Early Face Processing Specificity: It’s in the Eyes!

Roxane J. Itier1, Claude Alain1,2, Katherine Sedore1, and Anthony R. McIntosh1,2

Abstract

Unlike most other objects that are processed analytically, faces are processed configurally. This configural processing is reflected early in visual processing following face inversion and contrast reversal, as an increase in the N170 amplitude, a scalp-recorded event-related potential. Here, we show that these face-specific effects are mediated by the eye region. That is, they occurred only when the eyes were present, but not when eyes were removed from the face. The N170 recorded to inverted and negative faces likely reflects the processing of the eyes. We propose a neural model of face processing in which face- and eye-selective neurons situated in the superior temporal sulcus region of the human brain respond differently to the face configuration and to the eyes depending on the face context. This dynamic response modulation accounts for the N170 variations reported in the literature. The eyes may be central to what makes faces so special.

INTRODUCTION

The debate surrounding face specificity is still on. For more than three decades, behavioral studies have used the so-called inversion effect as a measure of specific mechanisms for processing faces compared to other visual object categories. Faces are harder to perceive, memorize, and recognize when presented upside-down compared to right-side up, and this effect is disproportionately larger than for common objects (Yin, 1969). Most researchers agree that inversion disrupts mainly the first-order relations between facial features common to all faces (eyes above a nose, itself above the mouth) and the second-order feature relations that make each face unique, such as, the distance between the eyes (Maurer, Le Grand, & Mondloch, 2002). These two feature relations are crucial for normal upright face recognition and belong to what is called configural processing, as opposed to featural processing related to information about the individual features, for instance, the shape of the nose (Maurer et al., 2002). Upright faces are thus thought to be processed mainly configurally, whereas inverted faces and objects are processed analytically, in a feature-based manner (Maurer et al., 2002; Rossion & Gauthier, 2002). This configural processing is regarded as a major aspect of face-specific processing (Maurer et al., 2002; Rossion & Gauthier, 2002; Rhodes, Brake, & Atkinson, 1993; Tanaka & Farah, 1993). Although less studied than the inversion effect, the contrast reversal (CR) effect (photo-negative) is another manipulation which keeps the first-order feature relations intact but also impairs face perception and recognition much more so than object recognition (Galper, 1970; Hochberg & Galper, 1967) by disrupting the second-order feature relations (Kemp, McManus, & Pigott, 1990). As inversion and CR result in very different stimuli yet both disrupt one same aspect of configural processing, using them together provides additional information on face processing.

Although over the past years the key areas of the brain networks subtending the processing of various aspects of the face have been found (Allison, Puce, & McCarthy, 2000; Haxby, Hoffman, & Gobbini, 2000), the neuronal dynamics of the face inversion and CR effects remain unclear. Studies recording electrical activity from the scalp have found an early face-sensitive marker, the negative N170 event-related potential (ERP) that occurs between 140 and 200 msec after stimulus onset (Bentin, Allison, Puce, Perez, & McCarthy, 1996) and is largest over parieto-occipital sites. The N170 is larger for faces than objects (Itier & Taylor, 2004b; Carmel & Bentin, 2002; Eimer, 2000a; Bentin et al., 1996), whether the stimulus is a schematic face, a sketch face, paintings, drawings or photographs of faces (Sagiv & Bentin, 2001), two-tone Mooney figures perceived as faces (George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) and is largest over parieto-occipital sites. The N170 is larger for faces than objects (Itier & Taylor, 2004b; Carmel & Bentin, 2002; Eimer, 2000a; Bentin et al., 1996), whether the stimulus is a schematic face, a sketch face, paintings, drawings or photographs of faces (Sagiv & Bentin, 2001), two-tone Mooney figures perceived as faces (George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) and is largest over parieto-occipital sites. The N170 is larger for faces than objects (Itier & Taylor, 2004b; Carmel & Bentin, 2002; Eimer, 2000a; Bentin et al., 1996), whether the stimulus is a schematic face, a sketch face, paintings, drawings or photographs of faces (Sagiv & Bentin, 2001), two-tone Mooney figures perceived as faces (George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) and is largest over parieto-occipital sites. The N170 is larger for faces than objects (Itier & Taylor, 2004b; Carmel & Bentin, 2002; Eimer, 2000a; Bentin et al., 1996), whether the stimulus is a schematic face, a sketch face, paintings, drawings or photographs of faces (Sagiv & Bentin, 2001), two-tone Mooney figures perceived as faces (George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) and is largest over parieto-occipital sites. The N170 is larger for faces than objects (Itier & Taylor, 2004b; Carmel & Bentin, 2002; Eimer, 2000a; Bentin et al., 1996), whether the stimulus is a schematic face, a sketch face, paintings, drawings or photographs of faces (Sagiv & Bentin, 2001), two-tone Mooney figures perceived as faces (George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) and is largest over parieto-occipital sites. The N170 is larger for faces than objects (Itier & Taylor, 2004b; Carmel & Bentin, 2002; Eimer, 2000a; Bentin et al., 1996), whether the stimulus is a schematic face, a sketch face, paintings, drawings or photographs of faces (Sagiv & Bentin, 2001), two-tone Mooney figures perceived as faces (George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) and is largest over parieto-occipital sites. The N170 is larger for faces than objects (Itier & Taylor, 2004b; Carmel & Bentin, 2002; Eimer, 2000a; Bentin et al., 1996), whether the stimulus is a schematic face, a sketch face, paintings, drawings or photographs of faces (Sagiv & Bentin, 2001), two-tone Mooney figures perceived as faces (George, Jemel, Fiori, Chaby, & Renault, 2005; Latinus & Taylor, 2005) and is largest over parieto-occipital sites.
suggests generators located in the inferior occipito-temporal cortex and/or superior temporal sulcus (STS). The N170 is thought to reflect the encoding stage of the face structure (Eimer, 2000a; Rossion, Campanella, et al., 1999), when the perceptual representation of a face is being created.

There is now considerable evidence that the N170 latency is delayed and its amplitude increased by face inversion (Itier, Latinus, et al., 2006; Itier & Taylor, 2002, 2004b, 2004c; de Haan et al., 2002; Taylor, Edmonds, McCrory, & Allison, 2001; Eimer, 2000b; Rossion et al., 2000; Bentin et al., 1996) and CR (Itier, Latinus, et al., 2006; Itier & Taylor, 2002, 2004c), an effect not seen for other object categories (Itier, Latinus, et al., 2006; Bentin et al., 1996), which show only a small latency delay and/or a small amplitude reduction (Itier, Latinus, et al., 2006) with these manipulations. These face inversion and CR effects on the N170 are markers of early face-specific processing (Itier, Latinus, et al., 2006; Rossion et al., 2000; Bentin et al., 1996; Itier & Taylor, 2002, 2004c) and may reflect the recruitment of object-selective neurons to process inverted and negative faces, due to the disruption of the face configuration (Itier & Taylor, 2002, 2004c; Rossion & Gauthier, 2002; Itier et al., 2006; Rossion et al., 1999). However, this larger N170 amplitude for inverted and negative compared to upright faces is at odds with other studies. In fact, data from several experiments and imaging modalities find that inverted and negative faces show a similar or smaller response than upright faces. These include monkey face-selective, single-unit recordings (Ito, Fujita, Tamura, & Tanaka, 1994; Hasselmo et al., 1989), human functional magnetic resonance imaging (fMRI) studies of the face-sensitive fusiform gyrus (FG) area (Yovel & Kanwisher, 2005; Aguirre, Singh, & D’Esposito, 1999; George et al., 1999; Haxby et al., 1999; Kanwisher, Tong, & Nakayama, 1998) and STS region (Leube et al., 2003), magneto-encephalography (MEG) recordings of the face-sensitive M170 (Itier, Herdman, George, Cheyne, & Taylor, 2006; Watanabe et al., 2003), and human intracranial recordings of the face-specific N200 from the FG surface (McCarthy, Puce, Belger, & Allison, 1999). Furthermore, source modeling studies of the N170 failed to find different sources for inverted, negative, and upright face N170s (Itier & Taylor, 2002, 2004a; Watanabe et al., 2003). According to all these data, the N170 should be similar or slightly smaller in amplitude for inverted and negative compared to upright faces, rather than larger.

In the present study, we develop a neuronal model for face processing that takes into account monkey cell recording data and explains much of the reported results on human face processing in the literature. This model, described in Figure 1, suggests a new explanation for the increase in N170 amplitude with face inversion and CR.

---

Figure 1. Simplified neural model of early face processing. Three sources are simultaneously active around 170 msec poststimulus onset. One source in the superior temporal sulcus (STS) region with a radial orientation generates the ERP N170 component. The combination of tangential sources in the fusiform gyrus (FG) and middle occipital gyrus (MOG) generates the MEG M170. The dynamic response modulation of eye- and face-selective neurons within the STS accounts for inversion and CR effects on the face N170 amplitude and for the other existing ERP data on the N170. The + signs represent the amount of activation of the neurons. The absence of + signs signifies that the neurons are not responding.
We assume that the N170 originates mainly from the STS region and that both face- and eye-selective neurons coexist in this part of the brain. According to single-unit recordings, when faces are presented upright, eye-selective neurons do not respond to the eyes of the face because of the face context (Perrett et al., 1985; Perrett, Rolls, & Caan, 1982). Mainly face-selective neurons respond to faces and the N170 would receive contribution primarily from these face-selective cells. However, in inverted and negative face conditions, the face configuration is disrupted (Maurer et al., 2002; Rossion & Gauthier, 2002; Kemp et al., 1990). The eyes are no longer perceived in a normal face context so eye-selective neurons now respond to the inverted or negative eyes of the face, in addition to face-selective neurons. Thus, rather than invoking the recruitment of additional object-selective neurons for the processing of inverted and negative faces, our model suggests that the increase in N170 amplitude observed with facial inversion and CR reflects mainly the processing of the eye region.

Accordingly, upside-down or photo-negative faces-without-eyes should yield N170s of similar amplitudes as normal upright faces because mainly face-selective neurons respond (Figure 1). We directly tested this prediction by removing the eye region from faces that were presented upright, inverted, and negative to a group of 13 participants. In two consecutive tasks, subjects discriminated either the orientation between upright and inverted stimuli (orientation discrimination task), or the contrast between positive and negative stimuli (contrast discrimination task), as previously done (Itier, Latinus, et al., 2006; Rossion et al., 2000), in blocks where faces, faces-without-eyes, isolated eyes and houses (Figure 2) were randomly mixed. Houses were used as the object category as they present a canonical orientation like faces and are commonly used as a control category. If processing the eye region is indeed responsible for the inversion and CR effects on the N170, then removing the eyes should prevent the N170 increase with inversion and CR. We thus predicted no amplitude increase between upright and inverted or negative faces-without-eyes. Furthermore, if processing inverted and negative faces reflects mainly the processing of the eye region, we should find no difference between the N170 to eyes and to inverted and negative faces, as well as no difference between their associated topographies, as topographies reflect the underlying generators of the signal recorded on the scalp.

**METHODS**

**Subjects**

Thirteen young and healthy volunteers from the Toronto area (21–35 years) were tested and paid for their participation. The data of one subject were discarded for the contrast categorization task due to acquisition problems. All participants had normal or corrected-to-normal vision. All signed a written informed consent that was approved by the Ethics Research Board of Baycrest.

**Stimuli and Design**

Gray-scale photographs of 48 different faces (half female) and 48 different houses were used. Photoshop 6.0 was used to create all final stimuli. The upper and lower parts of the faces were cropped so that for all 48 individuals, the eye region within the face was exactly in the center of the stimulus (12.3° × 6.4° visual angles). Faces were of neutral expression, showing all internal features and part of the hair (Figure 2). A set of 48 isolated eye stimuli was created by cropping the eye region from the faces in such a way that for all stimuli, the center of the rectangular eye region was approximately the center of the eyes themselves. The eye stimuli included the eyes, eyebrows, and the region between the eyes. Another set of 48 faces-without-eyes stimuli were created from the full face set by erasing the eye region and filling it with skin-like texture. Inverted and negative (contrast-reversed) stimuli were created from the original upright faces.
positive pictures. A simple 180° rotation was used to create the inverted stimuli, and the negative sets were created by inverting contrast information in the pictures without changing the upright orientation. All pictures were presented centrally on a gray background. The eye region was centered in all stimuli (including inverted items) and on the monitor.

Subjects sat in a dimly lit and sound-proof booth 60 cm in front of a computer monitor. They performed two tasks successively using two buttons of a keyboard. In the orientation judgment task, upright and inverted pictures were presented and subjects discriminated the orientation of each stimulus regardless of its category. In the contrast judgment task, positive contrast stimuli (the same as the upright stimuli in the orientation task) were discriminated from negative contrast pictures regardless of category. In each task, 12 blocks of 16 items (4 different pictures per category) were presented twice, yielding 96 trials per category per task. Stimuli were on for 250 msec with a 1000–1300 msec randomized interstimulus interval during which a central fixation cross appeared. Subjects had to respond as quickly and accurately as possible. Button keys were switched halfway through each task. The order of buttons, tasks, blocks, and stimuli within block was randomized and counterbalanced across subjects.

Electrophysiological Recordings

The electroencephalogram (EEG) was recorded with a NeuroScan 4.0 using 64-channel ElectroCaps (10/20 system) including two pairs of ocular sites monitoring vertical and horizontal eye movements from the outer canthi and infraorbital ridges. Electrode impedances were kept under 5 kΩ. The continuous EEG signal (500 Hz sampling rate, band-pass 0.01–100 Hz) was amplified using two SynAmps amplifiers. Trials containing artifacts (>±70 μV) on electrodes not adjacent to the eyes were rejected automatically after baseline correction. For each task, accepted trials were averaged into 600 msec sweeps (−200 to 400 msec) according to the four categories and the two formats (upright/inverted and positive/negative, respectively). Ocular artifacts were removed on each average by means of ocular source components with Brain Electrical Source Analysis (BESA) software (Picton et al., 2000). Averages were then digitally low-pass filtered (30 Hz).

Data Analysis

The N170 mean amplitude was measured using an automated procedure, at posterior sites P7/P8 and CB1/CB2 where it was maximal, using a ±40 msec window around the maximum of the grand-average mean for each condition. Peak latencies and amplitudes were taken at the maximum negativity between 130 and 210 msec after stimulus onset. They were measured at each electrode and were analyzed separately.

Repeated-measures analyses of variance (ANOVAs) were conducted separately for each task. Within-subject factors included category (4: eyes, face, faces-without-eyes, house) and orientation/contrast (2: upright/inverted or positive/negative) for accuracy and reaction time measures. Hemisphere (2) and electrode (2: parietal or cerebellar) were additional factors for peak latencies and amplitudes. Category effects on the N170 component were analyzed separately for upright, inverted, positive, and negative categories (Table 2), using a 4 (category) × 2 (hemisphere) × 2 (electrode) ANOVA. A 2 (orientation or contrast) × 2 (hemisphere) × 2 (electrode) ANOVA was also performed for each category separately (Table 3) to test for inversion and CR effects. Normalized mean amplitudes were calculated (McCarthy & Wood, 1985) between 120 and 220 msec and were compared between categories to assess topography differences using ANOVAs with category (2 levels) and electrodes (61—no ocular sites) as within-subject factors. Comparing topographies were particularly important to test our model predictions that postulate different contributions from face- and eye-selective neurons to the N170 recorded for the various stimulus categories. All ANOVAs used Greenhouse–Geisser adjusted degrees of freedom and pairwise comparisons used Bonferroni corrections for multiple comparisons. Planned comparisons used paired-sample t tests (two-tailed).

RESULTS

The behavioral results are summarized in Table 1. According to our model, one would predict interactions between orientation and categories, and contrast and categories, as inversion and CR should be maximal for faces. That is, for these manipulations, faces-without-eyes should show behavioral profiles similar to other objects. For accuracy, no interactions with category were found. However, reaction times revealed that the inversion effect and the CR effect were most pronounced for faces [F(3, 36) = 5.17, p = .011 for orientation; F(3, 36) = 5.12, p < .01 for contrast]. Although planned comparisons showed an inversion effect for faces-without-eyes [t(12) = −3.44, p < .005], but not for houses (p = .56) or isolated eyes which presented a trend (p = .062), the effect was largest for faces [t(12) = −5.38, p < .0001]. In the contrast task, slower RTs with CR were found for faces only [t(12) = −2.62, p < .05]; only a trend was found for houses [t(12) = −1.85, p < .09] and no effects were found for eyes (p = .18) or faces-without-eyes (p = .37).

Statistics for the ERP data are reported in Tables 2 and 3. In both tasks, the N170 was shortest for faces, paralleling the fastest reaction times found and reflecting faster processing of human faces (Itier, Latinus, et al.,...
As is often found, the N170 amplitude was larger to faces than houses and smaller to faces than isolated eyes. Inversion and CR for all categories are presented in Figures 3 and 4, with their associated topographies in Figure 5. As expected, inversion and CR increased N170 latency for all categories but increased its amplitude for faces only. Our prediction was thus confirmed: The N170 amplitude increased only

Table 1. Mean Accuracy Rates (%) and Reaction Times (msec) for Each Category, Both Orientations and Contrasts, for the Two Tasks

<table>
<thead>
<tr>
<th></th>
<th>Eyes</th>
<th>Faces</th>
<th>Faces-No-Eyes</th>
<th>Houses</th>
<th>Eyes</th>
<th>Faces</th>
<th>Faces-No-Eyes</th>
<th>Houses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hits (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Orientation Discrimination Task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upright</td>
<td>93 (1.4)</td>
<td>96.8 (0.7)</td>
<td>94.6 (0.9)</td>
<td>92.9 (1.7)</td>
<td>531 (26)</td>
<td>503 (28)</td>
<td>517 (27)</td>
<td>541 (26)</td>
</tr>
<tr>
<td>Inverted</td>
<td>93.4 (1.7)</td>
<td>93.3 (1.9)</td>
<td>93.3 (1.4)</td>
<td>93.8 (1.5)</td>
<td>542 (26)</td>
<td>523 (28)</td>
<td>533 (30)</td>
<td>538 (28)</td>
</tr>
<tr>
<td>Inverted–upright</td>
<td>0.4 (0.3)</td>
<td>−3.5 (1.2)</td>
<td>−1.3 (0.5)</td>
<td>0.9 (0.2)</td>
<td>11 (1)</td>
<td>20 (1)</td>
<td>16 (3)</td>
<td>−3 (2)</td>
</tr>
<tr>
<td><strong>Contrast Discrimination Task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>94.1 (1.5)</td>
<td>97 (0.8)</td>
<td>91.7 (1.9)</td>
<td>90.7 (1.6)</td>
<td>525 (24)</td>
<td>486 (26)</td>
<td>529 (25)</td>
<td>546 (24)</td>
</tr>
<tr>
<td>Negative</td>
<td>94.2 (1.1)</td>
<td>94.2 (1.3)</td>
<td>94.2 (1.3)</td>
<td>88.6 (2.6)</td>
<td>516 (25)</td>
<td>500 (23)</td>
<td>520 (29)</td>
<td>561 (27)</td>
</tr>
<tr>
<td>Negative–upright</td>
<td>0.1 (0.4)</td>
<td>−2.8 (0.5)</td>
<td>2.5 (0.6)</td>
<td>−2.1 (1)</td>
<td>−9 (1)</td>
<td>14 (3)</td>
<td>−9 (4)</td>
<td>15 (3)</td>
</tr>
</tbody>
</table>

Standard errors reported in parentheses.

In upright and positive contrasts, faces were responded to faster than all other categories. Most importantly, inversion and contrast-reversal effects on reaction times were largest for faces.

Table 2. Main Effects of Category on the N170 Component Measured at P7/8 and CB1/2

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>p</th>
<th>Pairwise Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orientation Discrimination Task</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upright</td>
<td>N170 latency</td>
<td>34.27</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>N170 amplitude</td>
<td>41.54</td>
<td>.0001</td>
</tr>
<tr>
<td>Inverted</td>
<td>N170 latency</td>
<td>7.49</td>
<td>.014</td>
</tr>
<tr>
<td></td>
<td>N170 amplitude</td>
<td>68.46</td>
<td>.0001</td>
</tr>
<tr>
<td><strong>Contrast Discrimination Task</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>N170 latency</td>
<td>16.45</td>
<td>.0001</td>
</tr>
<tr>
<td></td>
<td>N170 amplitude</td>
<td>53.75</td>
<td>.0001</td>
</tr>
<tr>
<td>Negative</td>
<td>N170 latency</td>
<td>14.74</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>N170 amplitude</td>
<td>47.33</td>
<td>.0001</td>
</tr>
</tbody>
</table>

Significant pairwise comparisons are reported at the various possible p values. Upright faces elicited smaller N170s than eyes, but similar than upright faces-without-eyes. N170s for inverted and negative faces were not significantly different from N170s to inverted or negative eyes, respectively.

* p < .05.
** p < .01.
*** p < .005.
† p < .0001.
slightly with inversion for faces-without-eyes and was not significantly modulated for negative faces-without-eyes (Figures 3 and 4, Table 3). The amplitude of the N170 to inverted faces-without-eyes was significantly smaller than that to inverted faces (Table 2), illustrating that the removal of the eyes greatly decreased the magnitude of the inversion effect on the N170. As seen for the N170 amplitudes, the topographies between faces and faces-
without-eyes presented upright did not differ significantly \([\text{Category} \times \text{Electrode interaction}, F(60, 720) < 2, p = .2; \text{Figure } 5]\), suggesting that upright faces-without-eyes are processed like normal faces. In contrast, along with the N170 amplitude modulations, topographies between faces and faces-without-eyes differed when the stimuli were inverted \([F(60, 660) = 5.2, p < .007; \text{Figure } 5]\), suggesting that

**Table 3.** Summary of the Effects of Inversion and Contrast Reversal on the N170 Component (Measured at P7/8 and CB1/2) for Each Category Analyzed Separately, in the Orientation and Contrast Discrimination Tasks, Respectively

<table>
<thead>
<tr>
<th></th>
<th>Eyes</th>
<th>Faces</th>
<th>Face-No-Eyes</th>
<th>Houses</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Orientation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Discrimination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Task (n = 13)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N170 latency</td>
<td>(F = 47.15^1), Inv &gt; up</td>
<td>(F = 415.18^1), Inv &gt; up</td>
<td>(F = 38.84^1), Inv &gt; up</td>
<td>(F &lt; 2, \text{ns})</td>
</tr>
<tr>
<td>N170 amplitude</td>
<td>(F &lt; 2, \text{ns})</td>
<td>(F = 35.84^1), Inv &gt; up</td>
<td>(F = 7.9^*, \text{inv &gt; up})</td>
<td>(F &lt; 1, \text{ns})</td>
</tr>
<tr>
<td><strong>Contrast</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Discrimination</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Task (n = 12)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N170 latency</td>
<td>(F = 20.16, \text{neg &gt; pos}^**)</td>
<td>(F = 87.6^1, \text{neg &gt; pos})</td>
<td>(F = 81.61^1, \text{neg &gt; pos})</td>
<td>(F &lt; 1, \text{ns})</td>
</tr>
<tr>
<td>N170 amplitude</td>
<td>(F = 30.93^1, \text{pos &gt; neg})</td>
<td>(F = 43.16^1, \text{neg &gt; pos})</td>
<td>(F = 3.98, \text{ns})</td>
<td>(F &lt; 1, \text{ns})</td>
</tr>
</tbody>
</table>

\(\text{inv} = \text{inverted}; \text{up} = \text{upright}; \text{neg} = \text{negative}; \text{pos} = \text{positive}. \text{No N170 amplitude increase was found for faces-without-eyes, unlike what was found for faces. Inversion increased N170 amplitude for faces-without-eyes but much less than for normal faces.}

\(~\text{nonsignificant.}\)

\(^*p < .05.\)

\(^{**}p < .01.\)

\(^1p < .0001.\)

---

**Figure 3.** Effects of inversion and CR on the N170 (shown here at CB2 site) in the two tasks for all categories. Horizontal lines show that the N170 amplitude for faces-without-eyes did not differ from that to upright normal faces in either task. Similarly, the amplitude recorded to upright or inverted eyes was not different from that recorded to inverted or negative full faces (see Table 3 for pairwise comparisons).
inverted and negative faces-without-eyes were no longer processed like normal faces.

Topographies for the N170 to upright eyes and faces differed \( F(60, 720) = 5.91, p < .001 \), confirming that eyes and faces originate from different generators. Isolated eyes did not differ in N170 amplitude whether upside-down or upright (Table 2, Figures 3 and 4). Although the amplitude for negative eyes was slightly smaller than for upright eyes, the topographies did not differ for eyes in any format \( F(60, 720) < 2, p = .3 \) for upright-inverted comparison; \( F(60, 660) = 2.3, p = .07 \) for upright-negative], suggesting that eyes are always perceived as eyes, regardless of the transformation (Itier, Latinus, et al., 2006) (Figure 5).

Topographies differed between upright and inverted faces \( F(60, 720) = 3.47, p = .01 \) and upright and negative faces \( F(60, 660) = 4.51, p < .01 \) (Figure 5), reflecting different underlying generators between these different face formats. In contrast to what was found for upright faces and eyes, the topographies for eyes were no longer different from that to negative faces \( F(60, 660) < 2, p = .15 \) (Figure 5). Topographies were different between inverted eyes and inverted faces \( F(60, 720) = 4.24, p < .003 \), suggesting a possible additional neuronal recruitment for inverted faces. However, N170 amplitudes to inverted and CR faces, which significantly differed from that to upright faces (Table 2), did not differ significantly from that elicited by isolated inverted and negative eyes.
respectively (Table 2, Figures 3 and 4). These topography and N170 amplitude results support the hypothesis that the N170 to inverted and negative faces reflects mainly the processing of the eyes. Although N170 was slightly larger for inverted than negative faces \(F(1, 11) = 10.3, p < .01\), topographies did not differ between the two face types \(F(60, 660) < 2, p = .4\) despite these stimuli being very different, further supporting this hypothesis.

**DISCUSSION**

As predicted by our model, we found behavioral and electrophysiological evidence for considerably reduced face-specific effects when the eyes were removed from the face. These results indicate that the eye region is the main contributor to the classic inversion and CR effects reported for human faces and that the N170s to inverted and negative faces reflect mainly the processing of the eyes once the face configuration has been disrupted. In other words, by disrupting the relationships between the eyes and the rest of the face (i.e., the configuration), inversion and CR somehow isolate the eyes from the face context, resulting in the recruitment of eye-selective neurons in addition to the face-selective neurons (Figure 1).

Our recent MEG work (Itier, Herdman, et al., 2006) suggests that there is at least three different brain areas synchronously active around the timing of the N170: the FG and an occipital area that are mostly recorded by MEG, and a more lateral temporal area that we call the "STS region" (Allison et al., 2000) that is likely the main contributor of the N170 (Itier, Herdman, et al., 2006; Itier, Latinus, et al., 2006; Itier & Taylor, 2004a; Watanabe et al., 2003). The assumption that the N170 is coming primarily from the STS region is supported by the fact that the N170 is larger to eyes than to faces (Itier, Latinus, et al., 2006; Taylor, Edmonds, et al., 2001; Jemel, George, Chaby, Fiori, & Renault, 1999; Bentin et al., 1996), whereas the M170, a magnetic component recorded around the same time as the N170 and which seems to come mainly from the FG (Itier, Herdman, et al., 2006; Watanabe et al., 2003; Halgren, Raij, Marinkovic, Jousmaki, & Hari, 2000), is not different between faces and eyes (Taylor, George, & Ducorps, 2001). The face-specific intracranial N200 component is also larger to eyes than to faces on the superior and middle temporal gyri (STS region) but not on the FG surface (McCarthy et al., 1999). Our present model assumes that in the human STS region, face- and eye-selective cells coexist, with similar properties as the ones found in the monkey brain. Their synchronous activation is recorded with EEG on the scalp as the N170, due to a radial orientation of the generators (Itier, Latinus, et al., 2006; Itier & Taylor, 2004a; Watanabe et al., 2003).

According to single-unit recordings, face-selective cells respond to the face configuration (Perrett, Oram, & Ashbridge, 1998; Perrett et al., 1982, 1984) and yield a large face-sensitive N170 at posterior lateral sites when a normal upright face is presented (Figure 1). Eye-selective neurons, however, do not respond to the eyes within the face because of the face context (Perrett et al., 1982, 1985). The face N170 thus reflects mainly the activation of face-selective cells. This explains why removing the eyes from the face does not modulate the N170 amplitude (Eimer, 1998). Even without the eyes, the general first-order relations between the remaining facial features and the face contour are preserved and are enough to define the stimulus as a face. The presence of the eyes within the face, however, speeds up the response (Itier, Latinus, et al., 2006) as shown by an earlier latency of the N170 for faces compared to faces-without-eyes as found here and in a previous study (Eimer, 1998).

Although eye-selective neurons do not respond to the eyes of the face because of the face context (Perrett et al., 1982, 1985), face-selective neurons can respond to isolated eyes presented outside the face context (Perrett et al., 1982). Thus, when isolated eyes are presented, both face-selective and eye-selective neurons respond (Figure 1), yielding a larger N170 to isolated eyes than full faces (Itier, Latinus, et al., 2006; Taylor, Edmonds, et al., 2001; Jemel et al., 1999; Bentin et al., 1996). The eye N170 is thus coming from close-by, yet different, generators than the face N170, explaining why face and eye N170s have significantly different topographies as found here, as well as different developmental courses (Taylor, Edmonds, et al., 2001). When inverted or in photo-negative, eyes are still perceived as eyes and both eye- and face-selective neurons respond, yielding similar N170s to upright, inverted, and negative eyes as found here and in our previous study (Itier, Latinus, et al., 2006).

With inversion and CR, the face configuration is disrupted (Maurer et al., 2002; Rossion & Gauthier, 2002; Kemp et al., 1990). The eyes are no longer perceived in a normal face context so eye-selective neurons now respond to the inverted or negative eyes of the face, in addition to face-selective neurons (Figure 1). Both face- and eye-selective neurons thus respond to inverted or negative faces, explaining the increase in N170 amplitude that is no longer significantly different from the N170 amplitude recorded to isolated eyes as shown here and in our previous study (Itier, Latinus, et al., 2006). In contrast, when a face-without-eyes is presented upside-down or in photo-negative, the eye-selective neurons do not respond, only the face-selective neurons do, which explains that no increase in N170 amplitude is seen. In other words, by disrupting the relationships between the eyes and the rest of the face, inversion and CR isolate the eyes from the rest of the face, resulting in the recruitment of eye-selective neurons in addition to the face-selective neurons (Figure 1).

Smaller responses of face-selective cells with CR have been reported in the monkey literature (Ito et al., 1994), and although there are no available data for eye-selective neurons, it is likely that the same decrease in response would be found. Indeed, CR is a strong
manipulation that has been shown to alter gaze processing greatly (Tipples, 2005; Ricciardelli, Baylis, & Driver, 2000; Sinha, 2000). As eyes are defined by their shape and by the contrast between the iris and the sclera, inverting this contrast should affect the response of eye-selective neurons. This could explain why the N170 is smaller for negative than positive eyes and why, although not significantly different, there is a trend for smaller N170 for negative than positive faces-without-eyes (Figures 2 and 4). This could also explain the slightly smaller N170 to negative than inverted faces even though both reflect the processing of the eyes, as further supported by no significant differences in their respective topographies.

The slight increase in amplitude seen for inverted faces-without-eyes associated with a different topography between inverted faces-without-eyes and isolated eyes suggest that there may be an additional recruitment of object-selective neurons to process inverted faces-without-eyes. The reason why the same amplitude increase was not found with negative face-without-eyes is not clear. This could be due to an additional recruitment of object-selective neurons compensating for the smaller response of face- and eye-selective neurons to CR. However, the absence of a significant difference between the N170 topographies for negative faces-without-eyes and negative eyes does not support this object-selective neuronal recruitment. Alternatively, this could simply be due to the fact that the overall shape of the face is preserved in CR faces-without-eyes while disrupted with inversion, and thus, no object-selective neurons are recruited. Future cell recording studies in monkeys could address these questions by studying in more detail neuronal responses to these stimuli.

As mentioned earlier, the proposed model aims to reconcile discrepant results obtained in ERP, MEG, and fMRI techniques in humans and cell recording data in monkeys. Our hypothesis that inverted and negative faces reflect mainly the processing of the eyes is also in agreement with an absence of an N170 amplitude increase with the inversion of Mooney (Latinus & Taylor, 2005) and schematic faces (Sagiv & Bentin, 2001) in which the eyes are not real features. It has to be noted, however, that the present model does not imply that eye-selective neurons never respond to full faces. What we suggest here is that they likely do not respond at the level of the N170 when the configuration of the face is normal (upright). They are likely recruited later on for a detailed processing of the various aspects of the face such as its identity, gaze direction, gender, or emotion, given the important role of eyes in identity recognition (Ellis, Shepherd, & Davies, 1979), gender discrimination (Schyns, Bonnar, & Gosselin, 2002), emotions, direction of attention, and intentions (Baron-Cohen, Joliffe, Mortimore, & Robertson, 1997; Baron-Cohen, 1995). Future research should address further the dynamics of the response of these eye- and face-selective neurons.

Our idea that there exists eye-selective neurons in the human brain just like in the monkey brain is in agreement with the eye direction detector hypothesis (Baron-Cohen, 1995). Infants are attracted to eyes (Maurer, 1985), and as early as 3 months of age they orient to gaze (Hood, Willen, & Driver, 1998). The early gaze discrimination is even seen in neonates (Farroni, Csibra, Simion, & Johnson, 2002) and a preference for direct gaze is observed supposedly because direct gaze signifies social contact (Farroni, Massacesi, Menon, & Johnson, 2007; Hood et al., 1998; Hains & Muir, 1996; Maurer, 1985). Interestingly, humans are the only species with white sclera and have the largest index of exposed sclera size (Kobayashi & Kohshima, 1997), an evolution of the eye morphology that seems related to the specific processing of gaze and the development of an elaborated social cognition system (Emery, 2000; Kobayashi & Kohshima, 1997 for a review).

The implications of the present model for clinical populations are important. The eyes are the most attended feature in any given face task (Itier, Villate, & Ryan, 2007; Henderson, Williams, & Falk, 2005; Sekuler, Gaspar, Gold, & Bennett, 2004; Vinette, Gosselin, & Schyns, 2004; Schyns et al., 2002; Janik, Wellens, Goldberg, & Dell’osso, 1978; Luria & Strauss, 1978; Walker-Smith, Gale, & Findlay, 1977; Laughery, Alexander, & Lane, 1971; Yarbus, 1967). They play a crucial role in identity (Ellis et al., 1979; McKelvie, 1976) and emotion recognition, as well as in theory of mind (Baron-Cohen et al., 1997), our capability to attribute mental states to others. Many disorders such as Autistic Spectrum Disorders (ASD), Williams Syndrome, and some cases of prosopagnosia are linked to configural processing impairment of faces (Elgar & Campbell, 2001; de Gelder, Bachoud-Lévi, & Degos, 1998). The inversion effect is often not found in these populations who present similar recognition rates for upright and inverted faces and sometimes better rates for inverted faces (de Gelder et al., 1998). ASD patients are furthermore known to avoid eye contact and cannot infer mental states of others from the eyes only (Baron-Cohen et al., 1997). Interestingly, although the inversion effect on the N170 has rarely been studied in these populations, abnormally small N170 responses have been reported compared to controls (Grice et al., 2001; Eimer & McCarthy, 1999) and sometimes no face inversion effects on the N170 (Grice et al., 2001). In addition to a lack of emotional recognition and/or configural processing, there could simply be an abnormal processing of the eye region in those populations.

To conclude, our model suggests that what was previously taken as reflecting the early face specificity, that is, the disproportionate amplitude increase of the N170 with facial inversion and CR not seen for objects, in fact, reflects mainly the processing of the eyes of the face. Of course, faces-without-eyes are a very unnatural stimulus and may not capture the exact neuronal modulations of the face system. Nevertheless, it seems from these data that the early face specificity may lie in the eyes.
REFERENCES


Itier, R. J., Latinus, M., & Taylor, M. J. (2006). Inversion and contrast polarity reversal affect both encoding and recognition...


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


Taylor, M. J., Edmonds, G. E., McCarthy, G., & Allison, T.

Itier et al. 1825
Eyes first! Eye processing develops before face processing in children. NeuroReport, 12, 1671–1676.


