Context Influences Early Perceptual Analysis of Faces—An Electrophysiological Study

Ruthger Righart[1] and Beatrice de Gelder[1,2]

[1]Department of Psychology, Cognitive and Affective Neurosciences Laboratory, Tilburg University, 5000 LE Tilburg, The Netherlands and [2]Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA 02129, USA

Electrophysiological and hemodynamic correlates of processing isolated faces have been investigated extensively over the last decade. A question not addressed thus far is whether the visual scene, which normally surrounds a face or a facial expression, has an influence on how the face is processed. Here we investigated this issue by presenting faces in natural contexts and measuring whether the emotional content of the scene influences processing of a facial expression. Event-related potentials were recorded to faces (fearful/neutral) embedded in scene contexts (fearful/neutral) while participants performed an orientation-decision task (face upright or inverted). Two additional experiments were run, one to examine the effects of context that occur without a face and the other to evaluate the effects of faces isolated from contexts. Faces without any context showed the largest N170 amplitudes. The presence of a face in a fearful context enhances the N170 amplitude over a face in neutral contexts, an effect that is strongest for fearful faces on left occipito-temporal sites. This N170 effect, and the corresponding topographic distribution, was not found for contexts-only, indicating that the increased N170 amplitude results from the combination of face and fearful context. These findings suggest that the context in which a face appears may influence how it is encoded.

Keywords: affective pictures, context effects, facial expressions, N170, P1, threatening scenes

Introduction

In everyday life, faces, just like objects, appear within a natural environment. It has been shown previously that natural contexts play an important role in the recognition of objects. A football player, for instance, is usually seen in a football stadium, and his presence is rather unusual in other contexts (Davenport and Potter 2004). If the probability is high that a certain context surrounds a visual object, the processing of that object is facilitated, whereas unexpected contexts tend to inhibit it (Palmer 1975; Ganis and Kutas 2003; Davenport and Potter 2004; but see also Hollingworth and Henderson 1998; see for a review Bar 2004).

Context influences have been investigated by varying the semantic relationship between a target object and its background context. In such experiments, an object is embedded either in a context in which it is highly expected (congruent object–context compounds) or in a context in which it is highly unexpected (incongruent compounds). Objects that were placed in semantically congruent contexts were recognized more accurately (Davenport and Potter 2004) and faster (Ganis and Kutas 2003) than objects in semantically incongruent contexts. Further, the time course of these context effects

was related to a decrease in the N400 when target object and context were congruent (Ganis and Kutas 2003). These findings indicate that knowledge about the visual world may influence our expectations of what objects should appear in a visual scene (Bar 2004).

A different factor that may influence stimulus processing besides semantic congruency between target object and context is the presence of emotional information (Smith, Dolan, and Rugg 2004; Smith, Henson, Dolan, and Rugg 2004). Emotional scenes are rapidly evaluated by the visual system for both biologically (e.g., snake) and artificially (e.g., gun) salient elements. Fast evaluation of salient stimuli is reflected in a prioritizing of threatening scenes, as is seen in rapid gaze shifts (Calvo and Lang 2004) and faster response times for the detection of fear-relevant pictures (Öhman and others 2001) as well as enhanced P1 amplitudes for threatening scenes (Smith and others 2003; Carretie and others 2004). Further, single-neuron responses in the right ventral prefrontal cortex for aversive scenes diverge from neutral scenes at around 120 ms (Kawasaki and others 2001).

Similarly, studies have shown that facial expressions can be discriminated from each other rapidly (White 1995). In electrophysiological studies, it was observed that the N170 amplitude is larger for faces than for other objects (Bentin and others 1996; Itier and Taylor 2004a). The N170 is considered to be associated with structural encoding of the face (Bentin and others 1996; Eimer 2000) and may be insensitive for facial expressions (Eimer and Holmes 2002; Holmes and others 2003). However, recent studies indicate that the N170 amplitude increases to fearful expressions (Batty and Taylor 2003; Stekelenburg and de Gelder 2004). The N170 is also sensitive to social relevant properties of the face (Pizzagalli and others 2002). Furthermore, effects of facial expression have also been observed at earlier components (~100 ms) at occipital (Pizzagalli and others 2002; Batty and Taylor 2003; Eger and others 2003) and frontal electrode sites (Eimer and Holmes 2002; Holmes and others 2003), and inferior occipital sources have been found by magnetoencephalography at 110 ms after stimulus onset (Halgren and others 2000). Further, the extrastriate regions of the brain respond to the emotional intensity of facial expressions, which implicates that visual areas are involved in the emotional analysis of stimuli (Surguladze and others 2003). Thus, the available evidence indicates that the emotional content of faces as well as of scenes is discriminated from neutral content at an early stage of processing. A few studies have already explored the electrophysiological (Guillaume and Tiberghien 2001; Tsivilis and others 2001; Rousselet and others 2004) and hemodynamic (Yi and others

2004) correlates of processing faces and objects in neutral contexts. The influence of emotional contexts on processing faces has not been investigated thus far.

In the present study, we focused on the early electrophysiological correlates (P1 and N170) of facial expressions embedded in emotional contexts. Our hypothesis was that because both emotional attributes of faces and of scenes influence perception, the combination of emotional faces and emotional contexts might increase processing of faces in contexts. If fearful faces are accompanied by a fearful context, the amplitudes of the P1 and N170 components may increase.

Method

Participants
Twelve neurological healthy participants (10 females) with normal or corrected-to-normal vision volunteered ($M = 21.1$ years). All gave informed consent. Ten were right handed.

Materials and Procedure
Stimuli consisted of 24 face photographs (6 females and 6 males posing a fearful or neutral expression) taken from a validated image database (Ekman and Friesen 1976) and 12 scenes used as contexts (6 fearful-related contexts, e.g., knife, crashed car, injection needle; 6 neutral contexts, e.g., house, sofa, guitar). Scenes were selected from the International Affective Picture System (Lang and others 1999) and complemented with scenes found on the Web. Participants were not familiar with the selected stimuli.

The pictures were evaluated by another group of subjects ($N = 15$) on arousal (9-point scale from 1 = calm to 9 = extremely arousing) and valence (1 = very unpleasant to 9 = very pleasant). Emotional arousal rates of the fearful pictures were reliably different from the neutral pictures (fearful $M = 7.31$, $SD = 0.64$; neutral $M = 3.77$, $SD = 0.64$; $t(14) = 14.33$, $P < 0.001$). Fearful pictures were evaluated significantly different on emotional valence than neutral pictures (fearful $M = 3.43$, $SD = 1.23$; neutral $M = 6.28$, $SD = 0.67$; $t(14) = 7.08$, $P < 0.001$). Every fearful picture was evaluated as being more unpleasant and more arousing than any of the neutral pictures.

All stimuli were gray-scale pictures. Faces were overlaid on the center of a context stimulus in such a way that the faces did not occlude the critical parts of the context (Fig. 1). The height and width of the facial images were 6.5 × 4.5 cm (4.7° × 3.2°) and for context images 20 × 30 cm (14.3° × 21.2°), respectively. Participants sat in an electrically shielded cabin at an 80-cm distance from the monitor.

This study comprised 3 experiments: a main experiment in which face–context compounds were presented and 2 additional experiments containing faces-only and contexts-only. The order of the 3 experiments was counterbalanced across participants. All individual experiments were preceded by a short training session to get familiarized with the procedures.

The face–context compounds were shown for 500 ms (Fig. 1). Stimuli were followed by a central fixation cross which was presented for 700 ms. Participants performed an orientation-decision task, in which they had to indicate the orientation of the face (upright or inverted) by
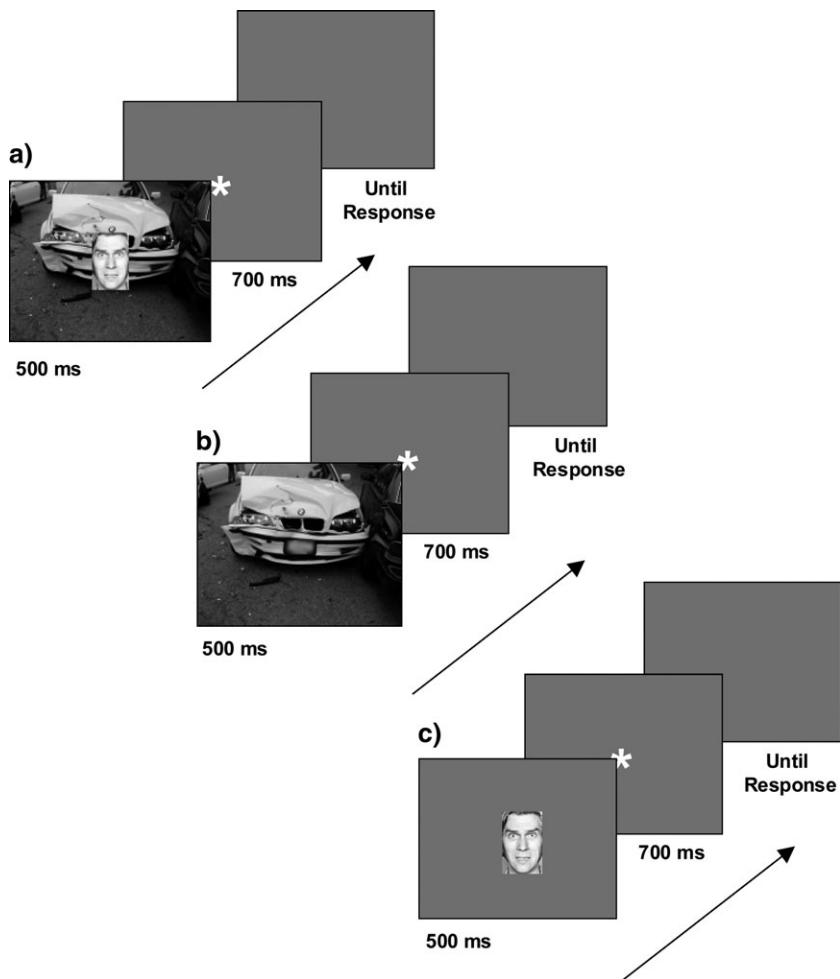
Figure 1. Designs of (*a*) faces in context, (*b*) contexts-only, and (*c*) faces-only experiments. Participants were instructed to indicate the orientation of the face in the faces in context and faces-only experiments. Participants had to detect a target trial (i.e., telephone) among context trials in the contexts-only experiment.

pressing 1 of 2 buttons after disappearance of the fixation cross. This delay in response was used in order to avoid contamination by motor-response-related artifacts. The orientation task was chosen to prevent any task-related effects for emotions in face or context. Compounds were created of fearful faces with fearful contexts and neutral contexts and neutral faces with the same fearful contexts and neutral contexts. These compounds were created for upright and inverted faces, resulting in 8 conditions. Each condition contained 72 unique face-context combinations. Stimuli were delivered in 3 separate blocks of 192 trials. Each of the blocks contained each condition equally (i.e., 8 × 24 face-context combinations).

In the contexts-only experiment, the same context stimuli were presented, but without faces, and target trials (i.e., a telephone) were presented. Participants were instructed to press one button when the target trial was presented and the other button for a nontarget context. The nontarget context stimuli were the same stimuli as the backgrounds in the face-context experiment. We hypothesized that if contexts influence event-related potential (ERP) components of face processing, then this ERP effect should be different from potentials generated by contexts-only.

Another experiment was run to control for the effects of face stimuli separate from context. The same face stimuli but without contexts (i.e., faces-only) were presented to provide a baseline against which faces with contexts could be compared. The face stimuli were overlaid on a gray background, similarly sized to the face and background in the face-context experiment. As in the face-context experiment, participants had to indicate the orientation of the face. All other specifications are similar to the face-context experiment described above. We hypothesized that if contexts affect face processing, the ERP components for faces-only could be differentiated from components generated for faces in contexts.

EEG Recording
EEG was recorded from 49 locations using active Ag–AgCl electrodes (BioSemi Active2) mounted in an elastic cap, referenced to an additional active electrode (Common Mode Sense) during recording. EEG signals were band-pass filtered (0.1–30 Hz, 24 dB/Octave). The sampling rate was 256 Hz. All electrodes were off-line referenced to an average reference. Horizontal electrooculographies (EOGs) were recorded from 2 electrodes placed at the outer canthi of both eyes. Vertical EOGs were recorded from electrodes on the infraorbital and supraorbital regions of the right eye in line with the pupil. Raw EEG data were segmented into epochs starting 100 ms before stimulus onset to 900 ms after the stimulus onset. Data were baseline corrected to the first 100 ms of the epoch.

After EOG correction using the algorithm of Gratton and others (1983), epochs with amplitude exceeding 100 µV at any channel were rejected from analyses. ERPs were averaged for each condition. Only trials on which subjects responded correctly were averaged. For the contexts-only experiment, nontarget trials were averaged in order to compare ERPs for contexts-only with faces in contexts. The target trials were shown infrequently to maintain fixation and attention and were

therefore not analyzed. Electrode selection was based on previous studies showing maximal amplitudes for P1 and N170 on these sites (Batty and Taylor 2003; Stekelenburg and de Gelder 2004). Peak latency and amplitude of P1 were scored at occipital sites (O1/2) and parieto-occipital sites (PO3/4) as the maximal positive peak in the time window 100–150 ms with respect to baseline. For N170, peak latency and amplitude were scored at occipito-temporal sites (P5/6, P7/8, and PO7/8) as the maximal negative peak in the time window 140–220 ms (for electrode positions, see Fig. 2).

Analyses
P1 and N170 latencies and amplitudes for faces in contexts were analyzed with multivariate analyses for repeated measures containing the within-subject factors' facial expression (fearful, neutral), context emotion (fearful, neutral), hemisphere (left, right), and electrode position.

Additional comparisons were made for the faces-only and contexts-only experiments. First, for faces-only, fearful/neutral faces in contexts were tested against fearful/neutral faces without a context by adding the factor context presence (present, absent). Second, for contexts-only, the effects of faces in fearful/neutral contexts were tested against fearful/neutral context without a face by adding the factor face presence (present, absent). Scalp topographic distributions were then analyzed for the N170 component for the comparison of faces in fearful/neutral contexts with fearful/neutral contexts-only. Difference waves (i.e., fearful contexts – neutral contexts) were calculated separately for faces in context and contexts-only. Amplitudes were vector normalized according to the method described by McCarthy and Wood (1985) and subjected to repeated-measures analysis of variance (ANOVA). Additionally, difference waves (i.e., faces in context – context-only) were calculated to test scalp distributions of fearful contexts against neutral contexts. P values were corrected by Greenhouse-Geisser epsilon correction, if appropriate. Statistics are indicated with original degrees of freedom (Picton and others 2000).

Finally, face inversion effects for P1 and N170 latencies and amplitudes were analyzed with repeated-measures ANOVA containing the within-subject factors' facial expression (fearful, neutral) and orientation (upright, inversion).

Results

Behavioral Results
Performance on the orientation-decision task (upright vs. inverted) was nearly flawless. The accuracy for upright faces was 99.2% and for inverted faces 99.1%. No significant differences were found among any of the conditions ($P > 0.05$). Response times were not analyzed because of the delayed-response paradigm.
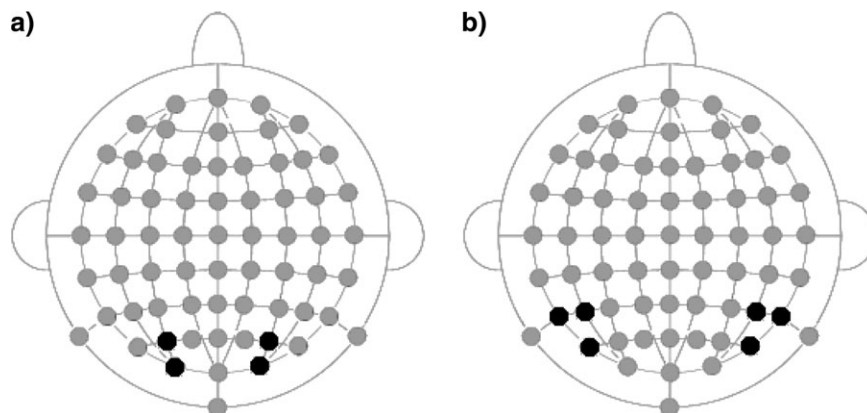
Figure 2. Scalp locations on which P1 (*a*) and N170 (*b*) were measured.

P1 Latency

There was a main effect of facial expression ($F(1,11) = 9.09$, $P < 0.05$) in that latencies were prolonged for fearful (135 ms) as compared with neutral (133 ms) faces. No main effect of context emotion was observed ($P > 0.05$). The factor face presence was added to the analysis to test whether contexts with faces could be differentiated from contexts without faces. The interaction between face presence and context emotion was near significance ($F(1,11) = 4.15$, $P = 0.07$). Post hoc comparisons showed that latencies were prolonged for neutral context containing a face (135 ms) as compared with neutral contexts without a face (129 ms) ($P < 0.05$). Furthermore, faces in contexts were compared with faces without a context. Latencies did not vary significantly as a consequence of context presence (all P values > 0.05). The latency difference between fearful and neutral faces was not significant for faces without any context ($P > 0.05$).

P1 Amplitude

No effects were found for facial expression ($P > 0.05$). The main effect for context emotion on P1 amplitude was marginally significant ($F(1,11) = 4.04$, $P = 0.07$), reflecting increased amplitudes for faces accompanied by fearful contexts (10.91 µV) as compared with neutral contexts (10.28 µV). However, this effect was not specific for faces in context as there was no significant interaction between context emotion and face presence ($P > 0.05$).

A main effect was found for face presence ($F(1,11) = 26.28$, $P < 0.001$) as amplitudes were larger for contexts in which a face was present (10.59 µV) than absent (9.52 µV). The amplitudes were marginally increased for faces without a context (11.76 µV) compared with faces in a context (10.59 µV) ($F(1,11) = 3.84$, $P = 0.07$), reaching significance on the posterior electrodes O1/2 ($P < 0.05$). Like faces in context, no main effects were found for facial expression without a context ($P > 0.05$).

N170 Latency

Latencies did not vary significantly as a consequence of context emotion ($P > 0.05$). A main effect was observed for facial expression ($F(1,11) = 12.13$, $P < 0.01$) as reflected by longer latencies for fearful faces (186 ms) than neutral faces (182 ms). The factor facial expression had a significant interaction with hemisphere and electrode position ($F(2,10) = 6.42$, $P < 0.05$). Post hoc comparisons showed that latencies for fearful faces were longer than neutral faces on each electrode but reached significance only for the right electrode P6 ($P < 0.05$). Latencies did not vary as a consequence of face presence ($P > 0.05$).

For faces-only, latencies were longer for fearful faces (189 ms) over neutral faces (182 ms) as well ($F(1,11) = 29.28$, $P < 0.05$). This difference was significant at each electrode ($P < 0.05$).

N170 Amplitude

The factor electrode position showed a main effect ($F(2,10) = 37.39$, $P < 0.001$) as N170 amplitudes were more negative for electrodes P7/8 (–2.22 µV) compared with electrodes P5/6 (+1.74 µV) and PO7/8 (+1.84 µV). No main effects were found for hemisphere or facial expression ($P > 0.05$). A main effect was observed for context emotion ($F(1,11) = 28.63$, $P < 0.001$), but this effect was qualified by an interaction with hemisphere ($F(1,11) = 12.16$, $P < 0.01$). Post hoc comparisons showed that

N170 amplitudes were more negative for faces in fearful contexts as compared with faces in neutral contexts, but only significantly for electrodes on the left hemisphere ($t(11) = 7.90$, $P < 0.001$). As indicated in Figure 3a, N170 amplitudes were more negative for faces accompanied by fearful contexts (–3.34 µV) than those accompanied by neutral contexts (–1.05 µV) on electrode P7 ($t(11) = 6.07$, $P < 0.001$) but not on electrode P8 (–2.21 and –2.29 µV, respectively). Indeed, all 12 participants showed this pattern on electrode P7 (binomial $P(12/12) < 0.001$).

To test whether this effect is truly explained by the face–context combination, comparisons were made to contexts-only. Therefore, the factor face presence (collapsed across facial expressions) was added to the repeated-measures model. First, a main effect was found for face presence ($F(1,11) = 8.60$, $P < 0.05$) in that amplitudes were more negative for contexts in which a face was present as compared with contexts in which a face was absent (Fig. 3b).

Second, an interaction was found between context emotion and face presence ($F(1,11) = 12.59$, $P < 0.01$), for which post hoc comparisons showed profound differences between fearful and neutral contexts when a face was present ($t(11) = 5.35$, $P < 0.001$), whereas no significant differences were observed between fearful and neutral contexts for which the face was absent ($P > 0.05$).

In addition, scalp topographies for this interaction were significantly different ($F(48,528) = 2.43$, $P < 0.05$) (Fig. 4). This interaction critically suggests that the observed effects were caused by the combination of face and context and not by contexts-only and may suggest that the underlying sources or relative source strengths are (at least partly) different (Picton and others 2000). The topography for faces in fearful contexts against faces in neutral contexts were not significantly different ($F(48,528) = 2.15$, $P > 0.05$) (see Fig. 4).

In addition to the 2-way interaction between context emotion and hemisphere, a 3-way interaction was observed for facial expression, context emotion, and hemisphere ($F(1,11) = 5.11$, $P < 0.05$). In post hoc comparisons, it was found that the N170 amplitudes were more increased for fearful faces than for neutral faces, when they were in the presence of a fearful context, but only on left hemispheric electrodes ($t(11) = 3.82$, $P < 0.01$). As illustrated for P7/8 in Figure 3, sharply enhanced negativities were found for fearful faces in the presence of a fearful context on electrode P7 (–4.11 µV) as compared with neutral faces in a fearful context (–2.57 µV) ($t(11) = 3.10$, $P < 0.05$). Ten out of 12 participants confirmed to this pattern (binomial $P(10/12) < 0.05$) on electrode P7, whereas 9 out of 12 participants (binomial $P(9/12) = 0.07$) showed such an effect on site P8 but that difference did not reach significance on post hoc testing ($P > 0.05$).

To interpret the effects of faces in context against faces without context information, the effect of context presence was analyzed. As indicated in Figure 3c, a main effect was found for context presence ($F(1,11) = 31.48$, $P < 0.001$). N170 amplitudes were more negative for faces in which the context (P7 = –5.73 µV; P8 = –5.92 µV) was absent than faces in which the context was present (P7 = –2.20 µV; P8 = –2.25 µV).

For faces without contexts, a main effect was found for facial expression ($F(1,11) = 5.38$, $P < 0.05$). Amplitudes of fearful expressions (–3.25 µV) were more negative than neutral expressions (–2.53 µV). However, the interaction with hemisphere was not significant ($P > 0.05$).
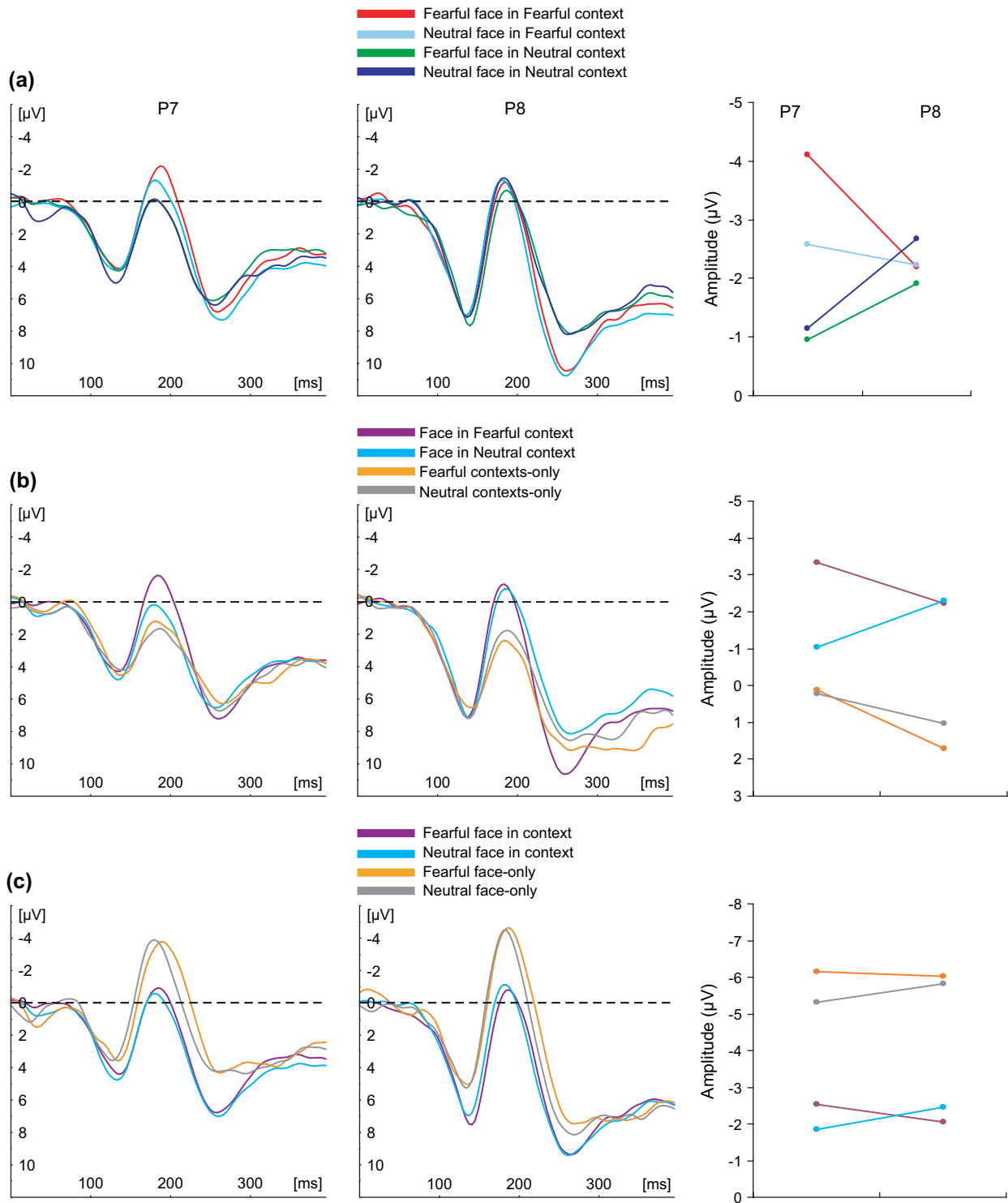
Figure 3. Grand average ERP waveforms and plots of N170 amplitudes. Electrode site P7 (left) and P8 (right) are displayed. Negative amplitudes are plotted upward. (*a*) Faces in context. N170 amplitudes were enhanced for faces in fearful contexts, particularly for fearful faces on left occipito-temporal sites. (*b*) Contexts-only. There were larger amplitudes for faces in fearful contexts than neutral contexts. These differences were not observed for contexts-only. (*c*) Faces-only. N170 amplitudes were larger for faces-only than faces in context.

Inversion Effects

No significant differences were found between upright and inverted faces on P1 latencies and amplitudes neither for faces in context nor for faces-only ($P > 0.05$). Although a significant interaction was observed between orientation and hemisphere ($F(1,11) = 5.26$, $P < 0.05$), this effect was not explained by inversion ($P > 0.05$) but by amplitudes being larger on right electrodes (12.55 µV) than left electrodes for upright faces (10.97 µV) ($P = 0.08$).

N170 latencies were prolonged for inverted (189 ms) as compared with upright faces in context (184 ms) ($F(1,11) = 15.26$, $P < 0.01$). An interaction between orientation and facial
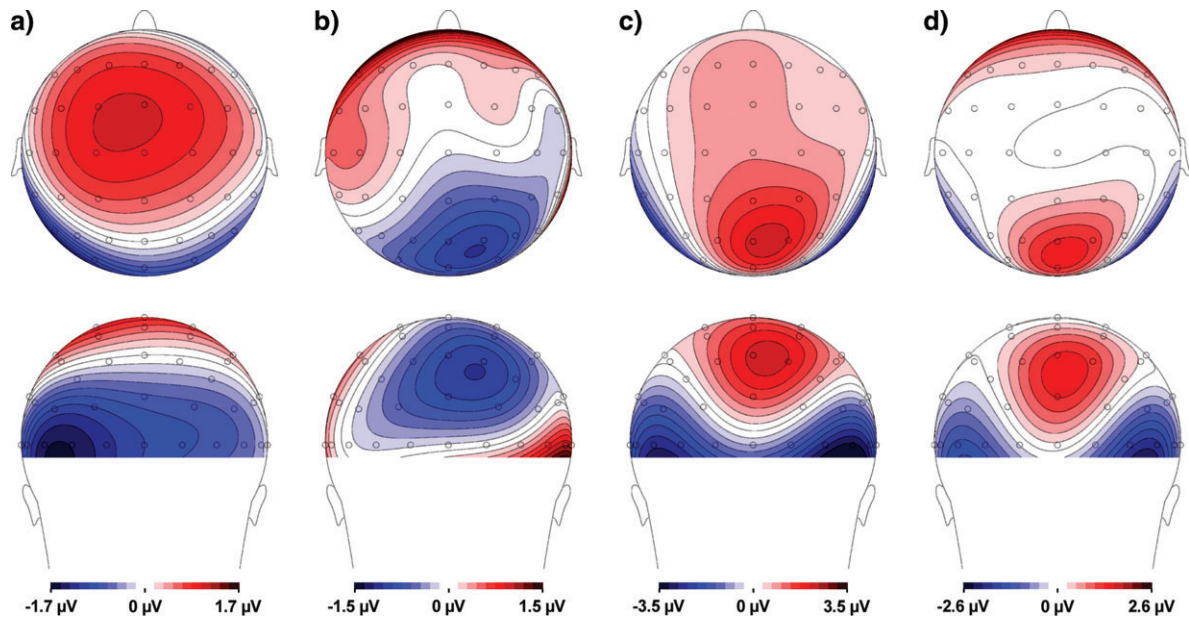
Figure 4. Scalp topography at 180 ms for (*a*) faces in fearful contexts minus faces in neutral contexts tested against (*b*) fearful contexts-only minus neutral contexts-only. Although a topographic difference is present between fearful and neutral contexts-only (in *b*), reflected as a parietal negativity, a similar topography is anticipated if presence of the face had no effect (in *a*). However, a different topography is seen, with a clear left occipito-temporal negativity and central positivity. (*c*) Scalp topography at 180 ms for faces in fearful context minus fearful contexts-only tested against (*d*) faces in neutral context minus neutral contexts-only. A similar topography with clear occipito-temporal negativities is seen.

expression was found for the N170 amplitude ($F(1,11) = 20.25$, $P < 0.001$). Inversion effects on N170 amplitude were found for neutral faces but not for fearful faces in context. For neutral faces, amplitudes were more negative for inverted (–0.26 μV) than upright faces (0.51 μV) ($t(11) = 2.67$, $P < 0.05$), whereas differences were not significant for fearful faces ($P > 0.05$).

Similar results were obtained for faces-only as faces in context. N170 latencies were longer for inverted (196 ms) than upright faces (186 ms) ($F(1,11) = 30.30$, $P < 0.001$). Again, an interaction between orientation and facial expression was found for the N170 amplitude ($F(1,11) = 7.65$, $P < 0.05$). For neutral faces, amplitudes were more negative for inverted (–3.69 μV) than upright faces (–2.53 μV) ($t(11) = 2.26$, $P < 0.05$).

Discussion

We investigated the electrophysiological correlates of perceiving faces in emotional contexts. Several important findings emerge from this study. Firstly, the N170 for face processing is sensitive to the presence of context information, which reduces the amplitude, compared with isolated faces. Secondly and more importantly, the N170 is influenced by the emotional information the context provides. The N170 amplitude was more negative when a face, particularly a fearful one, was presented in a fearful context as compared with a neutral context.

The Influence of Context on Face Processing

A comparison of the ERP components for faces without a context versus faces within a context indicates that early stages of face processing are sensitive to the presence of a context picture. Given the functional meaning of the stage of processing as reflected by the N170, the question arises whether the presence of a context leads to a reduced structural encoding. Previous findings indicate that frontal face views

compared with profiles and with faces not clearly standing off from the background yield higher N170 amplitudes (Eimer 2000). This may indicate better structural encoding for the isolated faces. Likewise, it has been found that N170 amplitudes increase linearly as the level of noise around a face is decreased stepwise (Jemel and others 2003). The data therefore suggest that structural encoding will be better when faces are shown without a context picture. This functional interpretation of the higher N170 amplitudes should be tested in a subsequent memory experiment. This interpretation is, however, consistent with behavioral data. Higher identification rates have been found for target objects that were isolated from any context information as opposed to targets in either congruent or incongruent contexts (Davenport and Potter 2004).

An alternative view is that the N170 amplitudes for faces in contexts are smaller than those for isolated faces because of the perceptual load of the context information. In terms of Lavie (1995), it may be hypothesized that load in processing the target-relevant stimulus (i.e., the face) determines the degree of processing target-irrelevant information (i.e., the context). It was shown that irrelevant distracter letters could interfere with the processing of relevant target letters only if perceptual load was low (Lavie 1995). Similarly, Yi and others (2004) observed that task-irrelevant scenes activated scene-sensitive areas of the brain (i.e., parahippocampal place area, a region of the medial temporal cortex) only if face discrimination was relatively easy. Perceptual load was in the present study low as well, which is reflected in the high accuracies on the face orientation task. Therefore, it may be hypothesized that the task-related stimulus did not demand all available attentional capacity. The irrelevant context may have captured the remaining capacity.

The context information may capture attention and therefore decrease the N170 amplitude (Holmes and others 2003, but see also Séverac-Cauquil and others 2000), which may influence face encoding. A related finding has been reported in a study by

Rolls and others (2003), in which macaque monkeys were shown objects against either a gray context or against a complex natural scene context. Single-neuron recordings showed that the receptive field (i.e., region of the visual field that causes a visual neuron to respond) of the inferior temporal cortex neurons was reduced if complexity of the surrounding context increased. It was suggested that this is due to attentional competition between object and context processing (Rolls and others 2003). A similar mechanism may underlie the effects for contexts on face processing in the present study. (It should be noted here that perceptual load could be better distinguished from structural encoding by using phase-scrambled scenes. According to the structural encoding hypothesis, face encoding should not be different in scrambled or intact scenes. The perceptual load hypothesis would still predict stronger competition from the intact scene. However, we decided to select gray backgrounds in order to reliably replicate the design of earlier studies. We thank an anonymous reviewer for this suggestion.)

However, it should be noted that it is unlikely that all attention was dedicated to the processing of contexts-only. The P1 and N170 amplitude were both larger for faces in contexts than contexts-only. Previous data have shown that the P1 and N170 are both increased for faces compared with other objects (Halgren and others 2000; Itier and Taylor 2004a) or scenes (Puce and others 1999). The time course of these stages of face processing is consistent with single-cell recordings (Sugase and others 1999). The data implicate that faces were effectively differentiated from their contexts. The data replicate previous findings for the N170 obtained for faces in neutral contexts (Guillaume and Tiberghien 2001; Rousselet and others 2004).

The Influence of Context Emotion on Face Processing

The central issue of the present study concerns not the extent of structural encoding in isolated versus contextualized faces but the differential effect of emotional contexts observed on P1 and N170. Comparing the ERP components for faces in fearful contexts with faces in neutral contexts showed that the N170 amplitude for faces is sensitive to the emotion in context. N170 amplitudes were more negative for faces in a fearful context as compared with neutral contexts. An additional experiment containing contexts without a face served as a control to show that the differential effects for emotion in face–context compounds were not generated by the contexts-only.

A number of arguments plead against an explanation of the observed N170 effects by contexts-only. First, whereas the critical difference was found on N170 amplitude for faces in fearful compared with neutral contexts, no significant difference was found between fearful and neutral contexts without a face. This difference between faces in contexts and contexts-only was also confirmed by scalp topography analyses. It is therefore unlikely that low-level features of the context have generated these differential ERPs. Second, participants made correct face orientation decisions and presentation of inverted faces resulted in the signature inversion effect of prolonged N170 latencies and increased N170 amplitudes (Bentin and others 1996; Itier and Taylor 2004a; Rousselet and others 2004). Therefore, these arguments do not support an explanation of differential emotion effects on N170 by contexts-only. We conclude that the findings are critically dependent on the combination of face and context.

An important 3-way interaction was observed between facial expression, context emotion, and hemisphere. N170 amplitudes for fearful faces differed significantly from neutral faces on left occipito-temporal sites, but only if faces were presented in a fearful context. One possible interpretation is that the neutral nature of the context had provided participants sufficient information that no adaptive action was required. Fearful contexts may add important information about the facial expression (see Discussion), which may influence encoding. Future work should address whether congruency effects are present for specific categories of facial expressions (e.g., fear, disgust, and happiness) that are either matched or mismatched with their accompanying contexts.

The finding that face processing is sensitive to the emotional context as reflected in increased N170 amplitudes is consistent with functional magnetic resonance imaging (fMRI) data. Facial expressions of fear and fearful scenes both increase amygdala responses (Hariri and others 2002) and increase the response in the fusiform gyrus (Lang and others 1998; Morris and others 1998; Surguladze and others 2003). Activation of the fusiform gyrus is associated with the intensity of fearful expressions (Morris and others 1998; Surguladze and others 2003). Enhanced responses in fusiform gyrus may indicate feedback modulation from the amygdala to the fusiform gyrus (Morris and others 1998; Vuilleumier and others 2004). As the N170 may have its source in the fusiform gyrus (Pizzagalli and others 2002) its enhancement may reflect feedback modulations from the amygdala to enhance structural encoding. The result that P1 and N170 latencies were prolonged for fearful facial expressions may at first seem counterintuitive. However, slower N170 peak latencies for fear over neutral faces have been reported previously (Batty and Taylor 2003) and may on the other hand reflect prolonged activation to encode available information. Future studies should establish the meaning of peak latencies in processing facial expressions.

The observed effects for the N170 on the left occipito-temporal electrode sites may be related to a predominant response of the left hemisphere to emotion. Although most neuroimaging studies have found a bilateral response of amygdala and fusiform gyrus to emotion in scenes (Lang and others 1998) and facial expressions (Morris and others 1998) or a predominant right fusiform gyrus response (Surguladze and others 2003), a few recent studies have shown left fusiform gyrus responses (Taylor and others 2003; Kim and others 2004). Kim and others (2004) have observed left fusiform gyrus responses to sad faces that were cued by negative sentences (e.g., about losing money) compared with positive sentences (e.g., about winning money). The authors propose that right hemisphere responses may be related to the ambiguous valence of some facial expressions, whereas left hemisphere responses may be evoked when valences are clearly determined (Kim and others 2004; see also Phelps and others 2001). In the present ERP study, fearful contexts possibly fill in the ambiguous nature of the facial expressions, which may have introduced the left hemispheric effects. However, further source analyses and fMRI studies should determine the role of the left hemisphere in processing emotions.

Alternatively, the N170 may have its source in the superior temporal sulcus (Henson and others 2003; Itier and Taylor 2004b), which has been related to social perception (Allison and others 2000). In threatening circumstances, finer encoding and sustained analysis of faces may be necessary to prepare an

appropriate reaction and act adaptively. In accordance with this, in earlier studies it has been shown that the N170 is sensitive to socially relevant cues, such as eyes (Bentin and others 1996), direction of eye-gaze (Watanabe and others 2002), facial motion (Puce and others 2003), biological motion (Jokisch and others 2005), affective facial features (Pizzagalli and others 2002), facial expressions (Batty and Taylor 2003; Stekelenburg and de Gelder 2004), and expressional change (Miyoshi and others 2004). Consistent with this, several neuroimaging and single-cell–recording studies have shown that the superior temporal sulcus is activated by socially relevant information (Allison and others 2000).

The N170 component has been related to structural encoding that is utilized for facial identification (Eimer 2000). The present results do not allow a definite answer to the question whether the larger N170 amplitudes reflect enhanced encoding of either identity or expression of the face, or both, and whether this predicts better recall of facial identity (see Haxby and others 2000). Similar to studies on the effect of emotional scene context on object recognition and memory, future studies of face recognition need to determine how the presence of an emotional context influences encoding and subsequent retrieval (Erk and others 2003; Smith, Dolan, and Rugg 2004; Smith, Henson, Dolan, and Rugg 2004).

Notes

References

Allison T, Puce A, McCarthy G. 2000. Social perception from visual cues: role of the STS region. Trends Cogn Sci 4:267–278.

Bar M. 2004. Visual objects in context. Nat Rev Neurosci 5:617–629.

Batty M, Taylor MJ. 2003. Early processing of the six basic facial emotional expressions. Cogn Brain Res 17:613–620.

Bentin S, Allison T, Puce A, Perez E, McCarthy G. 1996. Electrophysiological studies of face perception in humans. J Cogn Neurosci 8:551–565.

Calvo MG, Lang PJ. 2004. Gaze patterns when looking at emotional pictures: motivationally biased attention. Motiv Emotion 28:221–243.

Carretie L, Hinojosa JA, Martin-Loeches M, Mercado F, Tapia M. 2004. Automatic attention to emotional stimuli: neural correlates. Hum Brain Mapp 22:290–299.

Davenport JL, Potter MC. 2004. Scene consistency in object and background perception. Psychol Sci 15:559–564.

Eger E, Jedynak A, Iwaki T, Skrandies W. 2003. Rapid extraction of emotional expression: evidence from evoked potential fields during brief presentation of face stimuli. Neuropsychologia 41:808–817.

Eimer M. 2000. The face-specific N170 component reflects late stages in the structural encoding of faces. Neuroreport 11:2319–2324.

Eimer M, Holmes A. 2002. An ERP study on the time course of emotional face processing. Neuroreport 13:427–431.

Ekman P, Friesen WV. 1976. Pictures of facial affect. Palo Alto, CA: Consulting Psychologists Press.

Erk S, Kiefer M, Grothe J, Wunderlich A, Spitzer M, Walter H. 2003. Emotional context modulates subsequent memory effect. Neuroimage 18:439–447.

Ganis G, Kutas M. 2003. An electrophysiological study of scene effects on object identification. Cogn Brain Res 16:123–144.

Gratton G, Coles MGH, Donchin, E. 1983. A new method for off-line removal of ocular artifact. Electroencephalogr Clin Neurophysiol 55:468–484.

Guillaume F, Tiberghien G. 2001. An event-related potential study of contextual modifications in a face recognition task. Neuroreport 12:1209–1216.

Halgren E, Raij T, Marinkovic K, Jousmäki V, Hari R. 2000. Cognitive response profile of the human fusiform face area as determined by MEG. Cereb Cortex 10:69–81.

Hariri AR, Tessitore A, Mattay VS, Fera F, Weinberger DR. 2002. The amygdala response to emotional stimuli: a comparison of faces and scenes. Neuroimage 17:317–323.

Haxby JV, Hoffman EA, Gobbini MI. 2000. The distributed human neural system for face perception. Trends Cogn Sci 4:223–233.

Henson RN, Goshen-Gottstein Y, Ganel T, Otten LJ, Quayle A, Rugg MD. 2003. Electrophysiological and haemodynamic correlates of face perception, recognition and priming. Cereb Cortex 13:793–805.

Hollingworth A, Henderson JM. 1998. Does consistent scene context facilitate object perception? J Exp Psychol Hum Percept Perform 127:398–415.

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

Itier RJ, Taylor MJ. 2004a. N170 or N1? Spatiotemporal differences between object and face processing using ERPs. Cereb Cortex 14:132–142.

Itier RJ, Taylor MJ. 2004b. Source analysis of the N170 to faces and objects. Neuroreport 15:1261–1265.

Jemel B, Schuller A, Cheref-Khan Y, Goffaux V, Crommelinck M, Bruyer R. 2003. Stepwise emergence of the face-sensitive N170 event-related potential component. Neuroreport 14:2035–2039.

Jokisch D, Daum I, Suchan B, Troje NF. 2005. Structural encoding and recognition of biological motion: evidence from event-related potentials and source analysis. Behav Brain Res 157:195–204.

Kawasaki H, Adolphs R, Kaufman O, Damasio H, Damasio AR, Granner M, Bakken H, Hori T, Howard MA. 2001. Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. Nat Neurosci 4:15–16.

Kim H, Somerville LH, Johnstone T, Polis S, Alexander AL, Shin LM, Whalen PJ. 2004. Contextual modulation of amygdala responsivity to surprised faces. J Cogn Neurosci 16:1730–1745.

Lang PJ, Bradley MM, Cuthbert BN. 1999. International Affective Picture System (IAPS): technical manual and affective ratings. Gainesville, FL: NIMH Center for the Study of Emotion and Attention, University of Florida.

Lang PJ, Bradley MM, Fitzsimmons JR, Cuthbert BN, Scott JD, Moulder B, Nangia V. 1998. Emotional arousal and activation of the visual cortex: an fMRI analysis. Psychophysiology 35:199–210.

Lavie N. 1995. Perceptual load as a necessary condition for selective attention. J Exp Psychol Hum Percept Perform 21:451–468.

McCarthy G, Wood CC. 1985. Scalp distributions of event-related potentials: an ambiguity associated with analysis of variance models. Electroencephalogr Clin Neurophysiol 62:203–208.

Miyoshi M, Katayama J, Morotomi T. 2004. Face-specific N170 component is modulated by facial expressional change. Neuroreport 15:911–914.

Morris JS, Friston KJ, Buchel C, Frith CD, Young AW, Calder AJ, Dolan RJ. 1998. A neuromodulatory role for the human amygdala in processing emotional facial expressions. Brain 121:47–57.

Öhman A, Flykt A, Esteves F. 2001. Emotion drives attention: detecting the snake in the grass. J Exp Psychol Gen 130:466–478.

Palmer SE. 1975. The effects of contextual scenes on the identification of objects. Mem Cognit 3:519–526.

Phelps EA, O'Connor KJ, Gatenby JC, Gore JC, Grillon C, Davis M. 2001. Activation of left amygdala to a cognitive representation of fear. Nat Neurosci 4:437–441.

Picton TW, Bentin S, Berg P, Donchin E, Hillyard SA, Johnson Jr R, Miller GA, Ritter W, Ruchkin DS, Rugg MD, Taylor MJ. 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37:127–152.

Pizzagalli DA, Lehmann D, Hendrick AM, Regard M, Pascual-Marqui RD, Davidson RJ. 2002. Affective judgments of faces modulate early activity (~160 ms) within the fusiform gyri. Neuroimage 16:663-677.

Puce A, Allison T, McCarthy G. 1999. Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. Cereb Cortex 9:445-458.

Puce A, Syngeniotis A, Thompson JC, Abbott DF, Wheaton KJ, Castiello U. 2003. The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. Neuroimage 19:861-869.

Rolls ET, Aggelopoulos NC, Zheng F. 2003. The receptive field of inferior temporal cortex neurons in natural scenes. J Neurosci 23:339-348.

Rousselet GA, Mace MJ, Fabre-Thorpe M. 2004. Animal and human faces in natural scenes: how specific to human faces is the N170 ERP component? J Vis 4:13-21.

Séverac Cauquil A, Edmonds GE, Taylor M. 2000. Is the face-sensitive N170 the only ERP not affected by selective attention. Neuroreport 11:2167-2171.

Smith APR, Dolan RJ, Rugg MD. 2004. Event-related potential correlates of the retrieval of emotional and nonemotional context. J Cogn Neurosci 16:760-775.

Smith APR, Henson RNA, Dolan RJ, Rugg MD. 2004. fMRI correlates of the episodic retrieval of emotional contexts. Neuroimage 22:868-878.

Smith NK, Cacioppo JT, Larsen JT, Chartrand TL. 2003. May I have your attention, please: electrocortical responses to positive and negative stimuli. Neuropsychologia 41:171-183.

Stekelenburg JJ, de Gelder B. 2004. The neural correlates of perceiving human bodies: an ERP study on the body-inversion effect. Neuroreport 15:777-780.

Sugase Y, Yamane S, Ueno S, Kawano K. 1999. Global and fine information coded by single neurons in the temporal visual cortex. Nature 400:869-873.

Surguladze SA, Brammer MJ, Young AW, Andrew C, Travis MJ, Williams SC, Phillips ML. 2003. A preferential increase in the extrastriate response to signals of danger. Neuroimage 19:1317-1328.

Taylor SF, Luan Phan K, Decker LR, Liberzon I. 2003. Subjective rating of emotionally salient stimuli modulates neural activity. Neuroimage 18:650-659.

Tsivilis D, Otten LJ, Rugg MD. 2001. Context effects on the neural correlates of recognition memory: an electrophysiological study. Neuron 31:497-505.

Vuilleumier P, Richardson MP, Armony JL, Driver J, Dolan RJ. 2004. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. Nat Neurosci 7:1271-1278.

Watanabe S, Miki K, Kakigi R. 2002. Gaze direction affects face perception in humans. Neurosci Lett 325:163-166.

White M. 1995. Preattentive analysis of facial expressions. Cogn Emot 9:439-460.

Yi D-J, Woodman GF, Widders D, Marois R, Chun M. 2004. Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. Nat Neurosci 7:992-996.