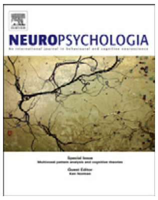
Neuropsychologia 50 (2012) 3461–3468

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia



Spatial attention modulates early face processing

Wenfeng Feng [a,*], Antigona Martinez [a,b], Michael Pitts [c], Yue-Jia Luo [d], Steven A. Hillyard [a]

[a] Department of Neurosciences, University of California, San Diego, 9500 Gilman Dr, La Jalla, CA 92093-0608, USA
[b] Nathan Kline Institute for Psychiatric Research, Schizophrenia Research Division Orangeburg, NY 10962, USA
[c] Department of Psychology, Reed College, Portland, OR 97202, USA
[d] National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China

ARTICLE INFO

Article history:
Received 23 March 2012
Received in revised form
13 September 2012
Accepted 17 September 2012
Available online 25 September 2012

Keywords:
Spatial attention
ERP
P1
N170
Face processing

ABSTRACT

It is widely reported that inverting a face dramatically affects its recognition. Previous studies have shown that face inversion increases the amplitude and delays the latency of the face-specific N170 component of the event-related potential (ERP) and also enhances the amplitude of the occipital P1 component (latency 100–132 ms). The present study investigates whether these effects of face inversion can be modulated by visual spatial attention. Participants viewed two streams of visual stimuli, one to the left and one to the right of fixation. One stream consisted of a sequence of alphanumeric characters at 6.67 Hz, and the other stream consisted of a series of upright and inverted images of faces and houses presented in randomized order. The participants' task was to attend selectively to one or the other of the streams (during different blocks) in order to detect infrequent target stimuli. ERPs elicited by inverted faces showed larger P1 amplitudes compared to upright faces, but only when the faces were attended. In contrast, the N170 amplitude was larger to inverted than to upright faces only when the faces were not attended. The N170 peak latency was delayed to inverted faces regardless of attention condition. These inversion effects were face specific, as similar effects were absent for houses. These results suggest that early stages of face-specific processing can be enhanced by attention, but when faces are not attended the onset of face-specific processing is delayed until the latency range of the N170.

1. Introduction

Facial recognition is one of the most important skills for human interaction and has been the focus of neuropsychological studies for decades. Three areas of the human brain have been identified as face-selective: the fusiform gyrus, known as the fusiform face area (FFA) (Kanwisher, McDermott, and Chun, 1997; McCarthy, Puce, Gore, and Allison, 1997; Sergent, Ohta, and MacDonald, 1992); the lateral inferior occipital gyri, known as the occipital face area (OFA) (Gauthier et al., 2000); and the superior temporal sulcus (Allison, Puce, and McCarthy, 2000; Hoffman & Haxby, 2000). Electroencephalographic (EEG) and magnetoencephalographic (MEG) studies aimed at measuring the timing of activation in these areas have consistently reported a prominent face-selective component (N170/M170), which typically peaks at around 170 ms after stimulus onset and is larger for faces than for non-face stimuli (houses, cars, scrambled faces, etc.) (e.g., Bentin, Allison, Puce, Perez, and McCarthy, 1996; Lu et al., 1991; Sams, Hietanen, Hari, Ilmoniemi, and Lounasmaa, 1997). Consistent with functional magnetic resonance imaging (fMRI) studies, event-related potential (ERP) and MEG studies have identified neuronal generators of the face-specific N170/M170 component in the fusiform gyrus and in the superior temporal sulcus (STS) (Liu, Harris, and Kanwisher, 2002; Itier & Taylor, 2004a).

There is evidence that an even earlier ERP/MEG component (P1/M100) may be increased in amplitude and latency for face stimuli, suggesting that face-specific neural processing is initiated earlier than the N170/M170. For example, Liu and colleagues (2002) found that the M100 component was enhanced by the successful detection of faces embedded in noise, and Debruille, Guillem, and Renault (1998) observed differential responses between known and unknown faces beginning at around 100 ms post stimuli onset.

Studies of inverted face processing have found that when faces were presented upside down their recognition was impaired to a greater extent than for other objects (Yin, 1969), a phenomenon known as the face inversion effect (FIE). Several MEG and ERP studies have found P1/M100 to be increased for inverted faces compared to upright faces (Itier & Taylor, 2002, 2004a, 2004b; Itier, Herdman, George, Cheyne, and Taylor, 2006; Linkenkaer-Hansen et al., 1998; Susac, Ilmoniemi, Pihko, and Supek, 2004; but

---

* Corresponding author. Tel.: +1 858 534 1566.
E-mail address: fengwfly@gmail.com (W. Feng).

0028-3932/$ - see front matter Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.neuropsychologia.2012.09.031

see Rossion et al., 1999). These early face processing effects imply that at least some aspects of first-order face processing, such as defining a stimulus as a face via detection of its component features (Maurer, Le Grand, & Mondloch, 2002; Diamond & Carey, 1986), may take place prior to the N170/M170 response. Studies attempting to localize the neural sources of the P1/M100 elicited by faces have produced inconsistent results; some reports have suggested a source in the medial occipital cortex (Halgren, Raij, Marinkovic, Jousmaki, & Hari, 2000; Itier et al., 2006), while others have indicated sources in posterior extrastriate cortex (Linkenkaer-Hansen et al., 1998) or in higher level visual cortex of the mid-fusiform gyrus (Herrmann, Ehlis, Muehlberger, & Fallgatter, 2005). Previous studies have also found that face inversion delayed the latency of the N170 component (Bentin et al., 1996; Rossion et al., 1999, 2000; Rossion, Joyce, Cottrell, & Tarr, 2003; Itier & Taylor, 2002, 2004a, 2004b; Itier et al., 2006), and most but not all these studies also found enlarged N170 amplitudes to inverted faces compared to upright faces (Rossion et al., 1999, 2000; Itier & Taylor, 2002, 2004a, 2004b; Itier et al., 2006). These results suggest that inversion of a face may modify its processing over multiple time ranges.

Face processing has been considered to occur automatically or pre-attentively on the basis of ERP and MEG studies that reported little or no modulation of the N170/M170 by various manipulations of attention (Cauquil, Edmonds, & Taylor, 2000; Carmel & Bentin, 2002; Holmes, Vuilleumier, & Eimer, 2003; Downing, Liu, & Kanwisher, 2001; Furey et al., 2006). Recently, however, Crist, Wu, Karp, and Woldorff (2008) reported that the N170 component could indeed be modulated by the allocation of spatial attention; specifically, the authors found no significant N170 amplitude difference between faces and houses when they were unattended, but N170 amplitude was significantly larger for faces than houses when these stimuli were attended. These results suggest that the face processing as indexed by the N170 component is not invariably automatic but may be modulated by spatial attention.

Previous studies have reported that the early P1/M100 component can be modulated by factors such as facial configuration (Halit, de Haan, & Johnson, 2000), familiarity (Debruille et al., 1998), and emotional expression (Pourtois, Grandjean, Sander, & Vuilleumier, 2004; Pourtois, Thut, Grave de Peralta, Michel, & Vuilleumier, 2005). However, it remains unknown whether the early face processing revealed by the P1/M100 can be modulated by the allocation of spatial attention. The present experiment aimed to investigate the effects of spatial attention on face-specific processing in the latency range of the P1/M100 and N170 components and to localize the sources of neural activity modulated by attention. The ongoing EEG was recorded while two streams of visual stimuli were presented to the participants, a steady-state alphanumeric stimuli stream to one visual field and a face/house image stream to the opposite field. ERPs were recorded separately to upright and inverted faces and houses, while steady-state visual evoked potentials (SSVEPs) were concurrently recorded in response to the alphanumeric stream. On separate runs spatial attention was directed to either the face/house stream or the alphanumeric stream. Previous studies have shown that SSVEP amplitudes are enhanced by the allocation spatial attention (e.g., Morgan, Hansen, & Hillyard, 1996; Müller & Hillyard, 2000; Mishra, Zinni, Bavelier, & Hillyard, 2011). Thus, modulations of SSVEP amplitudes here will demonstrate the effectiveness of the attention manipulation. The major question to be investigated in the present study is whether the effects of facial inversion on the P1 and N170 components can be modulated by spatial attention. Such modulations would indicate an influence of attention on the early processing of facial configurations.

2. Method

2.1. Participants

Thirteen right-handed subjects (five male, mean age 22 years) participated in this experiment after giving written informed consent. All subjects were recruited as volunteers and had normal or corrected-to-normal vision. The experimental procedures were approved by the Institutional Review Board of the University of California, San Diego, in compliance with the Declaration of Helsinki.

2.2. Experiment design

Two streams of stimuli were presented concurrently, one in the left and one in the right visual field. One stream consisted of a rapid serial visual presentation (RSVP) of alphanumeric characters ($2^\circ \times 2^\circ$ visual angle) that appeared 5° laterally to a central fixation cross (see Fig. 1). This text stream consisted of non-target letters (96% of stimuli) and target numbers (4%). Each alphanumeric character appeared for 150 ms before being replaced by a subsequent character; thus the rate of stimulation was 6.67/s. The other stream of stimuli consisted of images of faces (from Nimstim Face Stimulus Set, Tottenham et al., 2009) and houses (both $5^\circ \times 6^\circ$), which were presented in random order at an eccentricity of 4° from the center of fixation to the edge of the image. The face image set consisted of 35 faces (18 female), and the house image set consisted of 35 houses. The same face and house images were turned upside down to create matching inverted image sets. The stimuli onset asynchronies (SOAs) of successive face/house images were randomized between 600 and 900 ms in steps of 16.66 ms; specifically, each of the equiprobable SOAs was calculated as $600 + N \times 16.66$ ms, where N was a randomly chosen integer between 0 and 18. All images were presented for 100 ms. The content of the image (face or house) was task-irrelevant, as the subject's task was to detect infrequent (20%) blurry-image targets (either faces or houses). While the face/house and alphanumeric streams differed in their SOAs, the overall target probabilities per unit time were approximately equal for the two tasks (attend blurry images or attend numbers).

During each block, subjects were asked to maintain fixation on the central cross while covertly directing attention to the instructed side (left or right) in order to detect targets in the stream presented on that side (numbers if attending to the RSVP sequence or blurry images if attending to the face/house sequence). For half of the blocks, face/house images were presented on the right side, and the alphanumeric stream on the left, and for the other half, the left/right locations were reversed. The locations of stimuli were counter balanced with respect to whether they were attended or unattended. Each subject performed 28 blocks consisting of 80 face/house images each; overall, each image was presented 16 times to each subject, eight times on the left (four attended, four unattended) and eight times on the right (four attended, four unattended).

2.3. EEG recording and analyses

Subjects sat in a dimly lit recording chamber while the EEG was recorded from 62 electrode sites (Electro-Cap International, Eaton, OH) using a modified 10–20 system montage (Di Russo, Martinez, & Hillyard, 2003). Standard 10–20 sites were FP1, FPz, FP2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, Oz, O2 and M1. Additional intermediate sites were AF3, AFz, AF4, FC5, FC3, FC1, FCz, FC2, FC4, FC6, C5, C1, C2, C6, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP5, TP8, P5, P1, P2, P6, PO7, PO3, POz, PO4, PO8, I5, I3, Iz, I4, I6, SI3, SIz and SI4 (10–10 system, Nuwer et al., 1998). The electrode impedances were kept below 5 kΩ. All signals were amplified (SA Instruments, San Diego, CA) with a gain of 10,000 and digitized at 250 Hz with an amplifier band pass of 0.01–80 Hz. The right mastoid electrode served as the online reference during data acquisition. Horizontal eye movements were

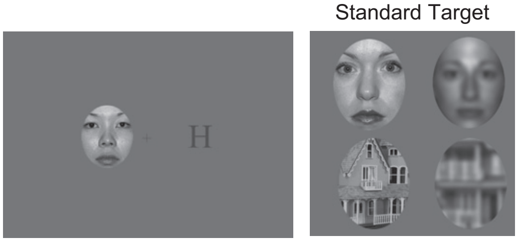


Fig. 1. Left: examples of stimuli in the condition where the face/house images were presented on the left and letters were presented on the right. The task was to detect blurry images in the face/house sequence or numbers in the alphanumeric sequence according to an instruction presented at the start of each block. The location of the face/house images was balanced on the left and right across blocks. Right: examples of standard stimuli (clear) and target stimuli (blurry) for faces and houses.

monitored via electrodes at the left and right outer canthi—the horizontal electrooculogram (HEOG). Vertical eye movements and blinks were monitored by an electrode below the left eye—the vertical electrooculogram (VEOG). To ensure that fixation was maintained and to remove trials contaminated with eye movements, an infrared eye monitor (ISCAN, Inc.) was also used throughout the recordings. Trials contaminated by eye movements, eye blinks, muscle activity, or amplifier blocking were rejected. On average, 14% of all trials were rejected due to these types of artifacts.

To ensure that the attention manipulation was effective, the steady-state visual evoked potential (SSVEP) elicited by the RSVP stream of characters was recorded and quantified. SSVEPs were first averaged in the time domain over 900 ms epochs that were time locked to each successive stimulus in the stream, separately for left RSVP stimuli and right RSVP stimuli. These epochs were initiated 300 ms after the onset of each block and terminated 300 ms before the end of each block, so the first two and the last two stimuli were not included in any epochs in order to exclude the visual evoked responses to the onset and offset of the stream. These SSVEP epochs were averaged for each subject and for the two attention conditions separately. The resulting averaged waveforms were then re-referenced algebraically to the average of the left and right mastoids. The SSVEP amplitudes were calculated from the time domain averages by Fourier transform and quantified as the absolute value of the complex Fourier coefficients. The SSVEP amplitudes were measured and averaged over a cluster of ten posterior electrode sites (PO3, PO4, PO7, PO8, O1, O2, I3, I4, I5, I6) and then tested by ANOVA with the factors of attention (attended vs. unattended), stimuli location (left, right), and hemisphere (left, right).

To analyze the ERPs elicited by the images, each scalp channel was re-referenced offline to the average of all channels, given that the face-elicited components have maximum amplitudes near the mastoids. Event-related potentials (ERPs) were time-locked to the onset of the face/house images and were later collapsed across the image locations (left and right visual field) and hemisphere of recording to obtain ERPs ipsilateral and contralateral to the side of image presentation. The ERP waveforms presented in Figs. 3 and 4 were averaged over homologous left and right electrode sites contralateral to image location (for example, PO7 when images were presented in the right visual field, and PO8 when images were presented in the left visual field). Mean amplitude of the P1 component (100–132 ms) was determined with respect to a 100 ms pre-stimulus baseline and averaged over six posterior electrode sites (P7/P8, PO7/PO8, O1/O2, I3/I4, I5/I6, SI3/SI4), separately for contralateral and ipsilateral sites. N170 amplitudes were measured as mean amplitudes between 172 and 212 ms over the same electrode sites as for the P1 component. N170 peak latencies were determined as the most negative time point in the latency range 172–212 ms.

An initial ANOVA confirmed that contralateral sites evoked larger P1 and N170 amplitudes than ipsilateral sites for both components [P1: $F(1, 12)=40.114$, $p<.0001$, N170: $F(1, 12)=17.167$, $p<.002$]. Therefore, subsequent analyses of attention effects were only carried out on the ERPs recorded over the contralateral electrode sites. To assess the effects of attention and face inversion, the face-evoked P1 (100–132 ms) and N170 (172–212) amplitudes and the N170 latencies were tested separately by ANOVAs with the factors of attention (attended, unattended) and orientation (upright, inverted). These same analyses were carried for the house stimuli as well. To allow comparisons with previous studies (for example, Crist et al., 2008), an additional ANOVA was carried out on the N170 difference amplitudes for faces minus houses. The strategy here was to first isolate face-specific processing (via the face-house subtraction) and to then assess differences based on attention and orientation.

The brain sources of the face inversion effects were estimated by dipole modeling using Brain Electrical Source Analysis (BESA version 5.1.8). The P1 (100–132 ms) difference waves (inverted minus upright faces) in the attended condition and N170 (172–212) difference waves (inverted minus upright faces) in the unattended condition were modeled in grand averaged waveforms (over all subjects) collapsed over right and left visual field stimuli to form contralateral and ipsilateral averages. All resulting solutions were transformed into the standardized coordinate system of Talairach and Tournoux (1988) and projected onto a structural brain image (MNI) using the Analysis of Functional NeuroImaging (AFNI) software (Cox, 1996).

3. Results

3.1. Behavioral results

When subjects attended to the alphanumeric stream, the mean reaction time (RT) for detecting the letter targets was 475 ± 24 ms, and the mean hit rate was 77.3% ± 7.2%. When subjects attended to the face/house image stream, mean RT for detection of the blurry targets was 461 ± 29 ms for the faces and 467 ± 29 ms for the houses; the corresponding hit rates were 89.9% ± 5.2% for the blurry face targets and 89.0% ± 7.7% for the blurry house targets.

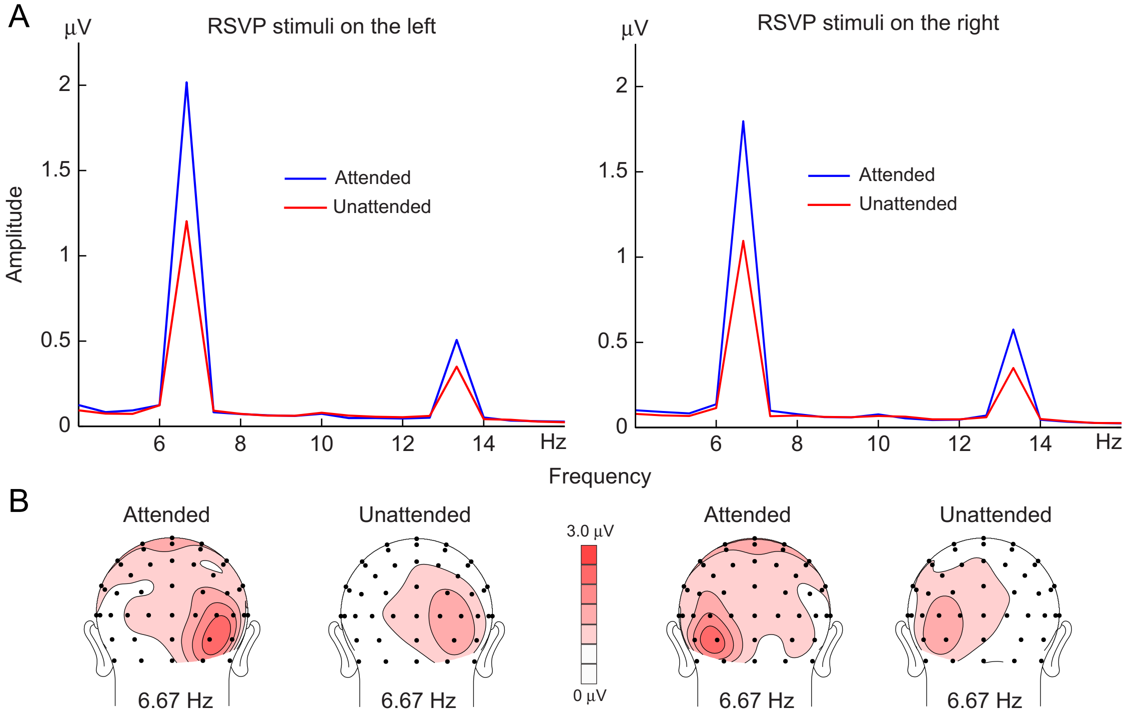


Fig. 2. (A) Grand average amplitude spectrum obtained by Fourier analysis of SSVEP waveforms averaged across a cluster of 10 electrodes for attended and unattended RSVP streams of characters presented in the left visual field (left panel) and right visual field (right panel). Note that SSVEP amplitudes are maximal at the stimulus presentation frequency (6.67 Hz), with a smaller response at the harmonic (13.34 Hz). (B) Topographical distributions of SSVEP amplitudes at 6.67 Hz to RSVP streams presented on the left and right sides are shown separately for the attended and unattended conditions. Enhanced SSVEPs were elicited by attended RSVP streams over the hemisphere contralateral to the stimulus location.

3.2. ERP results

Consistent with previous studies (Morgan et al., 1996; Müller & Hillyard, 2000; Mishra et al., 2011), the SSVEP amplitude to the RSVP stream of characters was enlarged by attention at posterior electrode sites [$F(1,12)=34.165$, $p<.0001$], indicating enhanced sensory processing of the RSVP stream when it was attended (Fig. 2A). The stimulus location × hemisphere interaction was also significant [$F(1, 12)=16.451$, $p<.002$]; analysis of the interaction revealed that the SSVEP to the RSVP stream on the left side was significantly larger over the right than the left hemisphere [$F(1, 12)=23.092$, $p<.001$], whereas the SSVEP to the RSVP stream on the right side was larger over the left than the right hemisphere [$F(1, 12)=5.012$, $p<.05$] (see Fig. 2B).

For the ERPs elicited by the images, comparisons were made between upright and inverted images, separately for faces and houses, with respect to whether the image stream was attended or unattended. ERPs and topographic maps of difference waves (inverted minus upright) elicited by faces and houses in both the attended and unattended conditions are shown in Figs. 3 and 4. There was no significant main effect of face orientation on P1 (100–132 ms) amplitude [$F(1, 12)=3.036$, $p=.107$], but a significant main effect of attention was observed, with attended faces eliciting much larger P1 amplitudes than unattended faces [$F(1, 12)=38.195$, $p<.001$]. Importantly, there was a significant interaction of attention × orientation [$F(1, 12)=31.835$, $p<.001$]; pairwise comparisons revealed that inverted faces evoked a larger P1 than upright faces when presented at the attended location [$F(1, 12)=24.092$, $p<.001$] but not at the unattended location [$F(1, 12)=2.657$, $p=.129$] (see Figs. 3 and 5). For houses, a significant main effect of attention was observed, with attended houses eliciting a larger P1 than unattended houses [$F(1, 12)=16.124$, $p<.002$]. However, neither the main effect of orientation [$F(1, 12)=.001<1$, $p=.980$] nor the attention × orientation interaction [$F(1, 12)=0.016$, $p=.902$] was significant for houses.

In the present study the N170 peak latency ranged from 170 to 200 ms in different individuals. A repeated measure ANOVA on the latency of N170 to faces revealed a significant main effect of orientation [$F(1, 12)=33.590$, $p<.0001$], with longer latencies elicited by inverted faces (187 ± 3.2 ms) than by upright faces (180 ± 2.7 ms). Neither the main effect of attention [$F(1, 12)=2.895$, $p=.115$] nor the attention × orientation interaction [$F(1, 12)=2.396$, $p=.148$] was significant for the N170 latency to faces. A similar analysis carried out on the latency of the house-evoked N170 revealed no significant main effects of attention [$F(1, 12)=0.663$, $p=.413$] or orientation [$F(1, 12)=0.818$, $p=.383$] nor an attention × orientation interaction [$F(1, 12)=0.194$, $p=.667$].

An ANOVA carried out on the amplitude of the face evoked N170 (172–212 ms) revealed no significant main effect of orientation [$F(1, 12)=2.513$, $p=.139$]. In contrast with the P1, however, the N170 amplitude (measured with respect to the pre-stimulus baseline) was significantly larger to unattended faces than to attended faces [$F(1, 12)=18.336$, $p<.002$] (see Figs. 3 and 5). Pairwise comparisons indicated that the N170 amplitude was larger to unattended faces than to attended faces regardless of the face orientation [for upright faces, $F(1, 12)=15.376$, $p<.003$; for inverted faces, $F(1, 12)=19.282$, $p<.002$]. The attention × orientation interaction was also significant [$F(1, 12)=8.418$, $p<.02$]; analysis of this interaction revealed that N170 amplitude to inverted faces was larger than to upright faces when faces were unattended [$F(1, 12)=4.897$, $p<.05$], but not when they were attended [$F(1, 12)=0.663$, $p=.431$]. For the N170 to houses, only the main effect of attention was significant [$F(1, 12)=15.414$, $p<.003$], with unattended houses eliciting a larger N170 than attended houses. Neither the main effect of orientation [$F(1, 12)=1.903$, $p=.193$] nor the attention × orientation interaction [$F(1, 12)=0.353$, $p=.563$] was significant.

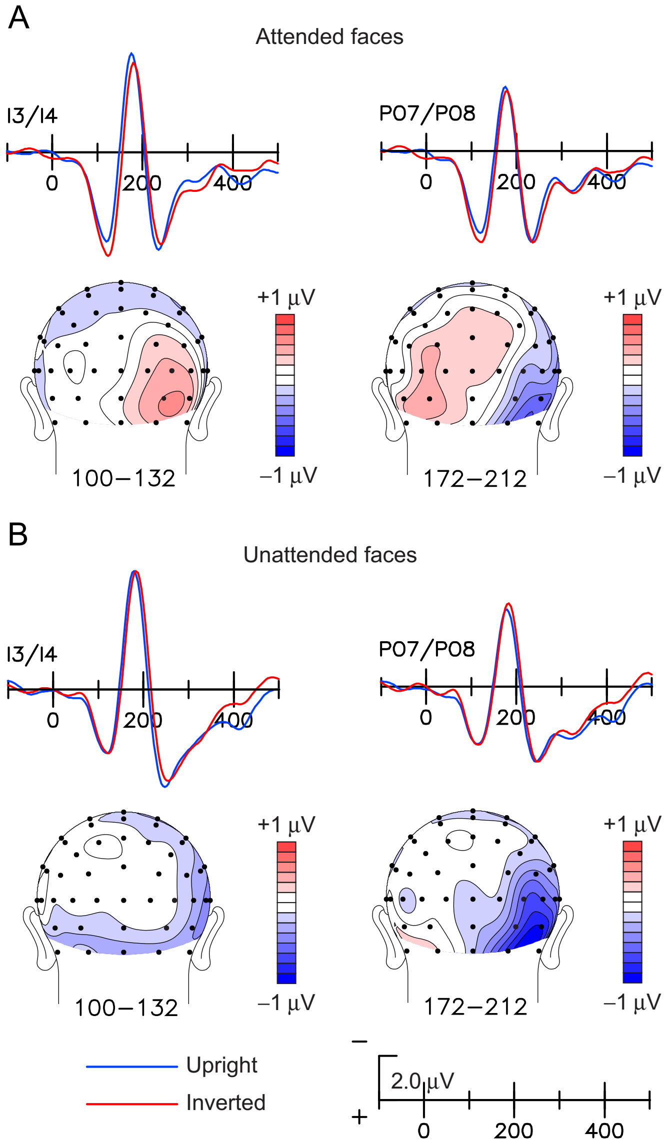


Fig. 3. Grand average ERPs to upright and inverted faces when they were attended (A) and unattended (B). ERPs shown are from contralateral electrode sites (I3/I4 and PO7/PO8), collapsed over left and right field stimulus streams. Voltage maps of the inversion difference waves (inverted faces minus upright faces) are shown for attended and unattended faces in the P1 (100–132 ms) and N170 (172–212 ms) latency ranges. Contralateral sites are shown on the right sides of the maps and ipsilateral sites on the left. Note that P1 was larger for inverted than for upright faces only when they were attended ($p<.001$).

To allow a comparison with previous studies (for example, Crist et al., 2008), the attention effect on the face specific N170 was calculated as the difference between the face-evoked N170 and the house-evoked N170 mean amplitudes (with respect to baseline). An ANOVA of this difference amplitude (over 172–212 ms) showed no significant main effect of attention [$F(1, 12)=1.047$, $p=.326$], orientation [$F(1, 12)=0.091$, $p=.796$], or attention × orientation interaction [$F(1, 12)=0.763$, $p=.400$]. These results suggest that the face specific N170 was not modulated by attention under the present conditions.

3.3. Topographical distribution and source localization

The neural sources of the face inversion effects evident in the P1 and N170 components were estimated using Brain

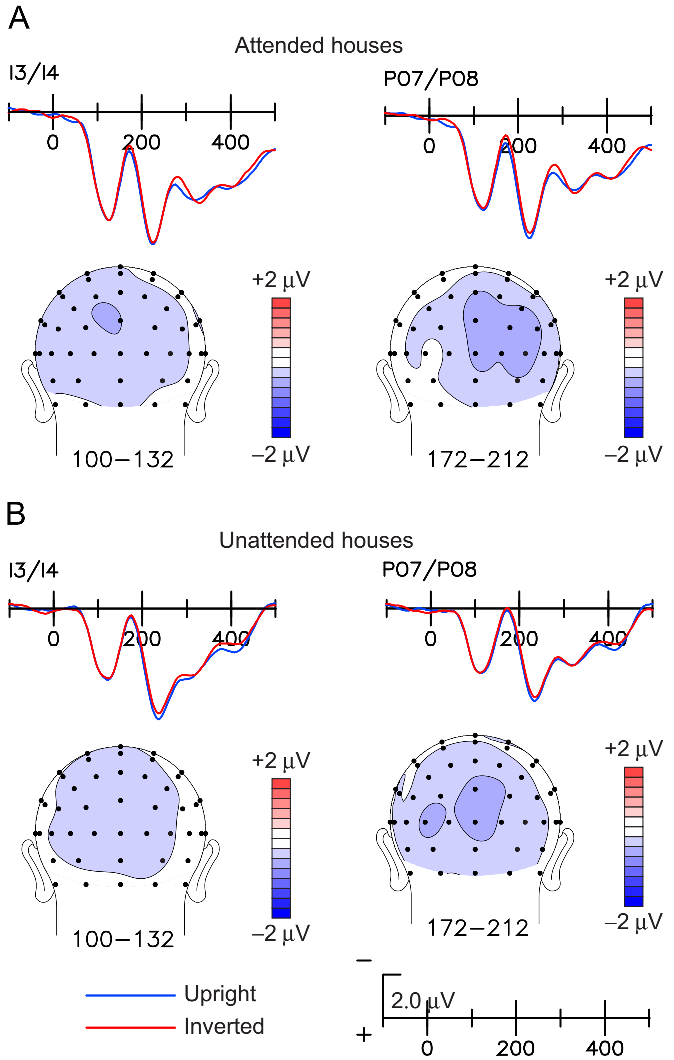


Fig. 4. Grand average ERPs to upright and inverted houses when they were attended (A) and unattended (B). ERPs shown are from contralateral electrode sites (I3/I4 and PO7/PO8), collapsed over left and right field stimulus streams. Voltage maps of the difference waves (inverted houses minus upright houses) are shown for attended and unattended houses in the P1 (100–132 ms) and N170 (172–212) latency ranges. Contralateral sites are shown on the right side of the maps and ipsilateral sites on the left.

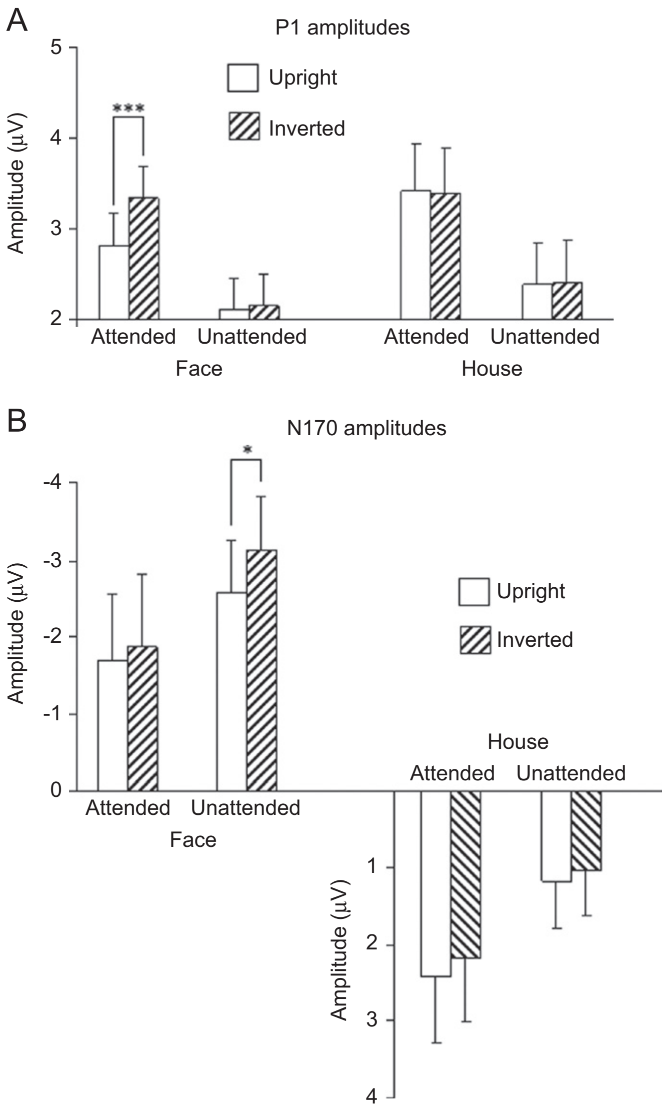


Fig. 5. (A) Mean amplitudes of P1 (100–132 ms) averaged over contralateral sites (O1/O2, P7/P8, PO7/PO8, SI3/SI4, I3/I4, I5/I6) elicited by upright and inverted faces (left) and houses (right) under attended and unattended conditions. *** indicates $p < .001$. (B) Same as A for the N170 (172–212 ms) amplitudes. * indicates $p < .05$.

Electrical Source Analysis (BESA version 5.1.8). For the P1, a pair of dipoles was fit to the scalp topography of the difference waveform (attended inverted face minus attended upright face, shown in Fig. 3A) within the specified time interval (100–132 ms). For the N170, a pair of dipoles was fit to the scalp topography of the difference waveform (unattended inverted face minus unattended upright face, shown in Fig. 3B) within the 172–212 ms time interval. Each pair of dipoles was constrained to be mirror-image in location only. The resulting dipole models revealed that the neural generators of the attended P1 inversion difference were estimated to lie within the ventral occipito-temporal cortex in correspondence of the fusiform gyrus (Talairach coordinates: $x = \pm 34$, $y = -58$, $z = -14$); the unattended N170 inversion difference was localized to a nearby region of the same gyrus (Talairach coordinates: $x = \pm 37$, $y = -60$, $z = -15$). These results suggest that the P1 and N170 inversion effects for faces were generated in the same or neighboring brain areas (Fig. 6).

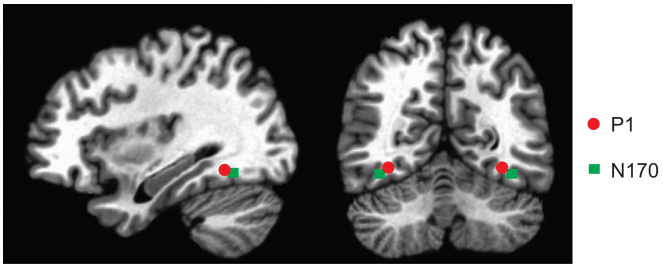


Fig. 6. Estimated dipolar sources of the P1 (100–132 ms) inversion difference waves (inverted minus upright) for attended faces ($x = \pm 34$, $y = -58$, $z = -14$) and of the N170 (172–212 ms) inversion difference waves for unattended faces ($x = \pm 37$, $y = -60$, $z = -15$).

4. Discussion

In the present study, spatial attention was manipulated by instructing participants to attend to one of two stimuli streams (image stream or RSVP stream) at the start of each block. The target detection accuracy was higher for the image task (89% hit rate) than for the RSVP task (77% hit rate), and both levels of performance were well above chance, indicating a strong allocation of attention to the assigned task. The ERP recordings verified that attention was allocated selectively according to the task instruction: faces and houses elicited larger P1 amplitudes when attended than when unattended (cf. Hillyard, Vogel, & Luck, 1998), and the SSVEP amplitudes were larger when the alphanumeric stream was attended than when unattended (cf. Morgan et al., 1996; Müller & Hillyard, 2000; Mishra et al., 2011). Thus the modulations of face-specific processing observed here as a function of task instruction can indeed be attributed to the selective allocation of spatial attention.

In the present study, the effect of attention on early face-specific processing was assessed by comparing ERPs elicited by upright faces to ERPs elicited by inverted faces according to whether they were attended and unattended. The results showed that inverted faces evoked larger P1 amplitudes than upright faces when they were attended, but this inversion effect on P1 was absent when faces were unattended. In contrast, the N170 component elicited by faces was larger for inverted than upright faces only when they were unattended. The peak latency of the N170 was delayed for inverted faces relative to upright faces, however, under both attended and unattended conditions. Topographical mapping showed that the P1 difference waves (inverted faces minus upright faces) in the attended condition and N170 difference waves (inverted faces minus upright faces) in the unattended condition had very similar scalp distributions. Source localization analysis suggested that these inversion effects for the P1 and N170 were generated in the same or neighboring regions of the ventral occipito-temporal cortex, and the Talairach coordinates of the best fitting dipolar sources pointed to a localization within the well-known fusiform face area (e.g., Kanwisher et al., 1997). For house stimuli, no significant inversion effect was found on P1 amplitude, and neither N170 latency nor the N170 amplitude showed significant inversion effects. This pattern of results suggests that the P1 and N170 inversion effects are face-specific and that early stages of face-specific processing can be enhanced by attention.

The P1 results observed here are consistent with previous ERP studies that found face inversion effects as early as the P1 component (Linkenkaer-Hansen et al., 1998; Taylor, Edmonds, McCarthy, & Allison, 2001; Itier & Taylor, 2002), suggesting an early stage of face-specific processing stage at around 100–132 ms (Halit et al., 2000; Itier & Taylor, 2004a, 2004b). MEG studies have also suggested an early face processing stage with larger amplitudes of M100 for faces compared to non-face stimuli (Liu et al., 2002). Liu et al. (2002) also found that the M100 component was localized within the same cortical areas as the M170 component. Thus, the early P1/M100 face processing effects appear to reflect an initial stage of face recognition or categorization. The present results extend these findings by showing that this early stage of face processing can be modulated by the allocation of spatial attention. This suggests that the initial neural processes involved in categorizing a face as a face are dependent on spatial attention.

For the N170 component, delayed latencies were found for inverted faces relative to upright faces both when they were attended and unattended, but no such inversion effect on N170 latency was found for houses. A delayed N170 for inverted faces has been consistently reported in previous studies (e.g., Bentin et al., 1996; Rossion et al., 1999, 2000). The amplitude of N170, however, only showed a significant face inversion effect in the present study when the faces were unattended. A similar pattern of inversion effects on N170 latency but not on N170 amplitude has been reported for attended stimuli in previous studies (Bentin et al., 1996; Rossion et al., 2003).

Since the P1 and N170 components are adjacent temporally, similar topographically, and maximal in the same EEG bandwidths (Rousselet et al., 2007), it has been suggested that N170 might be overlapped and modulated by variations in P1 amplitude (Rossion & Jacques, 2008; Goffaux, Gauthier, & Rossion 2003). In the present study such overlap between P1 and N170 may well account for the finding that the N170 amplitude (measured with respect to a pre-stimulus baseline) was significantly smaller to attended than to unattended faces. Such a result could ensue if the enhanced positivity to attended faces manifested in the P1 component extended into the N170 latency range. The fact that these adjacent effects (P1 larger to attended than to unattended faces; N170 smaller to attended than unattended faces) were of similar magnitude (about one microvolt) is consistent with the proposal that the enhanced early positivity to attended faces encompasses both the P1 and N170 latency ranges. An overlap between P1 and N170 might also account for the lack of a face inversion effect on the N170 to attended faces. That is, if the increased P1 positivity to inverted versus upright attended faces was prolonged into the N170 latency range, it could cancel out the increased N170 to inverted versus upright faces that was evident for unattended faces. This inversion effect seen for the N170 to unattended faces suggests that unattended faces might be partially processed with residual attention, perhaps because of the high salience of the face stimuli. The inversion effect on N170 but on inversion effect on P1 for unattended faces may also suggest that face-specific processing occurs in the N170 latency range even when the attentional allocation is insufficient to enable face-specific processing at the earlier P1 latency.

The proposal outlined above that the attention-related modulations of N170 observed here might actually be attributed to overlapping modulations of the P1 is consistent with previous conclusions that face processing in the N170 time range is not modulated by attention (Cauquil et al., 2000; Carmel & Bentin, 2002; Downing et al., 2001; Furey et al., 2006). This proposal is also consistent with the finding that face minus house differences in N170 did not differ between attended and unattended conditions which suggests that no attentional modulation of the face-specific N170 component took place in the current study.

As noted above, the effects of face inversion on the N170 component observed in the present study are consistent with those of previous studies, which found that face inversion delayed the latency of N170, and/or enlarged the N170 amplitude (Bentin et al., 1996; Linkenkaer-Hansen et al., 1998; Rossion et al., 1999, 2000, 2003; Taylor et al., 2001; Itier & Taylor, 2002, 2004a, 2004b; Itier et al., 2006). Also consistent with the present results are findings that inversion effects were not present for non-face objects (Bentin et al., 1996; Rossion et al., 2000; Rebai, Poiroux, Bernard, & Lalonde, 2001). This delay of N170 latency and enhancement of its amplitude by face inversion might be explained by the resulting disruption of processing of the facial configuration (Rossion et al., 1999, 2000).

Several previous ERP and MEG studies have found that the N170/M170 elicited by faces showed little or no effect of attention (Cauquil et al., 2000; Carmel & Bentin, 2002; Downing et al., 2001; Furey et al., 2006). More recently, however, Crist et al. (2008) reported a significant modulation of N170 amplitude by spatial attention, with larger N170 amplitudes for faces relative to houses when they were attended but not when unattended. In the present study, we also examined the N170 attention effect in the same way as Crist et al. (2008) by comparing the face minus house difference amplitude under attended and unattended conditions. No significant effect of attention was found on this face minus house N170 difference amplitude; faces elicited larger

N170 amplitudes than houses both when they were attended and unattended. The design of the present study differed from that of Crist and colleagues in several respects, most notably the images in the present study were presented closer to fixation (four degrees of eccentricity) than in the study of Crist et al. (nine degrees of eccentricity). Holmes et al. (2003) also observed a significant N170 enhancement to attended faces in a design where spatial attention was divided between a pair of concurrently presented faces or houses. These results suggest that when faces are farther away from fixation or when attention is spatially divided, the face-specific processing reflected in the N170 can be modulated by the allocation of spatial attention.

It has been suggested that the P1 may reflect the first order processing that categorizes a face as a face, while the N170 may reflect the subsequent processing of facial features, both of which can be disrupted by inversion (Itier & Taylor, 2002). The current results suggest that the initial face categorization process in the P1 latency range may depend on the allocation of attention to the face's location. However, when spatial attention is directed to another location, this first-order face processing may be attenuated or even absent, while the processing of facial feature reflected in the N170 may remain intact. The present results thus suggest that face-specific processing can begin as early as 100–132 ms when the allocation of attention is sufficient, but when attention is being diverted to a highly demanding task at another location, the face-specific processing may be delayed until a later time range reflected in the N170.

Acknowledgments

This research was supported by Grants from the National Science Foundation (BCS-1029084), the National Institute of Mental Health (1P50MH86385), the Office of Naval Research (N 00014-07-1–0937) and National Basic Research Program of China (973 Program, 2011CB711000).

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Science*, *4*, 267–278.

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.

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83*, 1–29.

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

Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computer and Biomedical Research*, *29*, 162–173.

Crist, R., Wu, C.-T., Karp, C., & Woldorff, M. (2008). Face processing is gated by visual spatial attention. *Frontiers in Human Neuroscience*, *1*, http://dx.doi.org/10.3389/neuro.09.010.2007.

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.

Di Russo, F., Martinez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex*, *13*, 486–499.

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

Downing, P., Liu, J., & Kanwisher, N. (2001). Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia*, *39*, 1329–1342.

Furey, M. L., Tanskanen, T., Beauchamp, M. S., Avikainen, S., Uutela, K., Hari, R., et al. (2006). Dissociation of face-selective cortical responses by attention. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 1065–1070.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*, 495–504.

Goffaux, V., Gauthier, I., & Rossion, B. (2003). Spatial scale contribution to early visual differences between face and object processing. *Cognitive Brain Research*, *16*, 416–424.

Halgren, E., Raij, T., Marinkovic, K., Jousmaki, V., & Hari, R. (2000). Cognitive response profile of the human fusiform face area as determined by MEG. *Cerebral Cortex*, *10*, 69–81.

Halit, H., de Haan, M., & Johnson, M. H. (2000). Modulation of event-related potentials by prototypical and atypical faces. *NeuroReport*, *11*, 1871–1875.

Herrmann, M. J., Ehlis, A. C., Muehlberger, A., & Fallgatter, A. J. (2005). Source localization of early stages of face processing. *Brain Topography*, *18*, 77–85.

Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: Electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *353*, 1257–1270.

Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*, 80–84.

Holmes, A., Vuilleumier, P., & Eimer, M. (2003). The processing of emotional facial expression is gated by spatial attention: evidence from event-related brain potentials. *Cognitive Brain Research*, *16*, 174–184.

Itier, R. J., Herdman, A. T., George, N., Cheyne, D., & Taylor, M. J. (2006). Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Research*, *1115*, 108–120.

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *NeuroImage*, *15*, 353–372.

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

Itier, R. J., & Taylor, M. J. (2004b). Effects of repetition learning on upright, inverted and contrast-reversed face processing using ERPs. *NeuroImage*, *21*, 1518–1532.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*, 4302–4311.

Linkenkaer-Hansen, K., Palva, J. M., Sams, M., Hietanen, J. K., Aronen, H. J., & Ilmoniemi, R. J. (1998). Face-selective processing in human extrastriate cortex around 120 ms after stimulus onset revealed by magneto- and electroencephalography. *Neuroscience Letters*, *253*, 147–150.

Liu, J., Harris, A., & Kanwisher, N. (2002). Stages of processing in face perception: An MEG study. *Nature Neuroscience*, *5*, 910–916.

Lu, S. T., Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Lounasmaa, O. V., Sams, M., et al. (1991). Seeing faces activates three separate areas outside the occipital visual cortex in man. *Neuroscience*, *43*, 287–290.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Science*, *6*, 255–260.

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

Mishra, J., Zinni, M., Bavelier, D., & Hillyard, S. A. (2011). Neural basis of superior performance of action videogame players in an attention-demanding task. *Journal of Neuroscience*, *31*, 992–998.

Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 4770–4774.

Müller, M. M., & Hillyard, S. A. (2000). Concurrent recording of steady-state and transient event-related potentials as indices of visual-spatial selective attention. *Clinical Neurophysiology*, *111*, 1544–1552.

Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guerit, J. M., Hinrichs, H., et al. (1998). IFCN standards for digital recording of clinical EEG. International Federation of Clinical Neurophysiology. *Electroencephalography Clinical Neurophysiology*, *106*, 259–261.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral Cortex*, *14*, 619–633.

Pourtois, G., Thut, G., Grave de Peralta, R., Michel, C., & Vuilleumier, P. (2005). Two electrophysiological stages of spatial orienting towards fearful faces: Early temporo-parietal activation preceding gain control in extrastriate visual cortex. *NeuroImage*, *15*, 149–163.

Rebai, M., Poiroux, S., Bernard, C., & Lalonde, R. (2001). Event-related potentials for category-specific information during passive viewing of faces and objects. *International Journal of Neuroscience*, *106*, 209–226.

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.

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.

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.

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.

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2007). Single-trial EEG dynamics of object and face visual processing. *NeuroImage*, *36*, 843–862.

Sams, M., Hietanen, J. K., Hari, R., Ilmoniemi, R. J., & Lounasmaa, O. V. (1997). Face-specific responses from the human inferior occipito-temporal cortex. *Neuroscience*, *77*, 49–55.

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

Susac, A., Ilmoniemi, R. J., Pihko, E., & Supek, S. (2004). Neurodynamic studies on emotional and inverted faces in an oddball paradigm. *Brain Topography*, *16*, 265–268.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: An approach to cerebral imaging*. New York: Thieme.

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

Tottenham, N., Tanaka, J., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., et al. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*, 242–249.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.