How affective information from faces and scenes interacts in the brain

Jan Van den Stock, Mathieu Vandenbulcke, Charlotte B. A. Sinke, Rainer Goebel, and Beatrice de Gelder

1Department of Neurosciences, Division of Psychiatry, KU Leuven, Brain and Emotion Laboratory Leuven (BELL), 2Old Age Psychiatry, University Hospitals Leuven, 3000 Leuven, Belgium, 3Department of Cognitive Neuroscience, Maastricht University, 6229 ER Maastricht, 4Laboratory for Cognitive and Affective Neuroscience, Tilburg University, 5037 AB Tilburg, the Netherlands

INTRODUCTION

Social animals have developed a wide range of communication abilities on which their wellbeing in a society depends. Among the cognitive skills of social species is the ability to monitor each other’s behaviour and to adapt continuously to the social signals of others, whether collaborative or competitive (Darwin, 1872). In line with this, the human brain shows functional specialization to process emotional expressions displayed by conspecifics (Zhu et al., 2013). Although recognition of threatening events in the natural environment is at least equally critical for survival, affective neuroscience has traditionally focussed on perception of facial expressions. There is ample evidence that the brain contains dedicated mechanisms to process faces and facial emotions (Calder et al., 2011). However, recent reports show that the affective meaning of a face is modulated by contextual stimuli, for instance a background scene (de Gelder et al., 2006; de Gelder and Van den Stock, 2011; Van den Stock and de Gelder, 2012; Wieser and Brosch, 2012).

In the present study, we used fMRI to investigate the brain areas that are modulated by the affective valence of the visual background scene in which a face is presented. In previous ERP-studies, we observed an influence of the background scene on the face-responsive N170 in the left hemisphere (Righart and de Gelder, 2006, 2008), presumably originating in the fusiform gyrus (FG) (Pizzagalli et al., 2002; Iidaka et al., 2006). In addition, previous reports of affective context modulation on faces by preceding pictorial (Mobbs et al., 2006) and verbal (Kim et al., 2004) contexts reported activity in the left FG. Based on these findings, our hypothesis is that affective context modulation of facial expressions by background scenes will involve the left FG. Second, in line with our previous findings showing increased EBA activation for neutral bodies against a threatening compared to a neutral background (Van den Stock et al., 2012), we conjecture that affective context modulation for faces may involve face-responsive regions like the occipital (OFA) and fusiform face area (FFA) and superior temporal sulcus (STS) (Haxby and Gobbini, 2011).

Another type of influence on perception of faces is related to the observer. There is evidence that (non-clinical) personality traits like empathy, neuroticism and anxiety have an effect on behavioural (Wieser et al., 2009a,b) and neural processing (Hooker et al., 2010) of social emotional expressions (Fox and Zougkou, 2011). We included a measure of trait empathy to investigate its influence on affective context modulation.

MATERIALS AND METHODS

Participants

Fifteen subjects (six male; age 26.2 ± 5.9 years; all right-handed) participated in the experiment. One subject was excluded due to excessive movement in the scanner.

Stimuli

Backgrounds

Scenes of happy, threatening, neutral, sad or disgusting everyday situations were taken from our own database (Van den Stock et al., 2012). We selected three familiar scene categories (buildings, cars and landscapes) that involve the same objects but with different affective significances. Examples of stimuli are: a house on fire or a holiday cottage, a damaged car in an accident or a shiny new convertible. None of the scenes displayed humans or animals. In a pilot study the scenes were presented one by one for 4000 ms with a 4000-ms inter-stimulus interval. Participants were instructed to categorize as accurately and as quickly as possible the pictures according to the emotion they evoked (anger, fear, happiness, disgust, sadness or neutral). Based on the results, we selected 24 scenes (12 threatening and 12 neutral) for the present experiment (all recognized correctly above 70%). Each category contained four exemplars with a car, four with...
a building and four landscapes. We created scrambled versions of every scene, by dividing the image in 10000 (100 × 100) squares and randomly rearranging the squares.

Faces
We selected 24 different face-stimuli (half male; half neutral, half fearful) from the Karolinska Directed Emotional Face database (Lundqvist et al., 1998).

Compound-stimuli
Scenes were combined with faces to create compound stimuli. In the centre of every scene, a face was overlaid. To avoid the impression of a ‘floating’ face, a geometrical torso-like shape was positioned underneath the face. Every scene stimulus was once overlaid with a fearful male face, once with a neutral male face, once with a neutral female face, resulting in 96 realistic compound stimuli. We used the scrambled images as controls for the scenes and triangles as control for the faces. The 24 scrambled scenes were combined once with a fearful face and once with a neutral face, leading to an additional 48 compound stimuli. We also paired every scene with a white triangle (24 intact scenes + 24 scrambled scenes). These scene-triangle combinations were used as a control condition instead of only the scenes in order to maintain the same task in all conditions (see below) and to ensure all stimuli had a clear foreground/background structure. This procedure results in 192 compound stimuli. For examples, see Figure 1.

Procedure stimuli were presented in blocks of 9000 ms, separated by fixation blocks of 15 750 ms. Within a block, eight stimuli were presented for 800 ms with an inter-stimulusinterval of 370 ms, during which a grey screen was shown. In fixation blocks, a grey screen with a black fixation cross was presented. We used a 3 (face: fearful, neutral and triangle) × 3 (scene: threatening, neutral and scrambled) factorial design. Participants were given an oddball detection task and instructed to press the response button when a stimulus was shown upside-down. A run lasted 711 s and consisted of 31 experimental blocks and 32 fixation blocks. The order of the blocks was randomized. In 4 of the 31 blocks (13%) an oddball stimulus occurred, while the remaining 27 blocks were divided in 3 blocks of every condition. The experiment consisted of four runs. The methods are also described in Sinke et al. (2012). The study was performed in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the faculty of Psychology and Neuroscience (ECP Maastricht, the Netherlands).

We used a validated Dutch translation (Lijffijt et al., 2005) of the widely used 17-questionnaire developed by Eysenck and colleagues (1985) as a measure of self-reported empathy. The questionnaire consists of 54 dichotomous items and includes three scales: impulsiveness (19 items), venturesomeness (16 items) and empathy (19 items). One subject did not complete the questionnaire.

RESULTS
Imaging data were analysed using BrainVoyager QX (Goebel et al., 2006). The first two volumes of every functional run were discarded to allow for T1 equilibration. Pre-processing of the functional data included slice scan time correction (cubic spline interpolation), 3D motion correction (trilinear/sync interpolation), temporal filtering [high-pass GLM (General Linear Model) – Fourier of 2 sines/cosines] and Gaussian spatial smoothing (6 mm). Functional data were then co-registered with the anatomical volume and transferred into Talairach space.

The statistical analysis was based on the GLM, with each condition defined as a predictor plus one for the oddball.

Whole-brain analysis
As categorical effects are typically larger than emotional effects, we used a threshold of \( P < 5 \times 10^{-4} \) (uncorrected) for the categorical comparisons and a threshold of \( P < 5 \times 10^{-3} \) (uncorrected) for the emotional comparisons (see also Van den Stock et al., 2012). We first examined category selective responses. All face-conditions compared with all triangle conditions revealed activation in the well-known face-selective (FFA, OFA) areas as well as the amygdala. All scene conditions compared with all scramble conditions activated the expected scene areas: parahippocampal place area (PPA) (Epstein and Kanwisher, 1998), retrosplenial cortex (RSC) (Bar and Aminoff, 2003) and the transverse occipital sulcus (TOS) (Hasson et al., 2003). Second, we identified the regions showing within-category emotion effects. Fearful compared

![Fig. 1 Schematic overview and stimulus examples of the conditions in the 3 (face) × 3 (background) factorial design.](image-url)
with neutral faces triggered activation in the right ventral premotor cortex and bilateral STS, while threatening vs neutral backgrounds activated the posterior PPA and region of the extrastriate body area (EBA) (Downing et al., 2001) in line with previous findings (Sinke et al., 2012; Van den Stock et al., 2012). The reverse contrasts did not reveal any significant activation. The results are displayed in the Supplementary Materials.

**Context effects on faces**

We identified the brain regions that respond differentially to faces in a threatening compared with a neutral background, controlled for isolated background effects: 

\[(\text{fearful face in threatening background} + \text{neutral face in threatening background}) > (\text{fearful face in neutral background} + \text{neutral face in neutral background})] > [(\text{triangle in threatening background} > \text{triangle in neutral background})].

Surprisingly, this revealed no significant results. We also made the reverse comparison to determine the brain regions that respond differently to faces in a neutral compared with a threatening background: 

\[(\text{fearful face in neutral background} + \text{neutral face in neutral background}) > (\text{fearful face in threatening background} + \text{neutral face in threatening background})] > [(\text{triangle in neutral background} > \text{triangle in threatening background})].

This revealed activity in the bilateral subgenual anterior cingulate cortex (sgACC) in the ventromedial prefrontal cortex (vmPFC) (Figure 2) and right anterior PPA (Figure 3).

To investigate the face-specificity of these effects, we performed an additional analysis on a dataset obtained with a similar study that used bodies instead of faces (Van den Stock et al., 2012). Note that the stimulus materials and design were exactly the same as in the present study, except for the face stimuli and participants. The results of the contrast 

\[(\text{fearful body in neutral background} + \text{neutral body in neutral background}) > (\text{fearful body in threatening background} + \text{neutral body in threatening background})] > [(\text{triangle in neutral background} > \text{triangle in threatening background})]

revealed activity in the right PPA, slightly overlapping the region obtained with the faces (8/368 voxels = 2%), but not in the sgACC, even under a more liberal threshold ($P < 0.10$, uncorrected) (Figure 3).

Summarizing the results from the present and previous (Van den Stock et al., 2012) study for the PPA, we observed:

- More activation for threatening than for neutral backgrounds in the right posterior PPA (current study and Van den Stock et al., 2012)
- More activation for neutral than for threatening backgrounds in the left anterior PPA (Van den Stock et al., 2012)
- More activation for neutral than for threatening backgrounds containing a person (face or body, irrespective of emotion) in the right anterior PPA (current study and Van den Stock et al., 2012)

**Context effects on facial expressions**

Next, we investigated background effects that are selective for facial expressions of fear. We identified the brain regions that responded more to fearful faces in a threatening compared with a neutral background, controlled for the same background effect on neutral faces: 

\[(\text{fearful face in threatening background} > \text{fearful face in neutral background}) > (\text{neutral face in threatening background} > \text{neutral face in neutral background})].

Note that this contrast also reflects the congruent vs incongruent face–background combinations. This revealed activity in the left FG. We examined whether this region falls within the face-responsive area of the FG by overlaying the smoothed left FFA obtained in another study (Frost and Goebel, 2012). The results are displayed in Figure 4 and show that the activation partly overlaps the posterior region of the face-responsive left FFA.

**Region of interest analysis**

We correlated the beta-differences of all the contrasts described above with the personality scales Empathy, Introversion and Venturesomeness obtained with the I7-questionnaire. This revealed only one significant correlation, i.e. between Empathy and the...
DISCUSSION

The primary research question of the present study was to identify the brain regions that are sensitive to affective valence of the visual background in which a face is presented. The results reveal a nuanced pattern of face–background interactions. First and contrary to our expectations, we found no evidence of brain areas responding more to faces in a threatening background than to faces in a neutral background, controlled for isolated background effects. On the other hand, the sgACC and right anterior PPA responded more to neutral than threatening backgrounds surrounding faces, but not surrounding geometrical figures (triangles). Additionally, the results show that while similar background effects also obtain on bodies in the right anterior PPA, the effect in sgACC is face-specific and furthermore correlates with the empathic ability of the observer: higher empathy scores are related to smaller background effects on faces. Moreover, empathy modulates the direction of the context effect: a high empathy score is associated with increased activation for threatening compared with neutral backgrounds, while lower empathy scores reveal more activity for neutral than threatening backgrounds. Below, we discuss the main findings point by point.

sgACC

As shown in Figure 2, the activity reported in the sgACC reflects reduced negative activity and not positive activity. This is not an uncommon finding in the medial prefrontal cortex (mPFC). It has been suggested that during baseline condition or ‘rest’, self-reflexive thoughts that activate mPFC are triggered, which may explain reductions in Blood Oxygenation Level Dependent responses for conditions of interest compared with baseline (Gusnard et al., 2001; D’Argembeau et al., 2005; D’Argembeau et al., 2007).

The subgenual portion of the ACC has been associated with emotional processing (Whalen et al., 1998; Bush et al., 2000), also in the context of faces (Bzdok et al., 2012). For instance, a recent report shows that matching facial expressions triggers activity in the sgACC compared with matching geometrical shapes (Ball et al., 2012). The present results add to these findings as they provide evidence that the face-specific sgACC response is modulated by the valence of the visual background scene. This is compatible with the notion that the sgACC processes face-specific information that goes beyond mere decoding of the facial expression and includes processing of contextual information like the visual background, concurrently presented faces or animacy (Ball et al., 2012; Van den Stock et al., 2013). The correlation with personality traits like empathy as we observe here and social anxiety as reported by Ball et al. (2012) is in line with this argumentation. By this logic, one would expect that the sgACC activation is not only modulated by the personality of the observer, but also by his emotional state. Indeed, the sgACC has been implicated in emotion regulation in the normal population (Zald et al., 2002; Kohn et al., 2013), but also in the context of affective disorders (Davidson et al., 2002; Etkin and Wager, 2007). For instance, reductions of both structure and function of the sgACC have been reported in depression (Drevets et al., 1997), while deep brain stimulation of the sgACC is associated amelioration of depressive symptoms (Mayberg et al., 2005). The emotion regulation function of the sgACC also applies to face-related activation, as abnormal sgACC activity to faces has been reported in mood disorders (Lennox et al., 2004; Gotlib et al., 2005; Drevets et al., 2008; Haldane et al., 2008).

A psychological explanation for the present results pertains to the notion that it reflects cognitive mentalizing processes. These presumably occur at the implicit level, as the task (oddball detection) was unrelated to social cognition (Frith, 2012; Frith and Frith, 2012). It may be that faces in a neutral scene elicit cognitive state attributions
What is he thinking?), while faces in threatening contexts primarily activate orienting responses (‘What is going on?’). The observed activity pattern is compatible with this hypothesis and also in line with clinical findings: syndromes characterized by deficits in empathy or other forms of social cognition like autism (Masten et al., 2011), schizophrenia (Koo et al., 2008) and fronto-temporal dementia (Rosen et al., 2005) have been associated with sgACC cortex abnormalities.

An alternative but not incompatible hypothesis relates to top–down mechanisms of object recognition. It has been suggested that perception of objects (for instance faces) triggers context-related processes, inter alia in the mPFC (Bar, 2003, 2004; Kveraga et al., 2011). Possibly, intrinsic contextual associations that are presumably richer for faces than for geometrical figures influence the visual background analysis. The results from the region of interest analysis are consistent with this hypothesis in the sense that faces may trigger stronger contextual associations in more empathic individuals, which may in turn reduce the effect of physically presented visual background. However, processing of contextual associations has been primarily associated with PPA and RSC and we will return to this notion below.

Right anterior PPA

Although the bar plot in Figure 2 reveals that the overall activation is higher for scenes containing triangles than for scenes containing faces, there is no significant effect of the affective valence of the background for triangles, contrary to faces. In other words, threatening backgrounds reduce activation in this area, but only when a face is present. However, it is less likely that this reflects face-specific processes, as this area partly overlapped with the same effect for bodies that we observed in a previous study (Van den Stock et al., 2012).

One possible explanation is that this region primarily sub serves topographic processing (Henderson et al., 2008; Kravitz et al., 2011; Troiani et al., 2012). It may be that the presence of a face triggers non-topographic processes inhibiting the spatial computations as they obtain for scenes without faces and that this is more pronounced when faces are presented against a threatening compared with a neutral background.

Another hypothesis is that context modulation is related to stronger affective contextual associations of faces and bodies, compared with geometrical figures. It has been reported that objects (Bar and Aminoff, 2003), faces (Bar et al., 2008a) and scenes (Bar et al., 2008b) that have elaborate contextual associations activate the PHC more than objects, faces or scenes with fewer contextual associations. In line with this, a recent study shows that the PPA is responsive to the object content of a visual scene (Harel et al., 2012). However, the finding that triangles trigger higher activation in the PPA is not compatible with the notion that faces and bodies presumably have more elaborate and stronger contextual associations than triangles.

The PPA as defined in the present study (all scenes compared with scrambled scenes) occupies a rather extensive area (see Figure 2 and Supplementary Materials). It is unlikely that this region as a whole constitutes a functional unit (Epstein, 2008). Although the PPA has been studied extensively, there have been few reports on functional subdivisions within. The present results add to our previous findings showing a posterior PPA activation for threatening compared with neutral scenes (Van den Stock et al., 2012) and additionally show that the more anterior region shows a reversed pattern as well as sensitivity to object content.

Left FG

One area in the left FG that falls largely within the face selective part of the FG responded to the face emotion–background emotion interaction with increased activity for threatening compared with neutral backgrounds on fearful facial expressions, while the background effect on the neutral faces was in the opposite direction with higher activation for neutral than for threatening backgrounds. This supports the hypothesis that the left FG is involved in processing the affective face–context congruency as has been previously reported for pictorial (Mobbs et al., 2006) and verbal semantic contexts (Kim et al., 2004).
It is also consistent with electrophysiological measures of visual background context effects (Righart and de Gelder, 2006, 2008).

**Right posterior PPA**

A remarkable result of the present study is that nearly all regions (except the left FG) responding to face–background combinations show higher activation for neutral than for threatening backgrounds. This may reflect that threatening information conveyed by the environment is little influenced by categorical information contained by the stimulus like faces or bodies. However, as threatening scenes presumably engage more processes related to attention and emotion (Vuilleumier, 2005) than neutral scenes, we expected that this would be reflected in activation of respective brain areas. The results reveal only two extrastriate visual areas, i.e. the region of the right EBA and posterior PPA (Sinke et al., 2012; Van den Stock et al., 2012). Little is known about how the affective information present in a complex scene is processed, when controlling for the object content of a scene. Our findings provide an indication that at least two areas respond to the affective content of a scene, although an important question concerns the time point of activation. It has been reported that the posterior PPA is activated at different time points and for different task demands during scene perception with an early activation for passive scene perception and later activation during a spatial judgement task (allocentric judgements) (Bastin et al., 2012). Methods with higher temporal resolution like ERP or MEG may provide important insights into the mechanisms underlying content-constrained affective scene perception.

**Amygdala**

In line with other recent reports (Zhu et al., 2013), we observed face-selective activity in the amygdala (fearful + neutral faces vs triangles), but not for fearful vs neutral faces. We believe that the absence of amygdala activation for fearful vs neutral faces contrast is at least partly due to a methodological factor. It has been reported that amygdala are subject to rapid habituation (Zald, 2003) and we used a blocked design, which is more susceptible to adaptation effects compared with an event-related design. However, ever since the landmark case study by Adolphs et al. (1994), facial emotion research has strongly focused on the amygdala, which may have resulted in a confirmatory publication bias. Although there is general consensus that the amygdala is involved in fear processing, there is an increasing number of reports challenging its presumed critical role (van der Gaag et al., 2007; Tsujiya et al., 2009; Terburg et al., 2012; Van den Stock et al., 2012; Zhu et al., 2013; Edmiston et al., 2013; Freeman and Luby, 2013). The present results add to this body of evidence and argue for a non-necessary involvement of the amygdala in fear processing.

**CONCLUSION AND FUTURE DIRECTIONS**

The present findings show that affective face–background congruence is primarily processed in the left FFA. In addition, the results reveal a pattern of complex scene perception with an increasing functional specialization along the posterior–anterior axis. The areas that respond to the affective content conveyed by the background are situated at the posterior extrastriate visual cortex: EBA and posterior PPA. This effect is modulated more anteriorly by the object content of the scene in the anterior PPA. A similar posterior to anterior gradient of object sensitivity has been reported in the medial temporal lobe (Liang et al., 2013) and this has been related to episodic encoding of an event (Preston et al., 2010). This may be particularly relevant for the understanding of the functional organization of occipito-temporal cortex relating to real-world perception. The reports of posterior–anterior gradient relating to object content are paralleled by studies reporting a mediolateral gradient relating to real world object size, with larger objects represented more medially (Konkle and Oliva, 2012). The present results are in line with this as the anterior PPA region we observe for faces is located laterally to the neighbouring region we observed for bodies. It has been postulated that this functional organization is shaped by a congenital connectivity to similar domain-specific regions (Mahon and Caramazza, 2011). The present results extent this notion by revealing that the object sensitivity gradient is modulated more anteriorly in the sgACC by the personality of the observer, i.e. the empathic ability.

We have previously provided evidence that perception of facial expressions is also modulated by other contexts, for instance bodies (Meeren et al., 2005; Van den Stock et al., 2007; Kret et al., 2013) and these findings have been confirmed by subsequent studies in other labs (Aviezer et al., 2008). In addition to bodies, other face-contexts include voices (de Gelder et al., 2006; de Gelder and Van den Stock, 2011). It would be interesting to investigate whether similar context effects apply. We have recently provided evidence that recognition memory for neutral faces is more influenced by affective bodies than by affective backgrounds (presented during encoding). Considering the face-specificity of the sgACC effect we observe here and the rapid integration of face–body stimuli (Meeren et al., 2005), we hypothesize that this region may respond stronger to body than to background scene modulation of faces. In addition clinical studies in populations with sgACC abnormalities would provide valuable insights in the role of this region in face–context integration.

**SUPPLEMENTARY DATA**

Supplementary data are available at SCAN online.

**Conflict of Interest**

None declared.

**REFERENCES**


Threat perception in faces and scenes


