suggesting that we are not facing only a modification in the speed of processing. The nature of the underlining mechanisms is probably different, as previously discussed. Thus, a functional approach of the ERP component properties suggests that the question of equivalence between the infant N290/P400 complex and the adult N170 cannot be given a simple yes or no answer. Some of the underlining visual processes, like eye processing or face orientation discrimination, suffer probably only anatomical changes (speed of processing, generator orientation) and we could therefore consider them as "equivalent" between the two populations. Other mechanisms change in nature, as is the case of view-invariant face perception.

In conclusion, the present study allowed us to characterize two learning stages. At 4 months, infants build a unitary percept of the human face, based on the frequent experience with co-occurring face elements. By the same age they have also noticed the temporal succession of profile and front faces and expect that after seeing a profile, they will be able to make an eye contact. Later on, the frequent visual exposure to this salient 3D object in rotation will lead to the construction of a view-invariant representation of the human head. Further studies should inquire whether this learning is mediated by general associative mechanisms or whether it uses specific knowledge about 3D object geometry and their physical appearance in rotation.

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.

Bar, M. (2003). A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neurosciences, 15*(4), 600–609.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience, 5*, 617–629.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition, 14*(2), 201–212.

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate module? Evidence from human neonates. *Infant Behavior and Development, 23*, 223–229.

Behl-Chadha, G. (1996). Basic-level and superordinate-like categorical representations in early infancy. *Cognition, 60*, 105–141.

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., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: The influence of perceptual experience on early visual processing of faces. *Cognition, 86*, B1–B14.

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

Biederman, I. (1987). Recognition-by-components: A theory of human image interpretation. *Psychological Review, 94*, 115–147.

Blass, E. M., & Camp, C. A. (2003). Biological bases for face preference in 6-week-old infants. *Developmental Science, 6*(5), 524–536.

Bonatti, L. E., Frot, R., Zangl, J., & Mehler, J. (2002). The Human First Hypothesis: Identification of conspecifics and individuation of objects in the young infant. *Cognitive Psychology, 44*, 388–426.

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.

Cave, B. C. (1997). Very long-lasting priming in picture naming. *Psychological Science, 8*, 322–325.

de Haan, M., Johnson, M. H., & Halit, H. (2003). Development of face-sensitive event-related potentials during infancy: A review. *International Journal of Psychophysiology*.

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., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience, 14*(2), 199–209.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*.

Dehaene-Lambertz, G., & Dehaene, S. (2004). Speed and cerebral correlates of syllable discrimination in infants. *Nature, 370*, 292–295.

Dehaene-Lambertz, G., & Gliga, T. (2004). Common neural bases for phoneme processing in adult and infant brain. *Journal of Cognitive Neuroscience, 16*(8), 1375–1387.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences of the United States of America, 93*, 13496–13499.

Desimone, R., Albright, T. D., Gross, C., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of Neuroscience, 4*(8), 2051–2062.

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.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13*(2), 1–10.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America, 99*(14), 9602–9605.

Farroni, T., Johnson, M., & Csibra, G. (2004). Mechanism of eye gaze perception during infancy. *Journal of Cognitive Neuroscience, 16*(8), 1320–1326.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 15822–15836.

Ganis, G., & Kutas, M. (2003). An electrophysiological study of scene effects on object identification. *Cognitive Brain Research, 16*, 123–144.

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. (2003). La reconnaissance des visages par le nourisson. *Medecine and enfance, 23*(9), 553–564.

Gliga, T., & Dehaene-Lambertz, G. (2005). Structural encoding of human body and face in infants and adults. *Journal of Cognitive Neuroscience, 17*(8).

Grill-Spector, K., Kushnir, S. E., Avidan, G., Itzchak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron, 24*, 187–203.

Grill-Spector, K., Kushnir, S. E., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping, 6*, 316–328.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica, 107*, 293–321.

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science, 198*, 853–855.

Halit, H., Csibra, G., Volein, A., & Johnson, M. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry, 45*(7), 1228–1234.

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*(3), 1180–1193.

Henson, R. N. A., & Rugg, M. D. (2003). Neural response supression, haemodynamic repetition effects and behavioural priming. *Neuropsychologia, 41*, 263–270.

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

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

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research, 78*(1), 193–202.

Jeffreys, D. A., & Tukmachi, E. S. (1992). The vetex-positive scalp potential evoked by faces and by objects. *Experimental Brain Research, 91*(2), 340–350.

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition, 40*(1-2), 1–19.

Johnson, S., Amso, D., & Slemmer, J. A. (2004). Development of object concept in infancy: Evidence for early learning in an eye-tracking paradigm. *Proceedings of the National Academy of Sciences of the United States of America, 100*(18), 10568–10573.

Johnson, S. P., Cohen, L. B., Marks, K. H., & Johnson, K. L. (2003). Young infant's perception of objects unity in rotation displays. *Infancy, 4*(2), 285–295.

Kellman, P. J., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology, 15*, 483–524.

Keysere, C., Xiao, D. K., Földiak, P., & Perrett, D. (2001). The speed of sight. *Journal of Cognitive Neuroscience, 13*(1), 90–101.

Kiefer, M. (2001). Perceptual and semantic sources of category-specific effects: Event-related potentials during picture and word categorization. *Memory & Cognition, 29*(1), 100–116.

Kobayashi, H., & Kohshima, S. (1997). Unique morphology of the human eye. *Nature Neuroscience, 387*, 767–768.

Kourtzi, Z., & Shiffrar, M. (1999). The visual representation of three-dimensional, rotating objects. *Acta Psychologica*, 265–292.

Lewis, M. B., & Edmonds, A. J. (2002). Localisation and detection of faces in naturalistic scenes. *Perception, 31*, 19.

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

Luck, S. J., & Hillyard, S. A. (1999). The operation of selective attention at multiple stages of processing. Evidence from human and monkey electrophysiology. In M. Gazzaniga (Ed.), *The new cognitive neurosciences*. Cambridge, MA: MIT Press.

Lueschow, A., Miller, E. K., & Desimone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex, 5*, 523–531.

Maurer, D., & Barrera, M. (1981). Infants' perception of natural and distorted arrangements of a schematic face. *Child Development, 52*(1), 196–202.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by neonates. *Science, 198*, 75–78.

Meltzoff, A. N., & Moore, M. K. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology, 25*(6), 954–962.

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

Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex, 11*, 966–974.

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

Pallier, C., Dupoux, E., & Jeannin, X. (1997). Expe: An expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, and Computers, 29*, 322–327.

Pascalis, O., de Haan, M., Nelson, C. A., & de Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 249–260.

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology, 93*, 603–608.

Perrett, D. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London, Series B, 223*, 293–317.

Pourtois, G., Schwartz, S., Seghier, M. L., Lazeyras, F., & Vuilleumier, P. (2005). View-independent coding of face identity in frontal and temporal cortices is modulated by familiarity: An event-related fMRI study. *Neuroimage, 24*(4), 1214–1224.

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception, 22*, 463–475.

Rousselet, G. A., Mace, M. J. M., & Fabre-Thorpe, M. (2003). Is it an animal? Is it a human face? Fast processing in upright and inverted natural scenes. *Journal of Vision, 3*(440-455).

Sitnikova, T., Kuperberg, G., & Holocomb, P. J. (2003). Semantic integration in videos of real-world events: An electrophysiological investigation. *Psychophysiology, 40*, 160–164.

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.

Sugase, Y., Yamane, S., Ueno, S., & Kawano, K. (1999). Global and fine information coded by single neurons in the temporal visual cortex. *Nature, 400*, 869–873.

Tanaka, K. (2000). Mechanisms of visual object recognition studied in monkeys. *Spatial Vision, 13*(2,3), 147–163.

Tanifuji, M., Tsunoda, K., & Yamane, S. (2002). Neural representation of object images in the macaque inferotemporal cortex. In N. Kanwisher & J. Duncan (Eds.), *Functional neuroimaging of visual cognition. Attention and performance XX*. Oxford: Oxford University Press.

Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T. (2001). Eyes first! Eye processing develops before face processing in children. *Neuroreport, 12*(8), 1671–1676.

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

Wang, G., Tanifuji, M., & Tanaka, K. (1998). Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging. *Neuroscience Research, 32*, 33–46.

Westermann, G., & Mareschal, D. (2004). From parts to wholes: Mechanisms of development in infant visual object processing. *Infancy, 5*(2), 131–151.

Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science, 256*, 1327–1331.