

# Inverting faces elicits sensitivity to race on the N170 component: A cross-cultural study

**Luca Vizioli**

Department of Psychology and Centre for Cognitive  
Neuroimaging (CCNi), University of Glasgow,  
United Kingdom



**Kay Foreman**

Department of Psychology and Centre for Cognitive  
Neuroimaging (CCNi), University of Glasgow,  
United Kingdom



**Guillaume A. Rousselet**

Department of Psychology and Centre for Cognitive  
Neuroimaging (CCNi), University of Glasgow,  
United Kingdom



**Roberto Caldara**

Department of Psychology and Centre for Cognitive  
Neuroimaging (CCNi), University of Glasgow,  
United Kingdom



Human beings are natural experts at processing faces, with some notable exceptions. Same-race faces are better recognized than other-race faces: the so-called other-race effect (ORE). Inverting faces impairs recognition more than for any other inverted visual object: the so-called face inversion effect (FIE). Interestingly, the FIE is stronger for same- compared to other-race faces. At the electrophysiological level, inverted faces elicit consistently delayed and often larger N170 compared to upright faces. However, whether the N170 component is sensitive to race is still a matter of ongoing debate. Here we investigated the N170 sensitivity to race in the framework of the FIE. We recorded EEG from Western Caucasian and East Asian observers while presented with Western Caucasian, East Asian and African American faces in upright and inverted orientations. To control for potential confounds in the EEG signal that might be evoked by the intrinsic and salient differences in the low-level properties of faces from different races, we normalized their amplitude-spectra, luminance and contrast. No differences on the N170 were observed for upright faces. Critically, inverted same-race faces lead to greater recognition impairment and elicited larger N170 amplitudes compared to inverted other-race faces. Our results indicate a finer-grained neural tuning for same-race faces at early stages of processing in both groups of observers.

Keywords: face recognition, face inversion effect, electrophysiology, culture

Citation: Vizioli, L., Foreman, K., Rousselet, G. A., & Caldara, R. (2010). Inverting faces elicits sensitivity to race on the N170 component: A cross-cultural study. *Journal of Vision*, 10(1):15, 1–23, <http://journalofvision.org/10/1/15/>, doi:10.1167/10.1.15.

## Introduction

Human beings are natural experts at recognizing faces. Brain imaging studies have shown that this visual expertise is implemented by a specialized cortical network located in the occipital temporal cortex (Haxby, Hoffman, & Gobbini, 2000), optimally tuned to process this particular visual category (Caldara & Seghier, 2009; Caldara et al., 2006). However, such a highly developed biological skill is markedly impaired in processing two specific face categories: other-race and inverted faces.

The face recognition impairment observed for other-race (OR) compared to same-race (SR) faces has long been reported in the literature (Feingold, 1914), a phenomenon often referred to as the Other-Race Effect (ORE—Malpass & Kravitz, 1969; for a review see

Meissner & Brigham, 2001). Although the neural mechanisms at play are yet to be understood, it is widely accepted that visual expertise plays a crucial role in shaping this recognition deficit. Many studies advocate the existence of finely tuned mechanisms to process SR faces, probably developed as a by-product of visual experience (Caldara & Abdi, 2006; Michel, Caldara, & Rossion, 2006; Michel, Rossion, Han, Chung, & Caldara, 2006; Tanaka, Kiefer, & Bukach, 2004; Walker & Tanaka, 2003), which can plausibly be held accountable for SR face identification advantage. This identification advantage is paired with faster performance in race categorization tasks for OR faces (Caldara, Rossion, Bovet, & Hauert, 2004; Levin, 1996; Valentine & Endo, 1992). To account for these findings, Valentine and Endo (1992) suggested a model in which individual faces are stored in a multidimensional space as a function of experience.

Within this space, OR faces are more closely clustered together than SR faces. These differences in the spatial distribution of the face exemplars would translate in be responsible for faster OR race categorization but impaired identification. A different account was provided by Levin (2000) who suggested that race information is a feature that is quickly extracted from OR faces at the cost of the information required for fine individual discrimination.

Picture plane inversion of faces also affects face recognition. Inverting faces has been consistently linked to a significantly greater recognition deficit compared to any other visual category (Yin, 1969): the so-called Face Inversion Effect (FIE). The FIE is thus regarded as one of the strongest evidence for specialized face processing. Although the putative neural mechanisms underlying this phenomenon have yet to be clarified, some level of consensus seems to reign around the idea that visual expertise may be accountable for the effects of face inversion (e.g. de Haan, Pascalis, & Johnson, 2002; Rossion & Gauthier, 2002). One of the main hypotheses involving *qualitative* processing differences holds that, as a byproduct of experience, the human brain engages into distinct mechanisms in processing upright and inverted faces: *holistic* mechanisms for upright and *featural* mechanisms for inverted faces; or at least that picture plane inversion significantly impairs holistic processing of faces (Rossion, 2008). One can thus infer that the greater level of expertise with SR faces, which translates in higher levels of holistic processing (e.g. Michel, Caldara et al., 2006; Michel, Rossion et al., 2006), should elicit a stronger FIE compared to OR faces. This prediction is supported by converging evidence showing that inverting SR faces leads to a greater impairment in recognition performance compared to OR faces (Buckhout & Regan, 1988; McKone, Brewer, MacPherson, Rhodes, & Hayward, 2007; Murray, Rhodes, & Schuchinsky, 2003; Rhodes, Tan, Brake, & Taylor, 1989). To date, only a single study reported the opposite pattern of results: Valentine and Bruce (1986) found a larger FIE for OR faces compared to SR faces. However, these authors attempted to equate performance by modulating the time of exposure for encoding SR and OR faces (of 2 and 5 seconds respectively), rendering these findings incomparable across groups of observers.

Electrophysiological studies in humans have significantly contributed to clarify the time course of visual processing. With regards to faces, a critical time window occurs between 130 to 200 ms after stimulus onset, in which the early Event-Related Potential (ERP) N170 component takes place. The N170 is a negative deflection in the ERP signal occurring roughly 170 ms after stimulus onset, peaking at occipitotemporal sites, which tends to be larger for faces than any other visual object categories (for a recent review see Rossion & Jacques, 2008). Since Bentin, Allison, Puce, Perez, and McCarthy (1996) seminal work, the speculative debate on the putative mechanisms underlying the N170 has yet to be resolved. While

originally this component was believed to reflect the structural encoding of faces prior to individual face recognition (Bentin et al., 1996), recent evidence has shown N170 modulation to identity (Jacques, d'Arripe, & Rossion, 2007; Jacques & Rossion, 2006) and levels of expertise (Rossion & Gauthier, 2002; Tanaka & Curran, 2001), suggesting that it may reflect more advanced stages of face processing. Earlier studies have identified a positive component peaking in the time window of the N170 at central and frontal sites: the Vertex Positive Potential (VPP—Botzel & Grusser, 1989; Jeffreys, 1989). Like the N170, the VPP has been reported to respond with larger amplitude to faces than other visual category, suggesting that both components reflect the same neural process (Joyce & Rossion, 2005).

Despite the large amount of behavioral findings supporting the ORE, electrophysiological studies have as yet failed to achieve a degree of consistency in the understanding of this phenomenon. The main ongoing debate lies in whether the early N170 face-preferential component is sensitive to race. Research has thus far produced contrasting results. A number of studies have failed to report sensitivity to race on the N170 (e.g. Caldara et al., 2003, 2004; Tanaka & Pierce, 2009), or on the related VPP (Ito, Thompson, & Cacioppo, 2004), suggesting that the race of the stimuli is not processed until later stages (i.e. roughly 250–300 ms after stimulus onset—Caldara et al., 2003, 2004; Tanaka & Pierce, 2009), or that the N170 is not be sensitive enough to capture race information (Caldara et al., 2003). However, other studies have found OR faces eliciting larger N170 amplitudes than SR faces (Stahl, Wiese, & Schweinberger, 2008; Walker, Silvert, Hewstone, & Nobre, 2008), with two studies reporting larger N170 amplitude for SR faces (Herrmann et al., 2007; Ito & Urland, 2005).

The heterogeneity of the N170 results described above may be explained by considerable methodological differences across studies. Indeed, race modulation of the N170 appears to depend on task demands (Caldara et al., 2003, 2004; Herrmann et al., 2007; Ito & Urland, 2003). Another point of interest is that albeit global low-level visual properties of face stimuli, such as amplitude spectrum, luminance and contrast, might affect early electrophysiological components such as P1 and N170 (Dakin, Hess, Ledgeway, & Achtman, 2002; Rousselet, Pernet, Bennett, & Sekuler, 2008), none of the previous studies attempted to control for or even address *all* these factors. This is surprising considering the low-level differences that co-vary with different races of faces. Furthermore, previous studies have used different types of OR stimuli, including EA, AA and Hispanic (e.g. Caldara et al., 2003, 2004; Herrmann et al., 2007; Ito & Urland, 2003, 2005; Tanaka & Pierce, 2009; Walker et al., 2008) faces, leading to contrasting results. Critically, all the studies that have reported modulation to race on the N170 have only relied on data gathered from a single population: the WC population. This methodological shortcoming undermines generalizations of the

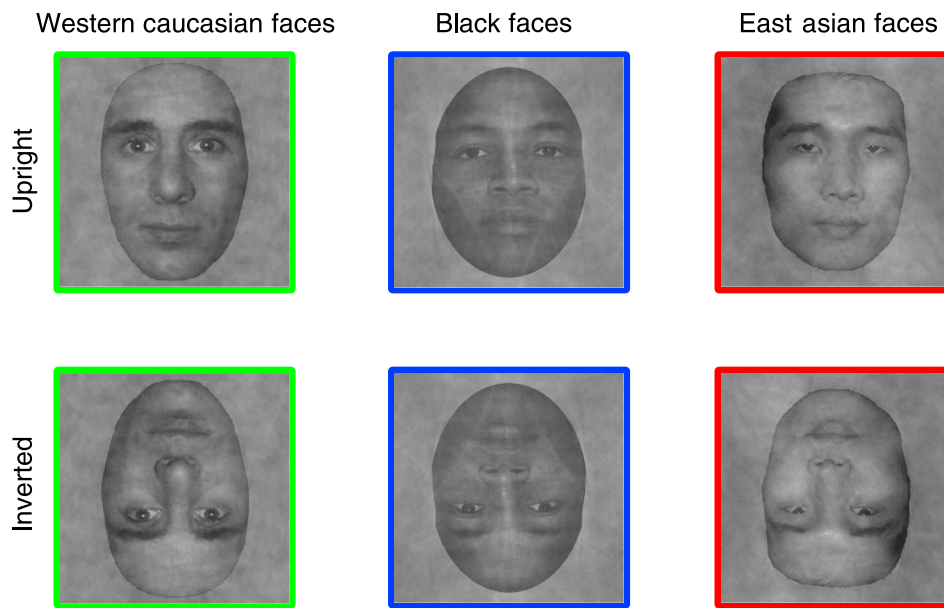


Figure 1. Example of the face stimuli used in the experiment. To avoid potential low level confounds on the ERP signal, amplitude spectrum, luminance and contrast were normalized to all images' mean.

results, because any effect could be confounded by differences in the face stimuli. A full cross-over interaction between races of observers and face stimuli is crucial to assess genuine ORE modulations on the ERP signal.

Although behavioral research has extensively investigated the mutual effects of ORE and FIE, to date only one electrophysiological study has exploited the conjoint advantages offered by the combinations of these two phenomena (Wiese, Stahl, & Schweinberger, 2009). Wiese et al. (2009) tested the effects of the ORE and FIE on the N170 using EA and WC faces on WC observers only. They reported a general amplitude increase for inverted compared to upright faces and latency delays on the N170 for inverted faces as a function of race. However, Wiese et al. (2009) only tested one group of observers (the WC population) and did not control for low-level visual properties of the stimulus' set.

In summary, The FIE is regarded as a well documented marker of visual expertise for faces (Yin, 1969) both at the electrophysiological (Bentin et al., 1996; Rossion et al., 1999, 2000) and the behavioral level (e.g. Freire, Lee, & Symons, 2000). Observers respond less accurately and with longer reaction times when recognizing inverted faces. Moreover face inversion elicits significantly delayed N170 (Bentin et al., 1996; Rossion et al., 1999, 2000) and VPP (Jeffreys, 1996), often accompanied by amplitude increases (Itier & Taylor, 2004a, 2004b; Rossion et al., 2000; Rossion, Joyce, Cottrell, & Tarr, 2003). These effects have been interpreted as a disruption of holistic processing for faces (Rossion & Gauthier, 2002).

Here we took advantage of the conjoint effects of the ORE and FIE to investigate whether the N170 is sensitive

to race in two groups of observers. We recorded scalp EEG in WC and EA observers presented with WC, EA and AA faces in upright and inverted orientations. Participants were required to perform an orthogonal task, consisting in the detection of colored faces in the stimulus sequence. Moreover, to control for potential low-level confounds in the ERP signals that would relate to differences in the low-level properties across faces from different races, we normalized the amplitude-spectrum, luminance and contrast of the entire image set (see Figure 1). In addition, we used a second “other race” category of stimuli (i.e., AA faces) as the absence of differences across the two OR faces (i.e. EA and AA for WC observers and WC and AA for EA observers) would further strengthen the claim that any observed ERP modulations are related to the ORE. A separate behavioral experiment was designed to assess the levels of ORE and FIE in both groups of observers. Possibly due to higher levels of expertise, SR faces have been reported to be processed more holistically than OR faces. We therefore hypothesized that SR faces would elicit a larger inversion effect compared to OR faces. At the electrophysiological level, we predicted FIE modulations as a function of the race of the faces on the face preferential N170 component.

We did not find any significant difference in the upright conditions. However, our data show sensitivity to race on the N170 across both groups of observers for inverted faces, with larger N170 amplitude modulations between upright and inverted faces (i.e. the magnitude of the FIE) for SR compared to OR faces. Moreover, although a significant delay for the onset of the N170 was observed for inverted compared to upright faces, we did not observe

latency differences across races. The electrophysiological modulation was paired with greater recognition impairment for inverted same-race faces compared to other-race faces.

## Methods

### Participants

30 subjects (15 East Asians (EA) and 15 Western Caucasian (WC), with an age range of 19–30 and a mean of 23) took part in the experiment. All the EA participants were Chinese; they had been in the country for less than 1 month and had previously never lived in a Western society. All participants provided written informed consent and had normal or corrected-to-normal vision. All subjects were right handed and 15 (8 EA and 7 WC) were female. The experiments received the approval of the local ethical committee.

### Stimuli

The stimuli consisted of 30 front-view grayscale photographs of WC, EA and African American (AA) faces (5 identities  $\times$  2 genders  $\times$  3 races) occupying approximately a visual angle of  $3.75^\circ \times 4.25^\circ$  (see Figure 1). We used the WC and EA face database used in previous studies (Michel, Caldara et al., 2006; Michel, Rossion et al., 2006). The AA face stimuli were obtained from a subset of identities of the Productive Aging Lab Face Database (Minear & Park, 2004). The inclusion of a second “other race” category (i.e. AA faces) represented a further control on the ERP signal.

Each identity was equally presented in 2 orientations (upright and inverted). All faces were cropped to remove external features; none had particular distinctive features and male faces were clean-shaven. The stimuli were centered in a  $5.2^\circ \times 5.2^\circ$  background of average luminance ( $25.4 \text{ cd/m}^2$ ,  $23.5^\circ \times 30.1^\circ$ ). The stimuli were equated in terms of spatial frequency content by taking the average of the amplitude spectra of all 30 stimuli and combining that average spectrum with the original phase spectra to reconstruct each individual stimulus. The RMS (root mean square) contrast (i.e. the standard deviation of the pixel intensities) was also kept constant across faces. Because form information is largely carried by phase rather than amplitude (Rousselet et al., 2008), individual faces remain easily recognizable after this manipulation, however still ensuring that any race related differences in the EEG is not simply a function of differences in the relative visibility of specific frequency components in the stimuli.

## Experimental design

### EEG study

Participants were seated in a dimly lit, sound-attenuated electrically shielded room. Viewing distance was maintained at 80 cm by use of a chinrest. Stimuli were randomly interspersed and sequentially presented for 200 ms (16 frames at 80 Hz) on a Samsung SyncMaster 1100 MB monitor (resolution  $2048 \times 1536$ ), with an inter-stimulus interval (ISI) randomly varying between 2200 and 3200 ms. A black  $0.3^\circ$  fixation cross was presented in the middle of the screen before stimulus onset and remained on the screen during the whole ISI duration. Each trial had thus a random duration ranging between 2400 and 3400 ms.

Participants were instructed to perform an orthogonal task that required pressing the “s” key on the keyboard every time a green face appeared on the screen and the “k” key for a red face. This orthogonal task was designed to avoid potential signal modulations due to attentional confounds linked to the race of the stimuli (Golby, Gabrieli, Chiao, & Eberhardt, 2001). Red and green faces appeared in approximately 8% of the trials (i.e. 15 red and 15 green faces). Each face appeared once either as a red or a green stimulus.

The experiment consisted of 5 blocks of 78 trials each (360 trials in total with 60 trials  $\times$  condition  $\times$  orientation and 30 red and green faces) and lasted approximately 20–25 minutes.

### Behavioral study

Subsequently to the EEG experiment, to directly assess the level of ORE and FIE, we recalled the same subjects asking them to take part in a small behavioral experiment. 9 out of 15 subjects per group participated in this experiment; the others failed to reply to our query. Participants sat in the same dimly lit room with the same viewing conditions, and were presented with 20 new faces for each race, one at the time (3s each, 5s ISI). The faces were obtained from the KDEF (Lundqvist, Flykt, & Öhman, 1998) and AFID (Bang, Kim, & Choi, 2001) databases. All faces were selected from a set of stimuli different from the set used in the EEG experiment, and were also equated for global amplitude spectrum, luminance, and contrast. Stimulus size subtended  $3.75^\circ \times 4.25^\circ$  of visual angle. Participants were told explicitly to memorize the faces. The encoding phase was followed by a forced-choice old-new recognition task whereby 40 faces (20 old and 20 new) were presented individually. Participants underwent two blocks of the old/new face recognition task per race with faces displayed in two orientations (i.e., upright and inverted). The blocks were counterbalanced across participants. Faces were blocked by race during both the encoding and the recognition phase. The encoding stage was followed by a 1 minute

pause, after which participants had to indicate whether each face was old or new by pressing the “s” key on the computer keyboard for old faces and the “k” key for new faces. Each face remained on the computer screen until the participant’s response, or for a maximum of 200 ms (3s ISI). Participants did not know the ratio of old to new faces and did not receive any feedback on their responses. The order and the number of presentation of SR and OR stimuli were pseudorandom, changing across subjects. As in previous studies (e.g. Carroo, 1986),  $d'$  indices (Swets, Tanner, & Birdsall, 1961) for WC and EA faces were calculated for each participant to assess accuracy.

## EEG recording and analysis

### EEG recording

EEG data were acquired by means of a 128-channel BioSemi Active Two EEG system (BioSemi, Amsterdam, Netherlands). Electrodes were placed in a nylon cap according to the 10–5 system (Oostenveld & Praamstra, 2001). Vertical and horizontal electro-oculograms were recorded by attaching 4 additional electrodes (UltraFlat Active electrodes, BioSemi) below and at the outer canthi of both eyes.

Analog signal was digitized at 1024 Hz and band-pass filtered online between 0.1 and 200 Hz. An active electrode (common mode sense—CMS) and a passive electrode (driven right leg—DRL) were used to comprise a feedback loop for amplifier reference, and electrodes impedances were kept between  $\pm 20$  k $\Omega$ . Subjects were asked to minimize blinking, head movement, and swallowing.

### EEG pre-processing

EEG analysis was performed using EEGLAB (Delorme & Makeig, 2004), Matlab 7.5 (2007b) and BESA 5.2. In BESA, EEG data were referenced on-line to an average reference. Noisy electrodes were rejected on a subject-by-subject basis. The signal was low-pass filtered offline at 40 Hz with a slope of 6 dB. Baseline correction was performed using 150 ms of pre-stimulus onset. Artifacts were rejected based on absolute abnormal values larger than 120  $\mu$ V. Trials were averaged across an epoch of  $-200$  ms to  $+600$  ms. Trials including red or green faces were excluded from the analysis.

Before averaging, single trials were corrected for horizontal and vertical eye movement and blinking artifacts by means of PCA. Due to their specific topographical configuration, we first identified these artifacts on the continuous signal. We then manually selected portions of the segment showing the topographical configuration of interest. We thus averaged the selected epochs to create a subject-specific template for that particular artifact. The first PCA component (accounting

for 89% to 99%—mean 92%—of the variance for individual subjects) was removed.

Across subjects and conditions, the minimum number of trials accepted after artifact rejection was 252, the maximum 360, and the mean 343. The rest of the analyses were performed in EEGLAB and Matlab.

Amplitudes and peak latencies were measured at electrodes O1 and O2 for the P1 and P2 components (i.e. where the amplitude of such components was maximal), and at 9 pairs of occipitotemporal electrodes over the left and right hemisphere for the N170, and N250. To select the electrodes of interest, we performed the mean average of the ERPs scalp topography at the mean latency of the N170 for upright and inverted stimuli across all the subjects and selected the electrodes with maximal amplitudes (see Figure 1 in the [supplementary](#) section). The channels previously excluded consequently to artifact rejection were interpolated using the EEGLAB *topoplot* function. The electrodes selected, consistently with the N170 literature (e.g. Bentin et al., 1996; Rössion & Jacques, 2008), were: P4, P6, PPO6h, P8, P8h, PO8, P10, PO10h, PO10 in the right hemisphere; and P3, P5, PPO5h, P7, P7h, PO7, P9, PO9h, PO9 in the left hemisphere (see Figure 1 in the [supplementary](#) section).

For each subject, each component amplitude was quantified as the maximal positive (for P1 and P2) or negative (for N170 and N250) voltage measured within a 40 ms time window centered on the Grand Average ERP latency.

## Statistical analyses

In terms of statistical analyses, we employed canonical methods (i.e. ANOVA and  $t$ -tests) as well as robust approaches (i.e. bootstrap). Bootstrap has higher statistical power and makes fewer assumptions on the distribution of the data compared to more canonical statistical methods. Our goal was to address a growingly popular issue that has often been overlooked within the electrophysiological literature: the examination of the number of subjects showing a similar pattern of results (e.g. Rousselet et al., 2008). It is worth noting that most EEG and MEG papers do not report the number of subjects showing a given effect, although there is a recent trend in the literature favoring this view (e.g. Philiastides & Sajda, 2006; Schyns, Petro, & Smith, 2007; Smith, Gosselin, & Schyns, 2007). Ensuring that an effect is observed across all or most subjects is essential because in some situations an effect might be driven by a minority of subjects (for instance the early P1 difference reported by Rousselet et al., 2008). This is not to say that such effects are not interesting, rather that they should be interpreted with caution.

A first statistical analysis was performed on the P1 N170, P2 and N250 peak amplitudes and latencies. We carried out a mixed model repeated measures 5 way

ANOVAs on the N170 and N250 components. The 5 factors included in the ANOVA were: 2 *groups* of observers (i.e. WC and EA participants); 3 *races* of the stimuli (i.e. *stimulus race*: WC, AA and EA); 2 *orientations* in which the stimuli were presented (i.e. upright and inverted); 2 *hemispheres*; and the *electrodes* of interest (as described above). Mixed model repeated measures 4 way ANOVAs (*group* × *stimulus race* × *orientation* × *hemisphere*) were carried out on P1 and P2 components. These analyses did not include the *electrode* factor as only one electrode per hemisphere was considered. A further 5 way ANOVA (*group* × *stimulus race* × *orientation* × *electrode* × *hemisphere*) was carried out with the latency of the above mentioned components as the dependent variable.

In light of the results, we then further carried out two 4 ways ANOVA (*hemisphere* × *electrodes* × *stimulus race* × *group*) on the N170 peak amplitude, one for the upright and one for the inverted conditions.

We finally assessed the difference between two conditions across subjects at each electrode independently using *t*-tests. The differences assessed were WC vs. AA, WC vs. EA, and EA vs. AA for WC Observers; and EA vs. AA, EA vs. WC, and WC vs. AA for EA Observers.

Given the rather heterogeneous, asymmetrical nature of the distributions of ERP differences (computed both between upright minus inverted faces per each condition independently—i.e. the magnitude of the FIE—and between conditions for each orientation independently), and the *d'* scores, which were also limited in sample size, we carried out percentile bootstrap analyses. We sampled subjects with replacement, averaging the means across participants independently for each condition, and then computing the difference between the means for the two conditions (for instance inverted WC vs. AA). This process was repeated 999 times, leading to a distribution of bootstrapped estimates of the mean difference between two ERP conditions, averaged across subjects. Then the 95% percent confidence interval was computed ( $\alpha = 0.05$ ). Finally, the difference between the two sample means was considered significant if the 95% confidence interval did not include zero. Note that this bootstrap technique, relying on an estimation of H1, tends to have more power than other robust methods like permutation tests and related bootstrap methods that evaluate the null hypothesis H0 (Wilcox, 2005).

We then carried out correlation analyses to assess the relationship between the N170 and behavioral magnitudes of the FIE. FIE indexes were calculated at the N170—by subtracting the N170 peak amplitude elicited by inverted to that elicited by upright faces—and behavioral—the *d'* scores for upright minus those for inverted faces—levels. Pearson correlation coefficients were calculated between the 2 FIE indexes for each group of observers and for each race of the faces. Again, because of the limited sample size and the asymmetrical nature of the distribution of the data, bootstrap tests of independence were also carried

out. We sampled with replacement pairs of data points (i.e. the N170 and *d'* FIE indexes), and then computed the Pearson correlation between these pairs, across subjects, for each stimulus race independently. This process was repeated 599 times, leading to a distribution of bootstrapped Pearson correlation coefficients between the two FIE indexes. Then, 95% confidence intervals were computed using the special adjustments suggested by Wilcox (2005). Finally, correlations were considered significant if their 95% confidence intervals did not include zero ( $\alpha = 0.05$ ).

Our second statistical analysis was data driven. Because we were interested in any significant interaction on the early cerebral dynamics evoked by faces of different races (across groups of subjects) (i.e. the first 600 ms after stimulus onset), we carried out an ANOVA independently at all the electrodes and all the time points. First we calculated the index of the magnitude of the FIE at the electrophysiological level by computing the differences between the amplitude of ERPs elicited by upright minus those elicited by inverted faces independently for each stimulus' race, at all the electrodes and all the time points. We then carried out a 2 way ANOVA (*group* × *stimulus race*) on these differences across all electrodes and all time points independently. Post-hoc bootstrap tests were subsequently carried out between paired conditions by sampling subjects with replacement, as described earlier. Again 95% percent confidence intervals were computed, with the difference between the means considered significant if a confidence interval did not include zero.

No correction for multiple comparisons was applied. However, significant differences were not randomly scattered (as it might be expected with the significant points being the result of random type I errors) but formed consistent spatiotemporal clusters.

## Results

We found four main results. First, both groups of observers were significantly less accurate at recognizing inverted compared to upright same race (SR) faces, but showed no behavioral inversion effect for other race (OR) and African American (AA) faces. Second, regardless of the group of observers, the race of the face stimuli modulated the magnitude of the N170 amplitude face inversion effect (FIE), which was largest for SR faces, with no differences between OR and AA stimuli. No race effects were observed on P1, P2, and N250. Third, no race effects were observed on the peak amplitude of the N170 for upright faces in both groups of observers. Finally, the behavioral FIE positively correlated with the N170 FIE for SR faces only. We report the analyses of all the behavioral and ERP results below.

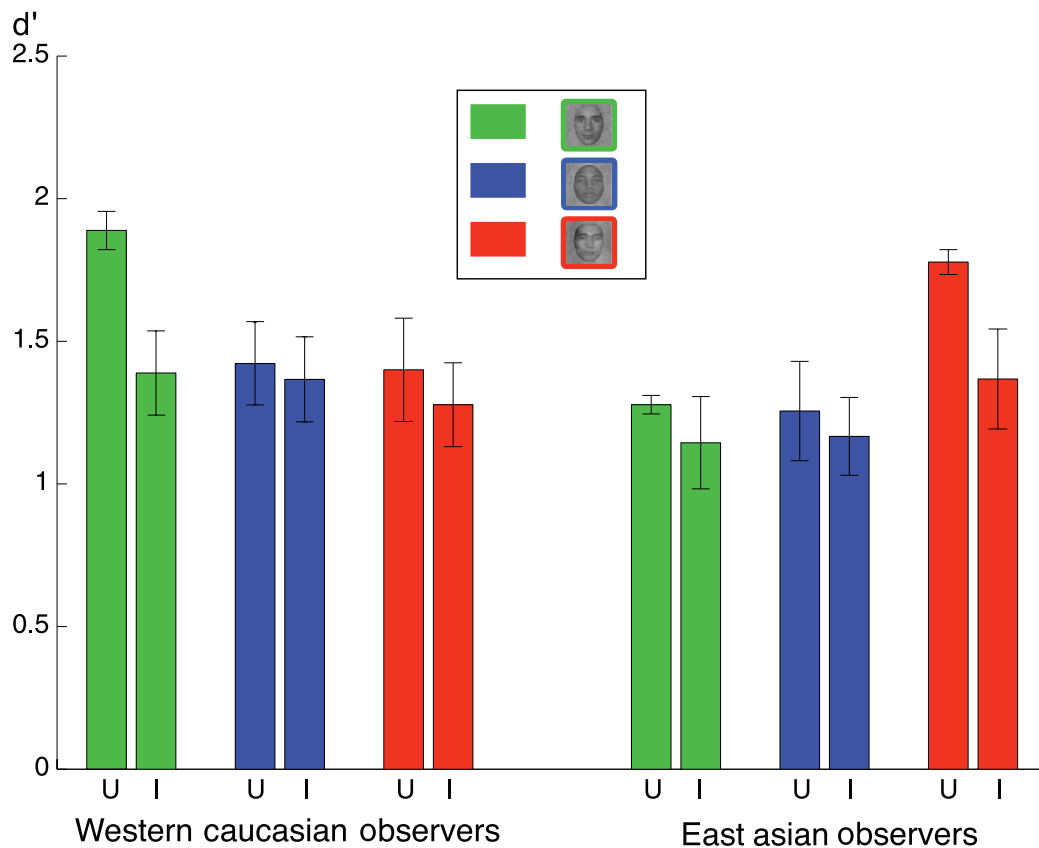


Figure 2. Mean  $d'$  scores and standard errors of the old/new face recognition task, for Western Caucasian (green), East Asian (red) and African American (blue) faces for the upright (U) and inverted (I) orientations.

## Behavioral results

We observed significant main effects for *orientation* ( $F(1,16) = 216.28$ ;  $p < .01$ ), with both groups being significantly more accurate at recognizing upright compared to inverted faces; and *race of the stimuli* ( $F(1,16) = 39.24$ ;  $p < .01$ ), with the lowest  $d'$  for AA faces.

As shown in Figure 2, both groups of observers exhibited a larger FIE (represented by the differences between  $d'$  scores for upright and inverted stimuli) for SR compared to OR and AA faces. Consistently, the 3-Way ANOVA (*groups*  $\times$  *orientations*  $\times$  *stimuli*) showed significant interactions between *Group*  $\times$  *Stimuli* ( $F(2,15) = 131.03$ ;  $p < .01$ ) and *group*  $\times$  *stimuli*  $\times$  *orientation* ( $F(2,15) = 16.11$ ;  $p < .01$ ). Although, within both groups, all subjects but one consistently displayed larger  $d'$  for upright compared to inverted SR faces only, a paired  $t$ -test contrasting upright and inverted SR faces showed a significant difference for WC ( $t(8) = 3.72$ ,  $p < .01$ ), but not for EA participants ( $t(8) = 2.16$ ,  $p = .06$ ). These low  $t$  values seem to be due to small sample sizes and violations of homoscedasticity. Bootstrap confidence intervals, which are less sensitive to these factors, showed that regardless of the group of observers, participants were significantly impaired at recognizing inverted compared to

upright SR faces (WC: mean difference = .53, bootstrap CI of the mean difference = [.27, .8],  $p < .01$ ; EA: 0.29, [.07, .59],  $p < .01$ ). No differences between upright and inverted  $d'$  score were observed for OR (WC:  $t(8) = .47$ ,  $p > .05$ ; mean difference = 0.12 [−.34, .58],  $p > .05$ ; EA:  $t(8) = .93$ ,  $p > .05$ ; mean difference = 0.13 [−.13, .39],  $p > .05$ ) and AA faces (WC:  $t(8) = .33$ ,  $p > .05$ ; mean difference = .06 [−.26, .37],  $p > .05$ ; EA:  $t(8) = .34$ ,  $p > .05$ ; mean difference = .09 [−.43, .56],  $p > .05$ ).

## EEG results

### P1

Table 1 shows P1 results. P1 amplitude was significantly modulated by face *orientation* ( $F(1,28) = 30.21$ ;  $p < .01$ —Figure 3). Inverted faces elicited significantly larger P1 (mean = 3.62  $\mu$ V; std = 1.63  $\mu$ V) compared to upright faces (mean = 2.96  $\mu$ V; std = 1.43  $\mu$ V). Importantly, no effect of stimulus race was observed on the amplitude elicited by faces presented in either orientation (*groups*  $\times$  *stimuli*  $F(2,27) = .56$ ;  $p > .05$ ; *groups*  $\times$  *stimuli*  $\times$  *orientations*  $F(2,27) = .741$ ;  $p > .05$ ).

P1 elicited by inverted faces (mean lat. = 111 ms; std = 7.4 ms) was significantly delayed by 3.4 ms compared to

Groups	Electrode	Orientation	Stimuli	Mean Amplitude	Std. Deviation	Mean Latency	Std. Deviation
Western Caucasian Observers	<b>O2</b>	<i>upright</i>	Western Caucasian	3.51	1.89	108	9
			African American	3.30	2.01	108	15
			East Asian	3.21	2.11	109	12
		<i>Inverted</i>	Western Caucasian	4.21	2.26	115	11
			African American	4.03	2.12	111	14
			East Asian	3.97	1.97	114	11
	<b>O1</b>	<i>upright</i>	Western Caucasian	4.27	4.10	109	13
			African American	3.87	4.07	109	14
			East Asian	3.58	3.41	108	11
		<i>Inverted</i>	Western Caucasian	4.67	4.32	110	14
			African American	4.51	4.08	109	14
			East Asian	4.63	4.70	112	13
East Asian Observers	<b>O2</b>	<i>upright</i>	Western Caucasian	2.71	1.41	104	11
			African American	2.69	1.16	107	11
			East Asian	2.52	1.23	107	9
		<i>Inverted</i>	Western Caucasian	3.16	1.88	110	9
			African American	3.00	1.25	109	9
			East Asian	3.02	1.56	108	11
	<b>O1</b>	<i>upright</i>	Western Caucasian	2.10	1.33	108	11
			African American	1.80	0.88	108	13
			East Asian	1.89	1.14	111	12
		<i>Inverted</i>	Western Caucasian	2.95	1.58	113	11
			African American	2.36	1.32	111	10
			East Asian	2.80	1.72	114	12

Table 1. P1.

that elicited by upright faces (mean lat. = 108 ms; std = 7.3 ms;  $F(1,28) = 31.907$ ;  $p < .01$ ). The latency of this component, like its amplitude, was not modulated by stimulus race in either orientation ( $group \times stimuli—F(1,28) = .35$ ;  $p > .05$ ;  $group \times stimuli \times orientations—F(2,27) = 1.42$ ;  $p > .05$ ). No other significant differences were observed.

### N170

The mean peak amplitudes, latencies and standard deviations in each condition are reported in [Table 2](#).

[Figure 4](#) shows the mean ERPs at the right hemisphere electrode PO8, where the N170 amplitude was largest ( $F(8,21) = 4.14$ ,  $p < .01$ ), and the left hemisphere electrode PO7, for both stimulus orientations and both subject groups. Our main result indicates that stimulus race systematically modulates the FIE magnitude of the N170 amplitude across groups, which was largest for SR compared to OR or AA faces ( $orientation \times stimulus \times group—F(2,27) = 4.62$ ,  $p < .05$ —[Figures 4](#) and [6](#)). Moreover, inverted faces elicited significantly larger N170 compared to upright faces regardless of stimulus race and the group of observers ( $F(1,28) = 60.41$ ,  $p < .01$ ).



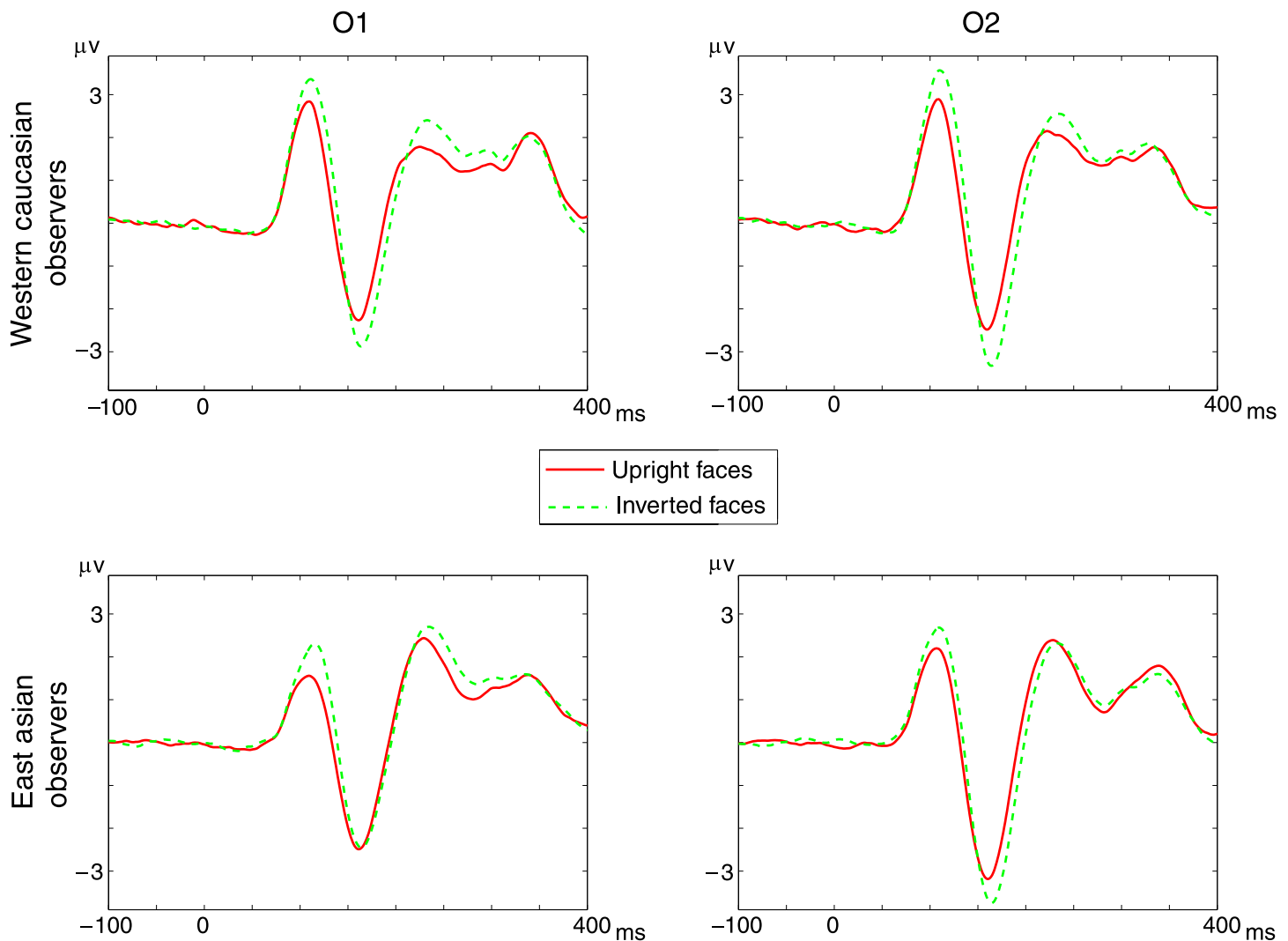


Figure 3. Grand average ERP waveforms of Western Caucasian and East Asian observers elicited by Upright (red) and Inverted (dotted green) faces at two occipital sites showing the largest P1 amplitude (O1, left hemisphere; O2, right hemisphere). The ERPs were obtained averaging together faces across conditions (i.e. Western Caucasian, East Asian and African American) independently per each orientation. While P1 was not modulated by the race of the faces in either orientation, a significant inversion effect was observed regardless of the race of the face stimuli, with inverted faces eliciting larger P1 compared to upright faces. No latency differences were observed across orientations.

Bonferroni corrected  $t$ -tests carried out between the peak amplitudes elicited by upright and those elicited by inverted faces independently for each stimulus race show that this effect was significant for all conditions ( $p < .05$ ).

We also observed a main *hemispheres* effect, with larger amplitudes over the right hemisphere ( $F(1,28) = 6.15, p < .05$ ); and a main effect of stimulus race ( $F(2,27) = 6.7, p < .01$ ). Bonferroni corrected paired  $t$ -tests showed that AA faces elicited the smallest amplitude ( $p < .05$ ), while no differences were observed between WC and EA faces. The 5 way ANOVA further showed significant interactions between *hemisphere*  $\times$  *orientation* ( $F(1,28) = 20.04, p < .01$ ), *orientation*  $\times$  *stimulus* ( $F(2,27) = 6.32, p < .01$ ); *electrode*  $\times$  *stimulus* ( $F(16,13) = 4.06, p < .01$ );

and *hemisphere*  $\times$  *stimulus*  $\times$  *orientation* ( $F(2,27) = 3.49, p < .05$ ).

To clarify these results, we carried out two 4 way ANOVAs on the N170 amplitudes (*hemisphere*  $\times$  *electrodes*  $\times$  *stimulus*  $\times$  *group*), one for each orientation. Crucially, this analysis indicated that for the upright orientation, stimulus race did not modulate the N170 amplitude (*stimuli*  $\times$  *groups*— $F(2,27) = 1.89, p > .05$ ). However, inverted SR faces elicited the largest amplitude regardless of the group of observers (*stimuli*  $\times$  *groups*— $F(2,27) = 9.15, p < .01$ ). A significant *electrode* effect indicated that the largest N170 was measured at PO8 for both orientations (upright:  $F(8,21) = 4.29, p < .01$ ; inverted:  $F(8,21) = 34.75, p < .01$ ). ERPs elicited by inverted, but not upright

Groups	Electrode	Orientation	Stimuli	Mean Amplitude	Std. Deviation	Mean Latency	Std. Deviation	
Western Caucasian Observers								
	<b>PO8</b>	<i>upright</i>	Western Caucasian	−5.20	4.77	159	10	
			African American	−5.37	4.69	159	14	
			East Asian	−5.06	4.25	158	11	
		<i>Inverted</i>	Western Caucasian	−7.56	5.20	165	11	
			African American	−6.40	4.89	164	15	
			East Asian	−6.90	5.16	165	13	
		<b>PO7</b>	<i>upright</i>	Western Caucasian	−4.45	3.81	160	12
				African American	−4.21	3.83	162	14
				East Asian	−4.42	3.54	162	13
			<i>Inverted</i>	Western Caucasian	−5.57	4.12	165	9
				African American	−4.81	4.29	164	11
				East Asian	−5.33	4.00	165	9
East Asian Observers								
	<b>PO8</b>	<i>upright</i>	Western Caucasian	−5.91	3.41	160	9	
			African American	−5.73	3.39	161	9	
			East Asian	−5.57	2.99	161	12	
		<i>Inverted</i>	Western Caucasian	−6.90	3.53	166	11	
			African American	−6.61	3.67	167	11	
			East Asian	−7.40	3.58	165	10	
		<b>PO7</b>	<i>upright</i>	Western Caucasian	−4.35	3.45	160	10
				African American	−4.30	3.59	161	9
				East Asian	−3.78	3.34	161	10
			<i>Inverted</i>	Western Caucasian	−4.33	4.40	165	9
				African American	−4.34	3.82	164	10
				East Asian	−4.43	4.42	165	10

Table 2. N170.

faces were also found to be significantly larger over the right *hemisphere* (upright:  $F(1,28) = 3.51$ ;  $p > .05$ ; inverted:  $F(1,28) = 8.72$ ;  $p < .01$ ). Finally, for inverted, but not upright, we observed a main effect of stimulus race (upright:  $F(2,27) = 2.38$ ,  $p < .05$ ; inverted:  $F(2,27) = 8.10$ ,  $p < .01$ ). Bonferroni corrected  $t$ -tests showed that inverted AA faces elicited the smallest N170 amplitude ( $p < .05$ ). This effect, as shown by the *stimuli*  $\times$  *hemisphere* interaction ( $F(2,27) = 5.46$ ,  $p < .01$ ) was stronger over the right hemisphere.

In light of the ANOVAs' results, paired sample  $t$ -tests and bootstrap confidence intervals were used to estimate

effects of stimulus races on the N170 peak amplitude elicited by inverted faces. No significant amplitude differences were observed (Figures 4 and 5) among stimulus races over the left hemisphere.

Over the right hemisphere, regardless of the group of observers, SR inverted faces elicited the largest amplitude. For both groups the amplitude differences between SR and AA *inverted* faces (at PO8:  $t(14) = -3.92$ ,  $p < .01$  for WC Observers, and  $t(14) = -3.66$ ,  $p < .01$  for EA Observers) was consistently larger than those between SR and OR (at PO8:  $t(14) = -3.24$ ,  $p < .01$  for WC Observers and  $t(14) = -3.81$ ,  $p < .01$  for EA Observers). No significant differ-

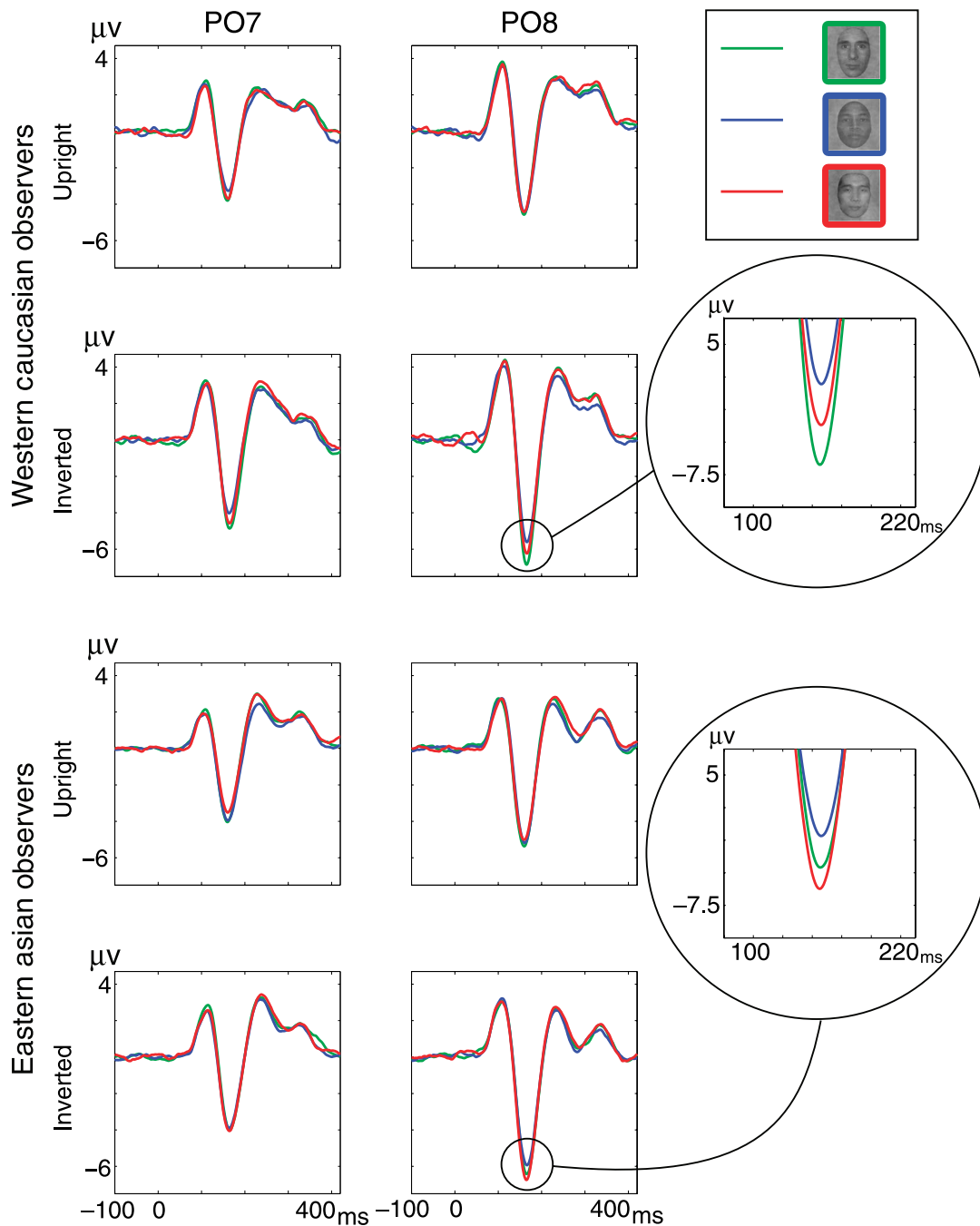


Figure 4. Grand average ERP waveforms of Western Caucasian and East Asian observers elicited by Western Caucasian (green), East Asian (red) and African American (blue) upright and inverted faces, at two occipitotemporal sites showing the largest N170 amplitude (PO7, left hemisphere; PO8, right hemisphere). While no significant differences were observed in the upright conditions, inverted faces consistently elicited larger N170 peaks for Same Race (SR) compared to Other Race (OR) faces on the PO8 electrode, with African American (AA) faces eliciting the smallest amplitudes.

ences were observed between OR and AA faces (at PO8:  $t(14) = 1.49$ ,  $p > .05$  for WC Observers, and  $t(14) = 1.61$ ;  $p > .05$  for EA Observers). Bootstrap confidence intervals showed the same patterns of results (Figure 5 displays effect size and data distribution at PO8 and PO7).

The latency of the N170 was significantly delayed by 6 ms for inverted (mean = 166 ms; std = 10 ms) compared to upright (mean = 160 ms; std = 10 ms) faces ( $F(1,28) = 18.83$ ,  $p < .01$ —see Figure 6). Importantly, there were no significant latency differences among the races of *stimuli*

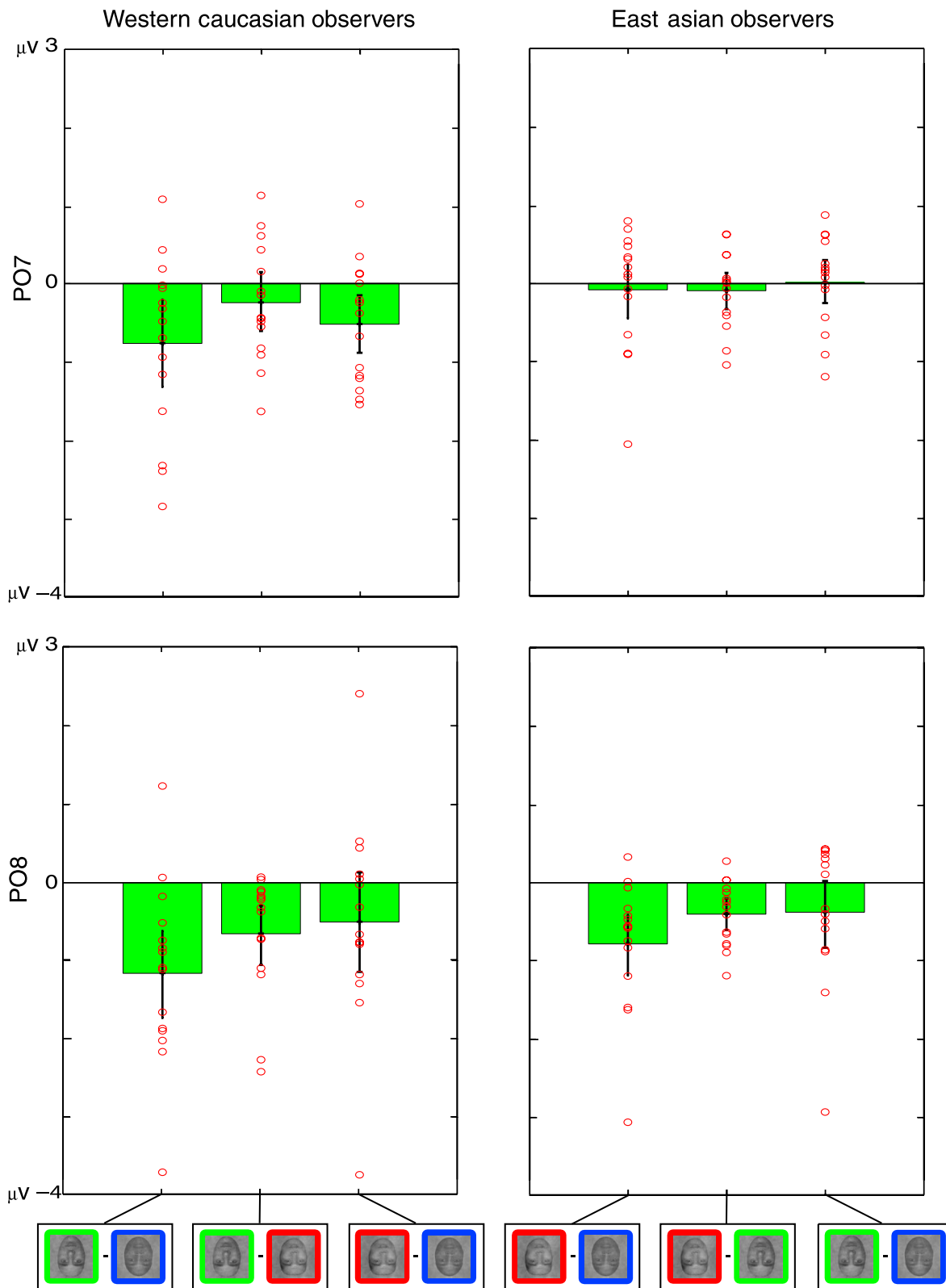


Figure 5. Bar plots of all paired differences of the N170 peak amplitudes at PO7 and PO8 for both groups of observers for the inverted condition. Red circles represent individual data points. The error bars denote the 95% bootstrapped (Wilcox, 2005) confidence interval (CI) of these differences. Statistically significant differences are conveyed by CI not including zero. Note that amplitudes differences at electrode PO8 show a consistent pattern of results, with both groups of observers responding to the law: Same-Race vs. African American > Same-Race vs. Other-Race > Other-Race vs. African Americans. The difference between Other-Race vs. African Americans is non-significant.

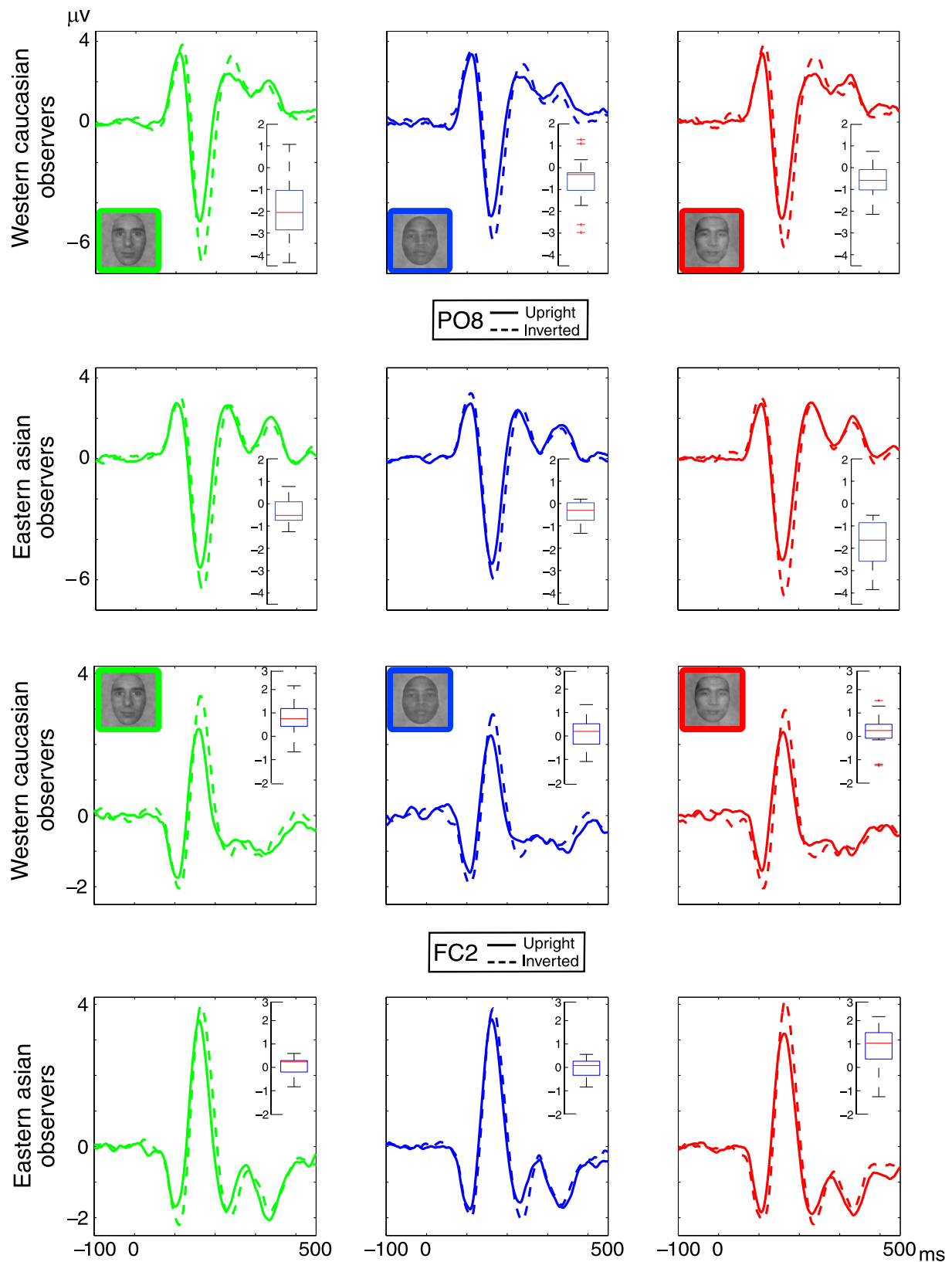


Figure 6. Grand average ERPs at PO8 and FC2 for the upright and inverted (dotted) conditions, electrodes showing respectively the largest amplitudes for the N170 and VPP components. Note that the maximal difference between Inverted and Upright faces occurs at the N170 latency for both the occipitoparietal and frontal site. Both groups of observers showed largest amplitudes for Same-Race > Other-Race > African Americans. The box plots represent the distribution of individual around the median (red line) of the amplitude difference between inverted and upright N170s and VPPs; red crosses indicate outliers.

( $F(5,24) = .33, p > .05$ ); and no significant interaction between the race of the stimuli and the groups of observers in either orientation ( $stimuli \times groups—F(2,27) = .07; p > .05$ ;  $stimuli \times groups \times orientation—F(2,27) = 1.2; p > .05$ ). No other significant latency differences were observed on this component.

## P2

P2 results are reported in Table 3. The absence of significant  $groups \times stimuli$  ( $F(2,27) = .33; p > .05$ ) and  $groups \times stimuli \times orientations$  ( $F(2,27) = .36; p > .05$ ) interactions on P2 amplitude indicated that stimulus race did not systematically modulate the amplitude of this component across groups in either orientation. Moreover, a main effect of stimulus race was observed on amplitude ( $stimuli—F(2,27) = 4.13; p > .05$ ). Pairwise comparisons revealed only one significant differences between EA and AA faces, which was true for upright and inverted faces (Bonferroni corrected  $t$ -tests,  $p < .05$ ). In addition, P2 amplitude was modulated by face orientation ( $F(1,28) = 8.48; p < .05—Figure 3$ ). Inverted faces elicited significantly larger amplitudes (mean =  $3.03 \mu V$ ; std =  $1.91 \mu V$ ) compared to upright faces (mean =  $2.74 \mu V$ ; std =  $1.75 \mu V$ ). However, this effect was observed only over the left hemisphere. ( $orientation \times hemisphere—F(1,28) = 5.08; p < .05$ ).

The latency of P2 elicited by inverted faces (mean = 235.5 ms; std = 14.85 ms) was significantly delayed by 5.8 ms compared to that elicited by upright faces (mean = 226.5 ms; std = 16.9 ms;  $F(1,28) = 22.96; p < .01$ ). Importantly, the race of the face stimuli did not modulate the latency of this component in either orientation ( $stimuli \times groups—F(2,27) = .55; p > .05$ ;  $stimuli \times groups \times orientation—F(2,27) = 1.14; p > .05$ ). No other significant differences were observed.

## N250

Table 4 shows the N250 results. Because all the electrodes of interest showed the same pattern of results, we only report the values for PO8 and PO7, those sites where the N250 was largest. We observed a main effect of stimulus race ( $F(2,27) = 3.87; p < .05$ ) on the N250 amplitude. However, Bonferroni corrected  $t$ -tests showed only a significant amplitude difference between EA and AA faces for the 2 orientations ( $p < .05$ ). Furthermore, unlike all the previous components, the N250 showed no significant amplitude modulations by face orientation ( $orientation—F(1,28) = .93; p > .05$ ). There was also no significant  $stimuli \times groups$  ( $F(2,27) = .09; p > .05$ ) and  $stimuli \times groups \times orientation$  ( $F(2,27) = 1.91; p > .05$ ) interaction.

Similarly to the N250 amplitude, the N250 latency showed no sensitivity to orientation ( $F(1,28) = .21; p > .05$ ) and to race ( $F(2,27) = 3.23; p > .05$ ). Race effects did

not interact with groups ( $stimuli \times groups—F(2,27) = .64; p > .05$ ) in either orientation ( $stimuli \times groups \times orientation—F(2,27) = 1.5; p > .05$ ). No other significant differences were observed for this component.

## Component free analyses

Figure 6 shows the N170 and the VPP.

The 2 way ANOVA ( $group \times stimuli$ ) carried out independently at each electrode and at each time point on the FIE index revealed significant interactions only in the N170 time window ( $p < .05$ ).

The scalp topography in Figure 7 shows the significant  $F$  values ( $p < .05$ ) for the  $group \times stimuli$  interaction. SR faces elicited the largest electrophysiological FIE at the N170 peak latency, regardless of the group of the observer. Significant effects were observed over four clusters of electrodes: left occipital, right occipital, left frontotemporal and frontal sites. Because all the electrodes within each cluster showed the same pattern of results, for display purposes, we only selected the electrodes with the largest amplitude in each group.

Bootstrap tests revealed a stronger FIE for SR compared to AA faces in the clusters of electrodes around PO8 and FC2 (Figure 7; PO8: mean difference = .82, CI = [.41, 1.18] for WC; mean difference = .65 [.10, 1.18] for EA; FC2: mean difference =  $-.55$  [ $-.20, -.90$ ] for WC; mean difference =  $-.20$  [ $-.02, -.45$ ] for EA). Similarly, the FIE was stronger in SR compared to OR faces (PO8: mean difference = .63 [.10, 1.11] for WC; mean difference = .31 [.05, .56] for EA; FC2: mean difference =  $-.30$  [ $-.10, -.50$ ] for WC; mean difference =  $-.10$  [ $-.05, -.30$ ] for EA). No significant differences were observed between OR and AA faces (PO8: mean difference = .27 [ $-.32, .67$ ] for WC; mean difference = .32 [ $-.15, .83$ ] for EA; FC2: mean difference =  $-.25$  [.05,  $-.54$ ] for WC; mean difference =  $-.10$  [.10,  $-.28$ ] for EA). Although the left occipital and frontotemporal clusters showed respectively the same patterns as the right occipital and left frontal clusters, pairwise comparisons did not reveal any significant differences (Figure 7).

## EEG and behavioral results

As shown in Figure 8, in WC observers there was a significant correlation between the N170 amplitude and the behavioral FIE for SR ( $r(8) = .82, p < .01$ ), but not for OR ( $r(8) = .56, p > .05$ ) and AA ( $r(8) = .23, p > .05$ ) faces. Moreover, while the same pattern of results was observed for EA participants, the correlation detected for SR faces was non significant ( $r(8) = .63, p = .06$ ), even though all subjects but one showed the effect; correlations for AA ( $r(8) =  $-.19, p = .62$ ) and OR faces ( $r(8) =  $-.50, p > .17—Figure 8$ ) were also non significant. A more robust percentile bootstrap test of independence showed significant$$

Groups	Electrode	Orientation	Stimuli	Mean Amplitude	Std. Deviation	Mean Latency	Std. Deviation	
Western Caucasian Observers	<b>O2</b>	<i>upright</i>	Western Caucasian	2.86	1.75	222	20	
			African American	2.71	1.88	223	22	
			East Asian	2.87	1.75	224	22	
		<i>Inverted</i>	Western Caucasian	3.23	2.18	238	25	
			African American	3.03	2.26	234	20	
			East Asian	3.15	2.09	229	18	
	<b>O1</b>	<i>upright</i>	Western Caucasian	2.70	1.88	220	23	
			African American	2.39	1.91	222	22	
			East Asian	2.47	1.68	228	32	
		<i>Inverted</i>	Western Caucasian	3.06	1.98	244	28	
			African American	2.96	1.87	238	19	
			East Asian	3.32	1.94	229	17	
	East Asian Observers	<b>O2</b>	<i>upright</i>	Western Caucasian	2.85	2.46	227	16
				African American	2.63	2.48	224	16
				East Asian	2.93	2.25	233	19
<i>Inverted</i>			Western Caucasian	2.72	2.30	236	17	
			African American	2.61	2.66	232	12	
			East Asian	2.93	2.29	236	19	
<b>O1</b>		<i>upright</i>	Western Caucasian	2.95	1.71	230	20	
			African American	2.57	1.48	235	16	
			East Asian	2.95	1.63	228	16	
		<i>Inverted</i>	Western Caucasian	3.06	1.81	238	20	
			African American	3.04	1.77	241	25	
			East Asian	3.19	1.78	233	20	

Table 3. P2.

correlations ( $p < .01$ ) between behavioral and electrophysiological FIE indexes, for both groups of observers (WC – CI = [0.28, 0.98]; EA – CI = [0.35, 0.94]).

## Discussion

We investigated whether the early ERP face-preferential N170 component (and its positive counterpart, the VPP)

was modulated by the race and orientation of faces as a function of the race of the observers.

Consistent with previous behavioral studies (e.g., Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Michel, Rossion et al., 2006; Rhodes et al., 1989; Tanaka et al., 2004; Walker & Tanaka, 2003), Western Caucasian (WC) and East Asian (EA) observers were more accurate at recognizing same-race (SR) compared to other-race (OR) faces. Importantly, in accordance with past reports (Buckhout & Regan, 1988; McKone et al., 2007; Murray et al., 2003; Rhodes et al., 1989) face inversion had

Groups	Electrode	Orientation	Stimuli	Mean Amplitude	Std. Deviation	Mean Latency	Std. Deviation
Western Caucasian Observers	<b>PO8</b>	<i>upright</i>	Western Caucasian	1.41	1.86	284	20
			African American	1.34	1.75	277	23
			East Asian	1.61	1.78	282	18
		<i>Inverted</i>	Western Caucasian	1.81	2.12	281	12
			African American	1.51	2.34	278	24
			East Asian	1.67	2.11	285	15
	<b>PO7</b>	<i>upright</i>	Western Caucasian	0.77	2.06	286	24
			African American	0.79	2.30	291	24
			East Asian	0.65	1.77	288	22
		<i>Inverted</i>	Western Caucasian	0.69	1.94	281	24
			African American	0.45	2.02	287	25
			East Asian	0.88	2.29	285	23
East Asian Observers	<b>PO8</b>	<i>upright</i>	Western Caucasian	0.04	1.56	280	15
			African American	−0.10	1.49	284	19
			East Asian	0.20	1.40	287	16
		<i>Inverted</i>	Western Caucasian	−0.11	1.38	288	18
			African American	−0.13	1.37	279	22
			East Asian	0.04	1.44	287	15
	<b>PO7</b>	<i>upright</i>	Western Caucasian	0.99	2.11	287	23
			African American	0.73	2.20	287	20
			East Asian	1.02	1.57	284	21
		<i>Inverted</i>	Western Caucasian	1.19	1.83	284	23
			African American	1.22	1.99	278	15
			East Asian	1.41	1.88	289	15

Table 4. N250.

greater detrimental effects on the recognition of SR compared to OR and African-American (AA) faces in both groups of observers. Moreover, in line with previous electrophysiological findings (Bentin et al., 1996; Itier & Taylor, 2004a; Jeffreys, 1996; Rossion et al., 1999, 2000; Rossion & Jacques, 2008), inversion increased and delayed the N170 and the VPP. Importantly, while the race of the faces did not affect the latencies of these components in either orientation, we observed a *modulation the magnitude of the face inversion effect (FIE)* as a function of race across groups of observers—this is the main novel finding of our study. Specifically, the

magnitude of the FIE (i.e., the amplitude difference between upright and inverted faces) was largest for SR, compared to OR and AA faces, with no differences between the latter two categories. This effect cannot be accounted for by low-level visual differences among faces from different races for three reasons. First, amplitude spectrum, luminance, and contrast were equated across all the stimuli; second, we observed a full crossover interaction between groups showing that, *regardless of the groups of observers*, SR faces elicited larger FIE compared to OR and AA faces; third, our statistical analyses showed no significant amplitude differences between the



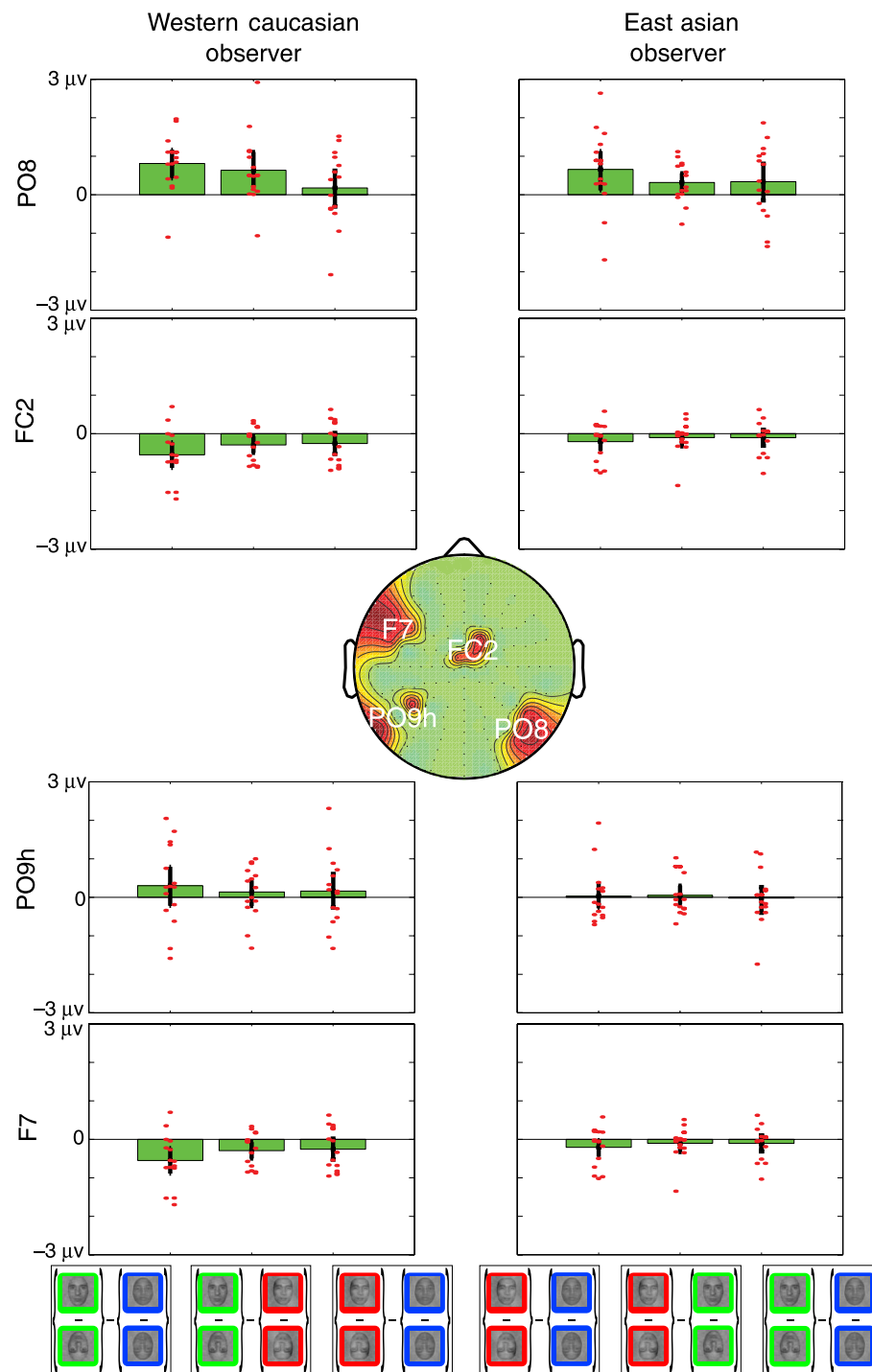


Figure 7. *Center*: Scalp topography at the N170 latency of significant  $F$  values for the ANOVA computed on the differences between Upright minus Inverted faces on all electrodes, across all time points and all conditions. The red clusters highlight significant  $F$  values distributions ( $p < .05$ ). Note that high significant  $F$  values clustered around the N170 latency at four sites (i.e. left and right occipitoparietal, frontal and right frontotemporal). *Top and bottom*: Bar plots reporting the results of the post-hoc tests computed on the pair-wise differences (i.e. Same-Race vs. African American, Same-Race vs. Other-Race, Other-Race vs. African Americans) for inverted vs. upright conditions at the time window where significant  $F$  values clustered (i.e. N170 window). Within each cluster all the electrodes show a similar pattern of results. We report here only the electrodes showing the largest differences per cluster. The red dots represent single data points. The error bars show the 95% bootstrapped (Wilcox, 2005) confidence interval (CI) of these differences. Statistically significant differences are conveyed by CI *not* including zero. Both groups of observers on these clusters respond to the law: Same-Race vs. African American > Same-Race vs. Other-Race > Other-Race vs. African Americans. The difference between Other-Race vs. African Americans is always non-significant.

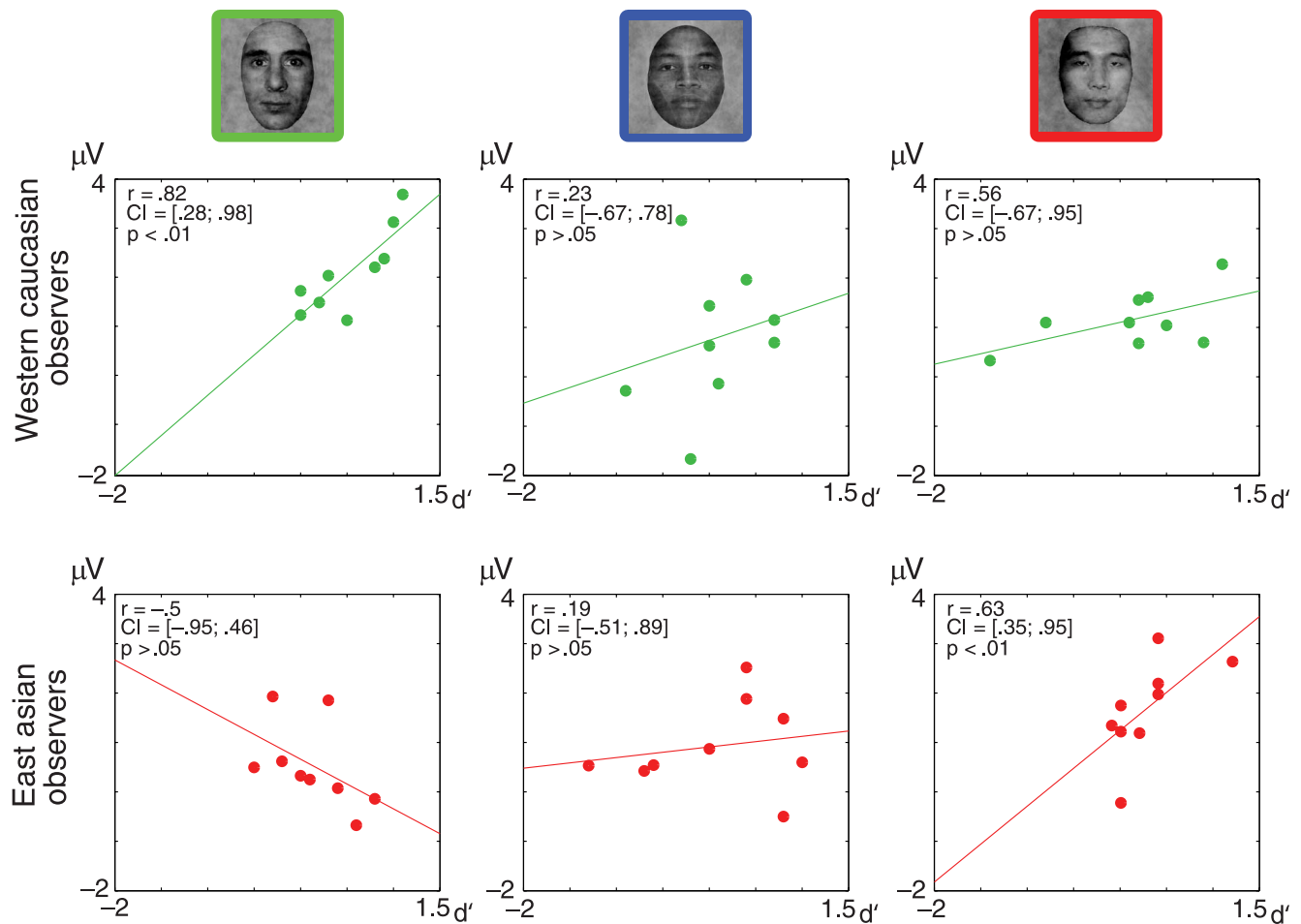


Figure 8. Correlations between the magnitude of the FIE on the N170 (i.e. N170 amplitude elicited by Inverted minus those elicited by upright faces per each condition) and the magnitude of the FIE on the  $d'$  scores. Pearson coefficient, bootstrap confidence interval and respective  $p$  values are reported in the top left corner.

two respective OR categories. These observations strengthen the claim that the modulations observed on N170 and VPP components are genuinely related to race.

Interestingly, while both P1 and P2 were sensitive to stimulus orientation, exhibiting larger amplitudes and delayed latencies for inverted compared to upright faces, the N250 was not modulated by this factor. Importantly, the P1, P2 and N250 components were not sensitive to race, as demonstrated by the absence of an interaction between the race of the faces and the groups of observers.

The electrophysiological results were consistent with the behavioral data. Both groups of observers showed a significant positive correlation between the magnitude of the FIE on the N170 and recognition accuracy only for SR faces. In line with previous observations (Jacques et al., 2007), these positive correlations indicate that larger N170 amplitudes elicited by inverted faces are associated with lower recognition accuracy.

It is worth noting however that these correlations are driven by different components of the inversion equation.

The behavioral FIE is due to the superior recognition of SR faces in their *upright* orientation. In contrast, the electrophysiological FIE is due to the enhanced N170 negativity elicited by SR faces in their *inverted* orientation. However, in our study the electrophysiological (i.e., orthogonal) and the behavioral (i.e., active) findings relied on different task constraints. Therefore, future studies are necessary to clarify the very nature of this paradox and precisely identify the mechanisms underlying the relationship between these measures.

The analyses of the conjoined effects of face inversion and race on the ERP signal were underlined by a rigorous and comprehensive statistical approach. We carried out both canonical statistical tests (i.e.  $t$ -tests) as well as robust statistics (i.e. bootstrap confidence interval) on behavioral data. From a methodological point of view, the importance of implementing such a statistical approach on our data is illustrated by the behavioral results. Canonical statistical tests (i.e.  $t$ -tests) did not capture significant effects (i.e.,  $p = .06$ ) for the FIE for SR

faces in EA observers, and the correlation between this index with the electrophysiological data. However, as clearly shown by the standard errors in Figure 2, the absence of significant effects seems to arise from the large variance for the inverted orientation (with all participants but one showing the effects). The power of parametric tests is significantly affected by both the limited sample size and the asymmetrical distribution of the population. By using bootstrap confidence intervals, which are less sensitive to sample size and outliers, both statistics reached significance.

In the face processing literature, it has been suggested that the recognition impairment observed for inverted faces reflects a *qualitative* switch from holistic to featural processing, or at least an impairment of holistic processing (e.g. Rossion & Gauthier, 2002). Undeniably, due to the lack of experience with inverted faces, these stimuli are more difficult to process than those presented in their canonical (upright) orientation. Concurrently, the amplitude enhancement on the N170 and VPP components could be related to the relative processing impairment associated with inverted faces (e.g. Rossion et al., 2000). It has also been argued that SR faces are processed more holistically than OR faces (e.g. Michel, Caldara et al., 2006; Michel, Rossion et al., 2006). Thus, if face inversion triggers a switch from holistic to featural processing mode, it should lead to a greater impairment for the category processed more holistically: SR faces.

Importantly, regardless of these theoretical interpretations, our data show that the relatively larger behavioral recognition impairment for SR faces due to stimulus inversion is associated with electrophysiological modulations, with both measures correlating positively. The sensitivity of the N170 to race during face inversion suggests that differences in processing SR and OR faces begin *early*, at the perceptual level. These early perceptual differences fit well with ORE theoretical frameworks. In the face space model proposed by Valentine and Endo (1992), which recently found objective support in neural network simulations (Caldara & Abdi, 2006), faces are encoded in an arbitrary psychological multidimensional space. Variations in exemplar density between races account for SR face recognition advantage, because SR face representations are more widely distributed in the face space than the representations of OR faces. The physical distance between points, symbolizing psychological representations of individual faces, is a function of the *perceived* difference between faces, which determines the density of the cloud. It would therefore make sense that differences in processing SR and OR faces arise at the *perceptual level*. Unlike Valentine's multidimensional face space model, Levin (1996, 2000) postulates that for OR faces, race is extracted as a feature at the cost of individuating information. Therefore, according to Levin, performance for SR and OR faces can be accounted for by differences in coding features for faces of different races

(Levin, 2000). This account would also be consistent with our data showing modulations of face processing as a function of race at early perceptual stages.

In keeping with several previous studies (e.g. Caldara et al., 2003, 2004; Tanaka & Pierce, 2009), race did not affect the N170 amplitude for upright faces. However, we failed to replicate the delayed N170 latency for inverted OR compared to SR faces (Wiese et al., 2009), as well as previous modulations of race on the N170 amplitude for upright faces (Herrmann et al., 2007; Ito & Urland, 2005; Stahl et al., 2008; Walker et al., 2008). We also failed to replicate race effects at earlier latencies than that of the N170. Ito and Urland (2003) reported race modulations between AA and WC faces in WC observers on early ERP components (i.e. 100 ms after stimulus onset). In their experiment, however they used color pictures of faces of different races that differed markedly in skin color and were not equated in spatial frequency content. Despite these physical differences, Ito and Urland (2003) interpreted their findings as evidence for early race categorization. Alternatively, P1 effects might be due to differences in global low-level visual properties, which are known to modulate early ERP components (e.g. Hillyard, Teder-Salejarvi, & Munte, 1998; Johannes, Munte, Heinze, & Mangun, 1995; Luck, 2005), even independently of attention (Hillyard et al., 1998). In our study, we equated some important global low-level visual properties across stimuli, and we failed to observe any P1 categorical race effect. In line with previous studies (e.g. Itier & Taylor, 2004a, 2004b), we only observed a general amplitude and latency increase on the P1 component for inverted compared to upright faces. Indeed, amplitude modulations on this component have been associated with low-level differences in the stimuli (Debruille, Guillem, & Renault, 1998; Halit, de Haan, & Johnson, 2000). In a nutshell, despite the fact that controlling for low-level visual properties of the stimuli may reduce their ecological validity, our findings suggest that this control abolishes potentially spurious effects related to salient differences in faces from difference race. In addition, these studies relied on only one group of observers. Importantly, we show that the investigation of sensitivity to race requires the use of at least two groups of observers and the presence of a crossover interaction.

Contrary to previous findings, our data failed to reveal P2 (Caldara et al., 2004) and N250r (Tanaka & Pierce, 2009) race effects. Although controlling for global low-level visual properties might have abolished P2 and N250 race effects, it seems more likely that the lack of race effects on these components in our experiment may result from differences in task constraints. Indeed, our subjects were instructed to perform a passive orthogonal task, whereas Caldara et al. (2004) used a direct race categorization task with interspersed catch trials, and Tanaka and Pierce (2009) employed an active old-new recognition task, directly tapping into face memory. The

N250 and the P2 have been related to the active categorization of visual stimuli (Harel, Ullman, Epshtein, & Bentin, 2007; Latinus & Taylor, 2005; Philiastides, Ratcliff, & Sajda, 2006; Philiastides & Sajda, 2006), and long term memory of faces (Schweinberger, Pickering, Jentsch, Burton, & Kaufmann, 2002). If the task demands do not entail active categorization of face stimuli, or retrieval of face representations, the effect on these later components may thus be abolished. Further research is needed to clarify this issue.

An important question that would require further investigation is whether the N170 is sensitive *only* to race for *inverted* faces. How can race affect the early electrophysiological dynamics only in a given unfamiliar upside-down orientation, while producing no effects in its more canonical upright orientation? This surprising result could be explained by a ceiling effect. First, humans' proficiency in processing upright faces is so high that the ERP signal might not be sensitive enough to detect any subtle differences that may exist. However, situations that are more demanding in terms of visual processing—as it is the case for inverted faces—may render such subtleties to become more easily identifiable. Second, the tasks and paradigms used so far may be unable to confer enough power to measure early amplitude modulations to race. Our data however do not allow speculation in either direction. Further research with more sensitive paradigms (e.g., neural adaptation) and task constraints (e.g., face identification) tapping into the roots of the ORE is necessary to clarify whether early neural markers of processing upright faces (i.e. the N170) are sensitive to race, or to confirm that race sensitivity occurs uniquely for inverted faces.

## Conclusion

Given its actuality in everyday life, the thoroughly documented ORE remains an interesting topic to be studied in the field of cognitive and vision science. Although, behavioral data have shown high levels of consistency, the neural mechanisms underlying this visual effect have yet to be fully understood. Electrophysiological studies have reported highly contrasting data. In the present study we sought to investigate whether it was possible to relate modulation of the ERP signal in the early neural dynamics of face processing to behavioral differences associated with the ORE. We addressed some of the methodological inconsistencies displayed in previous research by *controlling for low level visual properties* of our stimuli, which are known to modulate the ERP signal and by adding a set of other-race faces common to both group of observers (i.e., AA faces). Since any (significant) effect could be confounded by differences in the face stimuli, which would critically impact on the generalization of the results, *we adopted 2 groups of*

*observers*: EA and WC. We observed a full cross-over interaction between the races of the observers and the race of the faces, which positively correlated with our behavioral data. These findings indisputably represent a genuine ORE on the ERP signal. Our results showed that while race did not elicit differences across the whole ERP for upright faces, the race of the faces *systematically modulated the amplitude of the N170 and VPP in the inverted condition*. The neural face system is sensitive to race at early perceptual stages of processing, *at least* when faces are presented upside-down.

## Acknowledgments

This project was supported by a British Academy grant (SG-45080) awarded to RC. RC was also supported by The Economic and Social Research Council and Medical Research Council (RES-060-25-0010).

We would also like to thank Meike Ramon for insightful comments on previous versions of this paper.

Commercial relationships: none.

Corresponding authors: Luca Vizioli; Roberto Caldara.

Emails: lucav@psy.gla.ac.uk; r.caldara@psy.gla.ac.uk.

Address: Department of Psychology, University of Glasgow, 58, Hillhead Street, Glasgow, G12 8QB, UK.

## References

- Bang, S., Kim, D., & Choi, S. (2001). *Asian face image database*. In Lab IM, editor. Korea: Postech.
- Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, 8, 551–565. [[PubMed](#)] [[Article](#)]
- Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*, 3, e3022. [[PubMed](#)] [[Article](#)]
- Botzel, K., & Grusser, O. J. (1989). Electric brain potentials evoked by pictures of faces and non-faces: A search for “face-specific” EEG-potentials. *Experimental Brain Research*, 77, 349–360. [[PubMed](#)]
- Buckhout, R., & Regan, S. G. (1988). Explorations in research on the other-race effect in face recognition in Practical Aspects of Memory: Current Research and Issues. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Memory in everyday life* (vol. 1, pp. 40–46). Chichester, Sussex, UK: John Wiley and Sons.
- Caldara, R., & Abdi, H. (2006). Simulating the ‘other-race’ effect with autoassociative neural networks: Further evidence in favor of the face-space model. *Perception*, 35, 659–670. [[PubMed](#)]

- Caldara, R., Rossion, B., Bovet, P., & Hauert, C. A. (2004). Event-related potentials and time course of the “other-race” face classification advantage. *Neuroreport*, *15*, 905–910. [PubMed]
- Caldara, R., & Seghier, M. L. (2009). The fusiform face area responds automatically to statistical regularities optimal for face categorization. *Human Brain Mapping*, *5*, 1615–1625. [PubMed]
- Caldara, R., Seghier, M. L., Rossion, B., Lazeyras, F., Michel, C., & Hauert, C. A. (2006). The fusiform face area is tuned for curvilinear patterns with more high-contrast elements in the upper part. *Neuroimage*, *31*, 313–319. [PubMed]
- Caldara, R., Thut, G., Servoir, P., Michel, C. M., Bovet, P., & Renault, B. (2003). Face versus non-face object perception and the ‘other-race’ effect: A spatio-temporal event-related potential study. *Clinical Neurophysiology*, *114*, 515–528. [PubMed]
- Carroo, A. W. (1986). Other race recognition: A comparison of black American and African subjects. *Perceptual and Motor Skills*, *62*, 135–138. [PubMed]
- Dakin, S. C., Hess, R. F., Ledgeway, T., & Achtman, R. L. (2002). What causes non-monotonic tuning of fMRI response to noisy images? *Current Biology*, *12*, 476–477. [PubMed]
- Debruille, J. B., Guillem, F., & Renault, B. (1998). ERPs and chronometry of face recognition: Following-up Seeck et al. and George et al. *Neuroreport*, *9*, 3349–3353. [PubMed]
- de Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14*, 199–209. [PubMed]
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*, 9–21. [PubMed]
- Feingold, C. A. (1914). The influence of environment on identification of persons and things. *Journal of Criminal Law and Police Science*, *5*, 39–51.
- Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configural information: Direct evidence. *Perception*, *29*, 159–170. [PubMed]
- Golby, A. J., Gabrieli, J. D., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience*, *4*, 845–850. [PubMed]
- Halit, H., de Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *Neuroreport*, *11*, 1871–1875. [PubMed]
- Harel, A., Ullman, S., Epshtein, B., & Bentin, S. (2007). Mutual information of image fragments predicts categorization in humans: Electrophysiological and behavioral evidence. *Vision Research*, *47*, 2010–2020. [PubMed]
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233. [PubMed]
- Herrmann, M. J., Schreppel, T., Jager, D., Koehler, S., Ehlis, A. C., & Fallgatter, A. J. (2007). The other-race effect for face perception: An event-related potential study. *Journal of Neural Transmission*, *114*, 951–957. [PubMed]
- Hillyard, S. A., Teder-Salejarvi, W. A., & Munte, T. F. (1998). Temporal dynamics of early perceptual processing. *Current Opinion in Neurobiology*, *8*, 202–210. [PubMed]
- Itier, R. J., & Taylor, M. J. (2004a). Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. *Neuroimage*, *21*, 1518–1532. [PubMed]
- Itier, R. J., & Taylor, M. J. (2004b). N170 or N1? Spatiotemporal differences between object and face processing using ERPs. *Cerebral Cortex*, *14*, 132–142. [PubMed]
- Ito, T. A., Thompson, E., & Cacioppo, J. T. (2004). Tracking the timecourse of social perception: The effects of racial cues on event-related brain potentials. *Personality and Social Psychology Bulletin*, *30*, 1267–1280. [PubMed]
- Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, *85*, 616–626. [PubMed]
- Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective Behavioral Neuroscience*, *5*, 21–36. [PubMed] [Article]
- Jacques, C., d’Arripe, O., & Rossion, B. (2007). The time course of the inversion effect during individual face discrimination. *Journal of Vision*, *7*(8):3, 1–9, <http://journalofvision.org/7/8/3/>, doi:10.1167/7.8.3. [PubMed] [Article]
- Jacques, C., & Rossion, B. (2006). The speed of individual face categorization. *Psychological Science*, *17*, 485–492. [PubMed]
- Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research*, *78*, 193–202. [PubMed]

- Jeffreys, D. A. (1996). Evoked potential studies of face and object processing. *Visual Cognition*, 3, 1–38.
- Johannes, S., Munte, T. F., Heinze, H. J., & Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Brain Research Cognitive Brain Research*, 2, 189–205. [PubMed]
- Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology*, 116, 2613–2631. [PubMed]
- Latinus, M., & Taylor, M. J. (2005). Holistic processing of faces: Learning effects with Mooney faces. *Journal of Cognitive Neuroscience*, 17, 1316–1327. [PubMed]
- Levin, D. T. (1996). Classifying faces by race: The structure of face categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1364–1382.
- Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General*, 129, 559–574. [PubMed]
- Luck, S. J. (2005). *An introduction to event related potential techniques*. Cambridge, Massachusetts London, England: The MIT Press.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces*. Stockholm, Sweden: Karolinska Institute.
- Malpass, R. S., & Kravitz, J. (1969). Recognition for faces of own and other race. *Journal of Personality and Social Psychology*, 13, 330–334. [PubMed]
- McKone, E., Brewer, J. L., MacPherson, S., Rhodes, G., & Hayward, W. G. (2007). Familiar other-race faces show normal holistic processing and are robust to perceptual stress. *Perception*, 36, 224–248. [PubMed]
- Meissner, C., & Brigham, J. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, 7, 3–35.
- Michel, C., Caldara, R., & Rossion, B. (2006). Own-race faces are perceived more holistically than other-race faces. *Visual Cognition*, 14, 55–73.
- Michel, C., Rossion, B., Han, J., Chung, C. S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, 17, 608–615. [PubMed]
- Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods Instruments and Computers*, 36, 630–633. [PubMed]
- Murray, J., Rhodes, G., & Schuchinsky, G. (2003). When is a face not a face? The effects of misorientation on mechanisms of face perception In M. Peterson & G. Rhodes (Eds.), *Perception of faces, objects, and scenes: Analytic and holistic processes* (pp. 75–91). New York: Oxford University Press.
- Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, 112, 713–719. [PubMed]
- Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: A timing diagram. *Journal of Neuroscience*, 26, 8965–8975. [PubMed] [Article]
- Philiastides, M. G., & Sajda, P. (2006). Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cerebral Cortex*, 16, 509–518. [PubMed] [Article]
- Rhodes, G., Tan, S., Brake, S., & Taylor, K. (1989). Expertise and configural coding in face recognition. *British Journal of Psychology*, 80, 313–331. [PubMed]
- Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica*, 128, 274–289. [PubMed]
- Rossion, B., Delvenne, J. F., Debatisse, D., Goffaux, V., Bruyer, R., Crommelinck, M., et al. (1999). Spatio-temporal localization of the face inversion effect: An event-related potentials study. *Biological Psychology*, 50, 173–189. [PubMed]
- Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces? *Behavioral and Cognitive Neuroscience Reviews*, 1, 63–75. [PubMed]
- Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *Neuroreport*, 11, 69–74. [PubMed]
- Rossion, B., & Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *Neuroimage*, 39, 1959–1979. [PubMed]
- Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage*, 20, 1609–1624. [PubMed]
- Rousselet, G. A., Pernet, C. R., Bennett, P. J., & Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neuroscience*, 9, 98. [PubMed] [Article]
- Schweinberger, S. R., Pickering, E. C., Jentsch, I., Burton, A. M., & Kaufmann, J. M. (2002). Event-related brain potential evidence for a response of inferior temporal cortex to familiar face repetitions. *Brain Research Cognitive Brain Research*, 14, 398–409. [PubMed]

- Schyns, P. G., Petro, L. S., & Smith, M. L. (2007). Dynamics of visual information integration in the brain for categorizing facial expressions. *Current Biology*, *17*, 1580–1585. [[PubMed](#)]
- Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *Neuroimage*, *37*, 974–984. [[PubMed](#)]
- Stahl, J., Wiese, H., & Schweinberger, S. R. (2008). Expertise and own-race bias in face processing: An event-related potential study. *Neuroreport*, *19*, 583–587. [[PubMed](#)]
- Swets, J., Tanner, W. P., Jr., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, *68*, 301–340. [[PubMed](#)]
- Tanaka, J. W., & Curran, T. (2001). A neural basis for expert object recognition. *Psychological Science*, *12*, 43–47. [[PubMed](#)]
- Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: Evidence from a cross-cultural study. *Cognition*, *93*, B1–B9. [[PubMed](#)]
- Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive Affective and Behavioural Neuroscience*, *9*, 122–131. [[PubMed](#)]
- Valentine, T., & Bruce, V. (1986). The effect of race, inversion and encoding activity upon face recognition. *Acta Psychologica*, *61*, 259–273. [[PubMed](#)]
- Valentine, T., & Endo, M. (1992). Towards an exemplar model of face processing: The effects of race and distinctiveness. *Quarterly Journal of Experimental Psychology A, Human Experimental Psychology*, *44*, 671–703. [[PubMed](#)]
- Walker, P. M., Silvert, L., Hewstone, M., & Nobre, A. C. (2008). Social contact and other-race face processing in the human brain. *Social Cognitive and Affective Neuroscience*, *3*, 16–25. [[PubMed](#)] [[Article](#)]
- Walker, P. M., & Tanaka, J. W. (2003). An encoding advantage for own-race versus other-race faces. *Perception*, *32*, 1117–1125. [[PubMed](#)]
- Wiese, H., Stahl, J., & Schweinberger, S. R. (2009). Configural processing of other-race faces is delayed but not decreased. *Biological Psychology*, *81*, 103–109. [[PubMed](#)]
- Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing (2nd ed.)*. San Diego: Academic Press.
- Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology: Human Perception and Performance*, *81*, 141–145.