UNIVERSITY OF MILANO-BICOCCA
DEPARTMENT OF PSYCHOLOGY

DOCTORAL PROGRAM IN EXPERIMENTAL PSYCHOLOGY,
LINGUISTICS AND COGNITIVE NEUROSCIENCE
XXIV CYCLE

NEURAL CORRELATES OF FACE EVALUATION: EMOTIONAL EXPRESSIONS AND SOCIAL TRAITS

Tutor: Professor COSTANZA PAPAGNO

Doctoral thesis by:
GIULIA MATTAVELLI

ACADEMIC YEAR 2011/2012

That face, that face, that wonderful face!

It shines, it glows all over the place.

And how I love to watch it change expressions.

Each look becomes the pride of my possessions.

(That Face, Lew Spence and Alan Bergman)

Table of contents

Summary

Face processing is a crucial skill for human interaction. Accordingly, it is supported by a widely distributed fronto-temporo-occipital neural circuit (Haxby et al., 2000). The present work investigates the neural correlates of face expression processing by means of different neuroimaging and electrophysiological techniques. Using fMRI I investigated amygdala responses to basic emotions and activations in face-selective regions in response to social cues detected in faces (Study 1 and Study 2). These studies showed that the amygdala is highly responsive to fear expressions but has also a critical role in appraising socially relevant stimuli and together with the posterior face-selective regions it is sensitive to face distinctiveness as well as social meaning of face features. In Study 3 I demonstrated by means of TMS that the medial prefrontal cortex (mPFC) contains different neural representations for angry and happy expressions linked to lexical knowledge of emotions. Finally, the combined TMS-EEG experiment reported in Study 4 revealed interconnections between activity in the core and the extended system of face processing, and the interactions resulted to be modulated by the type of behavioural task.

Taken together the present results help to clarify the role of different regions as part of the face perception system and suggest that the coupling between cortical areas and the coordinated activity of different regions in the distributed network are crucial to recognize the multiplicity of information that faces convey.

Introduction

Face perception is a particularly high-developed skill in humans. Faces are multi-dimensional stimuli and together with specific features linked with personal identity they convey crucial information for social interaction and adaptive behaviour (Bruce and Young, 2012). Indeed, when presented with faces people are highly skilled at decoding face expression and making social judgements even with exposure times of few milliseconds (Kirouac and Doré, 1984; Edwards, 1998; Willis and Todorov, 2006). Converging evidence from neuroimaging, electrophysiological (Vuilleumier and Pourtois, 2007; Davidson and Irwin, 1999) and lesion studies (Adolphs et al., 1994; 1996; Philippi et al., 2009) have allowed identifying a cortical-subcortical neural network involved in the perception and processing of facial expressions. These circuits include cortical areas primarily implicated in face processing like the inferior occipital cortex, the superior temporal sulcus and the fusiform cortex, as part of the core system of face perception (Haxby et al., 2000), and the amygdala, prefrontal cortex, insula and cingulate cortex as part of the extended system involved in emotion and semantic processing (Adolphs, 1999). The present research is focused on the neural correlates of face expression recognition and processing of social traits. Indeed, despite the growing literature concerning affective and social neuroscience (Dalgleish, 2004; Haxby et al., 2002), many questions remain unsolved. For example, although the role of the amygdala in emotion expression recognition is well-established (Calder et al., 2001; Cristinzio et al., 2007), it is still unclear whether this subcortical nucleus is involved in processing emotional stimuli in general or selectively in threatening and negative stimuli (Adolphs, 2002); moreover, neuroimaging studies have provided mixed evidence concerning the amygdala role in perception of social cues in faces (Said et al., 2011). Therefore, by means of two fMRI studies I have investigated brain responses to faces displaying different basic emotions and faces varying in perceived trustworthiness, with the

analyses focused on activation in the amygdala and in the posterior face-selective regions (Study 1 and Study 2). Another issue of debate concerns how emotional expressions are represented in cortical regions of the extended system like the medial prefrontal cortex (mPFC) and the somatosensory cortex (SC). Studies on brain-damaged patients showed impairment in face expression recognition following damage to these cortical areas (Adolphs et al., 2000; Hornak et al., 2003; Heberlein et al., 2008). However the mechanisms underlying these deficits remain unclear as the possibility that different types of emotion are represented in distinct neural circuits (Adolphs et al., 1996; Phan et al., 2002). In Study 3, I have addressed these issues by means of Transcranial Magnetic Stimulation (TMS). A priming task was used to modulate the activation state of the mPFC and the right somatosensory cortex (rSC) during emotional expressions discrimination and to analyse how TMS interacts with the ongoing neural activity depending on the type of emotion presented. Finally, only few studies have directly investigated connections among the areas part of the distributed face perception system (Summerfield et al., 2006; Fairhall and Ishai, 2007) and there is still disagreement about the organization and the interactions within the network (Pessoa and Adolphs, 2010). Recently, the development of the combined TMS-EEG technique has introduced a non-invasive method to measure directly and with high temporal resolution cortical excitability and effective connectivity among areas (Rosanova et al., 2012). Thus, in Study 4, I have explored by means of TMS-EEG whether the local cortical excitability in the mPFC and long-range connections between the extended and the core system of face perception are modulated by different behavioural tasks of face processing.

Neural system for perception and evaluation of faces

In the last decade, the neuroscience literature on face perception has been largely influenced by the neural model proposed by Haxby et al. (2000). Based on functional neuroimaging evidence, this model includes the inferior occipital gyri, the superior temporal sulcus and the lateral fusiform gyrus in a core system specialized in face processing, with a specific role in the perceptual analysis and structural encoding of face stimuli. Other areas located in different brain regions are included in an extended system, which is not selectively involved in face processing, but contributes in representing additional cues from faces like semantic information about the personal identity, interpretation of emotions and lip-reading (see Figure 1).

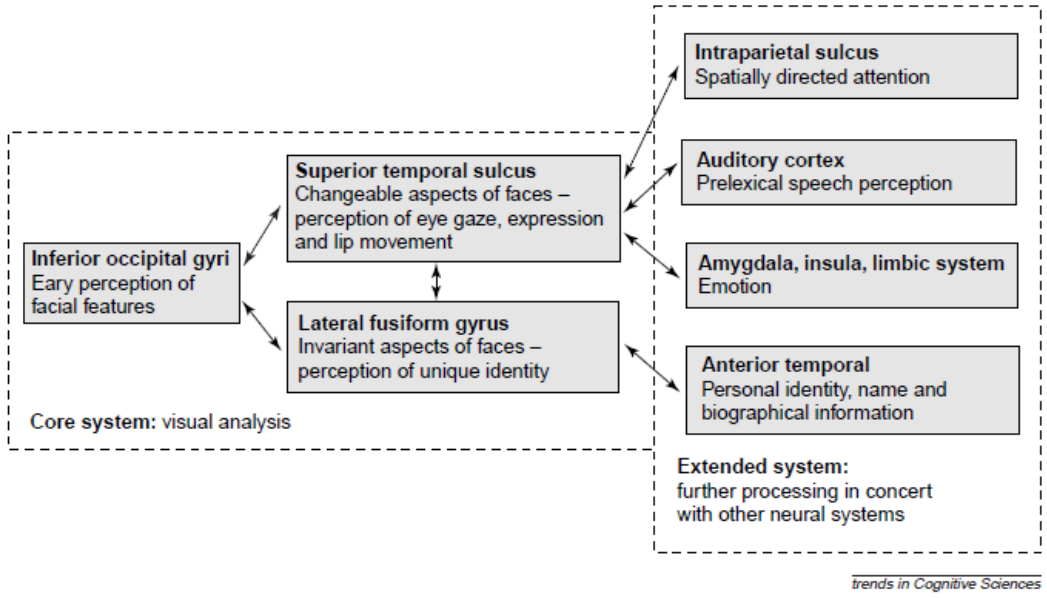

Figure 1. Neural model of the face perception system from Haxby et al., 2000.

The hierarchical organisation of this model assumes that the extended system extracts information from faces following the early visual processing mediated by the core system. Moreover, Haxby et al.'s (2000) neural model was inspired by the cognitive model of Bruce and Young (1986), which, supported by data from behavioural experiments and prosopagnosic patients, distinguished between processes involved in identity and expression

recognition. Indeed, it is hypothesized that the areas part of the core system differently contribute in representing changeable rather than invariant features of faces, which constitute two cognitive independent and anatomically dissociable aspects of face processing. In particular, the superior temporal sulcus is considered specifically involved in representing changeable cues like expression, eyes gaze or lip-movement whereas the lateral fusiform gyrus contributes to discriminate facial identity. Then, the evaluation of facial expressions depends on the activity of regions more generally associated with emotions like the amygdala, the insula, the somatosensory and orbitofrontal cortices, which are also involved in reactions to emotional stimuli and reward mechanisms (Haxby et al., 2002).

The Haxby et al.'s (2000) model is still the dominant framework for face perception studies and the critical role of the neural correlates evidenced above has been demonstrated by a wide number of neuroimaging and neuropsychological data (Adolphs, 2002; Haxby et al., 2002; Rossion et al., 2003; Grill-Spector et al., 2004; Andrews and Ewbank, 2004). However, recent studies have reported that activity in the posterior face-selective regions can be affected by attentional, cognitive and emotional modulation (Vuilleumier and Pourtois, 2007), questioning the hypothesis of a hierarchical information processing. This provides support to a new functional model (Ishai, 2008), which assumes the synchronised activity of multiple regions within the face processing network. Based on functional neuroimaging studies investigating effective connectivity (Summerfield et al., 2006; Fairhall and Ishai, 2007), Ishai (2008) predicted feed-forward and top-down connections modulated by the type of stimuli presented and the facial features requiring specific evaluation. For example, connectivity between the superior temporal sulcus and frontal regions depends on processing of animated faces; famous and attractive faces increase connectivity between the orbitofrontal and the fusiform gyrus, whereas the amygdala and the fusiform gyrus enhance interactions when emotional faces are presented.

Amygdala and basic emotions

The amygdala is one of the key components in the circuit of emotional face perception. Neuropsychological studies have demonstrated that patients with amygdala damage are impaired in emotion recognition (Adolphs et al., 1994; Anderson et al., 2000; Young et al., 1995; 1996). Deficits are reported as particularly severe for fear perception (Adolphs et al., 1994; Calder et al., 1996; Broks et al., 1998), and are often accompanied by an attenuated experience of fear and a reduced reaction to potential threats (Broks et al., 1998; Sprengelmeyer et al., 1999; Feinstein et al., 2010). Functional neuroimaging studies support the hypothesis that the amygdala is involved in processing fearful expressions and threatening stimuli (Calder et al., 2001). In particular, greater amygdala activation has been found when fearful faces were presented as compared with happy (Morris et al., 1996; 1998), angry (Whalen et al., 2001) or neutral faces (Phillips et al., 1998a).

Despite this converging evidence from neuropsychology and functional brain imaging, it is still a matter of debate whether the role of the amygdala in evaluating emotional expressions is specific for fear. For example, in patients with amygdala lesions, the deficit is not usually restricted to fear; most patients show impaired recognition of more than one emotion, even though the deficit in fear recognition tends to be the most severe (Adolphs et al., 1999). Accordingly, neuroimaging studies have reported amygdala activation for emotions other than fear, including sadness (Blair et al., 1999) and happiness (Breiter et al., 1996). Indeed, some neuroimaging studies report amygdala responses to several facial expressions without any specific effect of emotion type (Fitzgerald et al., 2006; Winston et al., 2003), a pattern that might be consistent with the idea that the amygdala is activated by emotionally salient stimuli and is involved in detecting relevant stimuli regardless of their positive or negative valence (Sander et al., 2003).

Amygdala and social traits

Beyond basic emotions faces can convey other crucial information for social interaction like cues linked with personality traits. For example, judgements of trustworthiness from facial appearance are extremely consistent across different observers, and can even be made with very brief presentations (Bar et al., 2006; Willis and Todorov, 2006). Theories of social perception link the fast evaluation of trustworthiness to a more general conception of primate behaviour in which individuals who are part of a social group are evaluated for potential threat (warmth, or approachability) and their capacity to enact any such threat (competence, or dominance) (Fiske et al., 2007; Todorov, 2008). In these models the evaluation of trustworthiness is closely linked to approachability; indeed ratings of trustworthiness and approachability were highly correlated (Oosterhof and Todorov, 2008; Santos and Young, 2008).

Neuropsychological studies have shown a role for the amygdala in processing trustworthiness and approachability (Adolphs, 1999; Adolphs et al., 2002; Cristinzio et al., 2007). Patients with amygdala damage rate untrustworthy-looking faces as more approachable and trustworthy than do neurologically unimpaired participants, consistent with a more general problem in the evaluation of potential threat and risk in the environment (Adolphs et al., 1998; Feinstein et al., 2010). The role of the amygdala in evaluating trustworthiness has also been supported by functional neuroimaging studies, though with mixed results. Early studies showed greater response in the amygdala for untrustworthy as compared to trustworthy faces (Winston et al., 2002) with a linear trend in amygdala activation for increasing untrustworthiness (Engell et al., 2007). Other studies have found U-shaped, quadratic responses in the amygdala (Said et al., 2008; Todorov et al., 2008), with increased responses at the extremes of the trustworthiness dimension; however, these studies found both linear and non-linear components in amygdala activation, preventing unequivocal

conclusions concerning how the amygdala processes this social dimension. Interestingly, U-shaped functions have been reported also in response to faces that vary along other social dimensions such as dominance (Said et al., 2010). These contrasting results lead to different interpretations of the role of the amygdala in social evaluation. A linear response is in line with the hypothesis that the amygdala is activated by arousing and threatening signals (Gläsher and Adolphs, 2003; Lane et al., 1997) and involved in evaluating the valence of negative stimuli (Todorov and Engell, 2008). On the other hand, a U-shaped quadratic pattern is more consistent with the hypothesis that the amygdala is more generally activated by salient social cues (Said et al., 2008). Thus, as seen above for the evaluation of basic emotion, it remains unclear whether there is a specific type of stimuli capable of eliciting higher amygdala activation, or its response reflects a more general evaluation of facial features relevant for social interaction.

The distributed cortical system

Together with the perceptual processing supported by the face-selective regions in the occipital and temporal lobes and emotional processing supported by the amygdala, recognition of facial expressions requires also a contribution from additional cortical regions involved in the cognitive evaluation and interpretation of the perceived expression (Haxby et al., 2002). Indeed, brain regions like the mPFC and the SC have been found to play a critical role in emotion discrimination (Dolan et al., 1996; Kesler/West et al., 2001; Adolphs, 2002; Winston et a., 2003). Several studies suggest that the prefrontal cortex is involved in emotional stimuli processing (Hornak et al., 1996), representation of affective states (Davidson and Irwin, 1999) and in processes that allow using emotional stimuli as cues for social behaviour (Damasio, 1994). In particular, the prefrontal cortex is connected with the amygdala and is thought to modulate emotional responses throughout cognitive control

(Hariri et al., 2000; Nomura et al., 2004). Neuropsychological studies support these hypotheses by demonstrating that patients with mPFC damage are impaired in recognizing emotional expressions and this deficit is associated with abnormal social behaviour (Mah et al., 2005) and reduced emotional responsiveness (Heberlein et al., 2008).

Recent theories of embodied cognition also emphasize the role of the SC. This area enhances facial emotion recognition through simulation processes of reactivation of somato-visceral responses associated with early acquisition and production of the perceived emotion (Niedenthal, 2007). Indeed, Adolphs et al. (2000), testing a large sample of brain-damaged patients, demonstrated that the integrity of the rSC was necessary for normal recognition of facial expressions. TMS studies have supported this conclusion by showing that stimulation of the rSC affects facial expressions discrimination (Pourtois et al., 2004; Pitcher et al., 2008).

Despite the evidence pointing to a crucial role of the mPFC and the rSC in facial emotion processing, one unsolved issue concerns whether these regions contain distinct neural circuits representing different types of emotion. Lesion studies have shown that brain damage can differently affect the ability to recognize specific emotions (Heberlein et al., 2008; Adolphs et al., 1996), but neuroimaging studies have provided conflicting results about the different circuits involved in the processing of specific facial expressions (Winston et al., 2003; Kesler/West et al., 2001). In particular, while evidence concerning amygdala and insula contribution respectively to fear and disgust processing is consistent (Morris et al., 1998; Adolphs et al., 1994; Phillips et al., 1997; Calder et al., 2001), the cortical areas involved in other basic emotions, like anger and happiness, are less clearly defined. For example, Kesler/West et al. (2001) found that angry faces activated the medial region of the superior frontal gyrus while happy faces activated the medial frontal/cingulate sulcus region; similarly, Phillips et al. (1998b) found a specific signal increase in the anterior and posterior

cingulated gyri and in the mPFC when happy facial expressions were presented, while no brain region showed signal increase for sad expressions. In a different fMRI study (Blair et al., 1999), without explicit emotion discrimination, the right orbitofrontal cortex responded to angry, but not sad faces. In contrast with these results of prefrontal activation for both happy and angry expressions, TMS applied over the mPFC increased response times in discriminating angry, but not happy faces (Harmer et al., 2001).

Concerning rSC, lesion and functional studies suggest that this area contributes to facial expression processing regardless emotion type (Adolphs et al., 2000; Winston et al., 2003). Consistently with this, repetitive TMS over rSC disrupted accuracy in discriminating all the six basic emotions (Pitcher et al., 2008). However, Pourtois et al. (2004) found that single pulse TMS over rSC selectively interfered with fear but not happy expressions. Conversely, happiness expressions were more affected compared to other emotions in a recognition task when subjects' facial mimicry was blocked by an irrelevant task (as bite a pen with the teeth or the lips) supposed to involve the rSC (Oberman et al., 2007). These contrasting results may depend on different emotions requiring different levels of somatic representation. The rSC activation may thus vary depending on the perceived facial expression. Accordingly, different TMS effects on rSC in emotional processing may depend on the interaction between the specific stimulation parameters (i.e. intensity, frequency) and the specific level of activation of the rSC region (Hussey and Safford, 2009).

Taken together these data contribute to describe the neural substrates that mediate emotion processing; however further studies could clarify whether different emotions are represented in segregated circuits and how they are functionally organized.

Faces in the cortical network

The influent Haxby et al.'s (2000) model suggests that the integrated activity of multiple regions, part of the distributed face perception system, is crucial for processing different facial features and achieve a comprehensive representation of face stimuli. Nevertheless, most neuroimaging studies focused on the functional selective role of each discrete brain area (Calder and Young, 2005), thus functional interactions within the system are still largely unknown.

However, recent fMRI studies emphasized the distributed nature of brain activation in response to presentation of faces and found larger activation in the temporal face-selective regions for emotional than neutral faces (Ishai et al., 2005), supporting the hypothesis of interactions between the face perception system and emotion processing (Vuilleumier and Pourtois, 2007). The functional association of different regions has been underlined also in fMRI studies which reported negative correlations between activity in the prefrontal cortex and the amygdala suggesting a modulatory influence of the prefrontal cortex (Hariri et al., 2000; Nomura et al., 2004). Although these studies showed important results about the functional network involved in emotion processing, they do not explain the connectivity and causal links between areas. These issues have been recently addressed in fMRI studies which used Dynamic Causal Modeling (DCM) to estimate brain activity considering not only the stimuli presented by the experimenter but also interconnections with other brain regions whose activity correlates with the task (Friston et al., 2003). Indeed, Summerfield et al. (2006) found that perceptual decision about faces rather than other objects enhanced feedback signals from the prefrontal cortex towards the amygdala and the fusiform face area; similarly, Fairhall and Ishai (2007) reported increased connectivity between the amygdala and the fusiform gyrus when emotional faces were presented. The combined participation of different structures from the early stage of face processing has been shown also during

intracerebral recording in epileptic patients, in which simultaneous responses in the fusiform gyrus and in the inferior frontal cortex have been detected during a face recognition task (Barbeau et al., 2008).

These studies provided new evidence concerning the integrated system which mediates face processing and have promoted the development of new neural models, which take into account bidirectional connections between the core and the extended system (Ishai, 2008). The investigation of interactions between different regions could be a key aim for future research, in order to understand the dynamic of the connectivity among areas and how stimuli and task type modulate the information processing in the system.

1. Study 1: Amygdala responses to basic emotions

1.1. Introduction

Recent meta-analyses have considered the patterns of findings across functional neuroimaging studies investigating emotional processing, without resolving the debate concerning whether the amydala plays a specific role in fear recognition (Phan et al, 2002; Vytal and Hamann, 2010), or only shows a general activation in response to emotional faces (Sergerie et al., 2008). What these meta-analyses do agree upon, though, is that interpretation is limited by the fact that there are few studies, which compare several expressions within a single experiment (Vytal and Hamann, 2010). Moreover, variable results in the literature may be due to the use of different non-emotional stimuli (neutral faces or non-face images) as comparison condition whose activation is subtracted from the emotion conditions (Sergerie et al., 2008).

In light of the above, the present study aimed at clarifying amygdala responses to faces displaying different emotions. Following Sergerie et al.'s (2008) advice, a block design was used to take advantage of the greater statistical power as compared to an event-related paradigm, and it was examined whether the adoption of different statistical thresholds affected the pattern of results. The fMRI block design experiment included expressions of six basic emotions (fear, anger, disgust, sadness, happiness and neutral), and two different comparison conditions. Pictures of buildings were used as a non-face contrast, to find amygdala responses to faces in general. Mildly happy (a 25% morph along the neutral to happy continuum) expressions were used to identify emotion-specific activations. The mildly happy face was used as previous studies have suggested that neutral faces can appear slightly cold and hostile (Ekman and Rosenberg, 1997). Finally, fear expressions were compared with the other emotions to test the hypothesis of a greater amygdala response to fearful faces.

1.2. Method

Participants

Twenty-four healthy volunteers (12 male, 12 female, mean age = 24.3 years, range 19-35) took part in the experiment. All participants were right-handed, with a western cultural background, and had normal or corrected to normal vision with no history of neurological illness. The study was approved and conducted following the guidelines of the Ethics Committee of the York Neuroimaging Centre, University of York. All participants gave written consent prior to their participation.

Stimuli

Face stimuli were greyscale images from the FEEST set (Young et al., 2002). Five models (F5, F6, F8, M1, M6) were selected on the basis of the visual similarity of the posed expression across different models, the percent recognition rate of each model's expression, and the similarity of the action units (muscle groups) used to pose the expressions. For each model the neutral pose and the expressions of fear, anger, disgust, sadness and happiness were used. An additional condition was created with faces with a 25% happy expression produced with computer manipulation by morphing the images along the neutral-happy continuum for each model (Calder et al., 1997). Previous studies have used an equivalent mild happiness expression as a contrast condition (Phillips et al., 1998a; 1999) because it looks more socially neutral than a completely neutral pose, which can appear slightly cold and a little hostile (Ekman and Rosenberg, 1997). Stimuli for the buildings condition were greyscale pictures of houses matched for luminance, size and resolution.

Imaging parameters

Scanning was performed at the York Neuroimaging Centre at the University of York with a 3 Tesla HD MRI system with an eight channels phased array head coil (GE Signa Excite 3.0 T, High resolution brain array, MRI Devices Corp., Gainesville, FL). Axial images were acquired for functional and structural MRI scans. For fMRI scanning, echo-planar images were acquired using a T2*-weighted gradient echo sequence with blood oxygen level-dependent (BOLD) contrast (TR = 3 sec, TE = 32.7 msec, flip-angle = 90°, acquisition matrix 128 x 128, field of view = 288 mm x 288 mm). Whole head volumes were acquired with 38 contiguous axial slices, each with an in-plane resolution of 2.25 mm x 2.25 mm and a slice thickness of 3 mm. The slices were positioned for each participant to ensure optimal imaging of the temporal lobe regions, where the amygdala is situated. T1-weighted images were acquired for each participant to provide high-resolution structural images using an Inversion Recovery (IR = 450 msec) prepared 3D-FSPGR (Fast Spoiled Gradient Echo) pulse sequence (TR = 7.8 sec, TE = 3 msec, flip-angle = 20°, acquisition matrix = 256 x 256, field of view = 290 mm x 290 mm, in-plane resolution = 1.1 mm x 1.1 mm, slice thickness = 1 mm). To improve co-registration between fMRI and the 3D-FSPGR structural a high resolution T1 FLAIR was acquired in the same orientation planes as the fMRI protocol (TR = 2850 msec, TE = 10 msec, acquisition matrix 256 x 224 interpolated to 512 giving effective in plain resolution of 0.56 mm).

fMRI Experiment

The experiment investigated brain responses to different basic emotions taking into account distinct comparison conditions. A block design was used with eight conditions comprising six basic emotions (fear, anger, disgust, sadness, happiness and neutral), a mildly happy face condition and a non-face condition (buildings). Within each block 5 images from each

condition were presented in a pseudorandom order for 1 second followed by a 200 msec fixation cross, giving a block duration of 6 seconds; blocks were interleaved with a 9 seconds fixation cross on a grey screen. Blocks corresponding to each of the eight conditions were repeated eight times in a counterbalanced order, resulting in a total of 64 blocks and scan duration of 16 minutes. A red spot detection task was used to monitor attention during the fMRI session. In one or two images per block a small red spot appeared; participants were instructed to look at the stimuli and press a response button whenever they saw the red spot. Other than this red spot detection task, the requirement was simply passive viewing of the stimuli. Experiments were run using Neurobehavioural System Presentation 13.0 software.

After the MRI scans, participants were asked to complete a behavioural task to check that they correctly recognized the facial expressions. The same stimuli used in the experiment were presented on a computer screen and participants were required to sort the face images into six emotional expressions (fear, anger, disgust, sad, happy and neutral).

fMRI data analysis

Image analyses were performed by means of FEAT (FMRI Expert Analysis Tool) Version 5.98, part of FSL (http://www.fmrib.ox.ac.uk/fsl). For each participant the following pre-statistic processing was applied: motion correction using MCFLIRT (Jenkinson et al., 2002), slice-timing correction using Fourier-space time-series phase-shifting, non-brain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel (FWHM 5 mm), grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 60 sec).

First level analyses were performed by modelling the hemodynamic BOLD response for each condition. The response to the six emotion conditions was compared to building and mildly happy control conditions. In addition, the response to fear was compared to each of the other emotion conditions. FLIRT (Jenkinson et al., 2001; 2002) was used to register participants' fMRI images onto their T1 FLAIR, then onto their high resolution T1 structural images, and finally onto Montreal Neurological Institute (MNI152 with $2mm^3$ voxels) standard space.

Statistical analysis at group level (higher level analysis) was carried out using FLAME (FMRIB's Local Analysis of Mixed Effects) stage 1 (Beckmann et al., 2003; Woolrich et al., 2004; Woolrich, 2008). Since the amygdala was a priori region of interest (ROI), the Harvard-Oxford sub-cortical probability atlas was used to anatomically mask the right and left amygdala at group level. This atlas represents each structure as a standard space image with value from 0 to 100, according to the cross-population probability of a given voxel being in that structure. Analyses were run using both liberal 5-100% amygdala masks and more conservative 50-100% amygdala masks to take into account any possible difference between amygdala responses and activations in the peri-amygdalar regions. Results of the significant activations within the amygdala are reported using a less conservative statistical threshold of $p < .005$ ($Z > 2.6$) and a more stringent threshold of $p < .001$ ($Z > 3.1$). Percent signal change in the masked amygdala was extracted for each condition using the Featquery tool. Activations in other brain regions based on a threshold of $p < .001$ ($Z > 3.1$) and a minimum extension of 20 voxels are also reported.

1.3. Results

Behavioural data

A post scan behavioural task was used to check that participants in the fMRI experiment correctly categorised the facial expressions. Results confirmed that all participants recognised the different expressions with high accuracy (90.1 % s.d. = 6.2, fear = 95.4 s.d. = 7.8, anger = 93.3 s.d. = 12, sadness = 80 s.d. = 18.2, disgust = 77.5 s.d. = 22.5, happiness = 98.3 s.d. = 3.8). The neutral face was also classified as neutral (96.3 s.d. = 8.2) and the 25% happy faces was mainly categorised as neutral (63.3 %) or happy (32.9%), consistent with their position on the neutral-happy continuum (Young et al., 1997).

fMRI analysis

Figure 1.1 shows the spatial extent of voxels in the amygdala that were more active when viewing faces posing emotional expressions compared to the control conditions. Results show a significant difference in the response to each expression compared to buildings within the amygdala. When emotion conditions were compared to the 25% happy face, the magnitude and extent of the significant voxels were less than the buildings contrasts, but still there was a significant activation for all conditions except sadness and neutral.

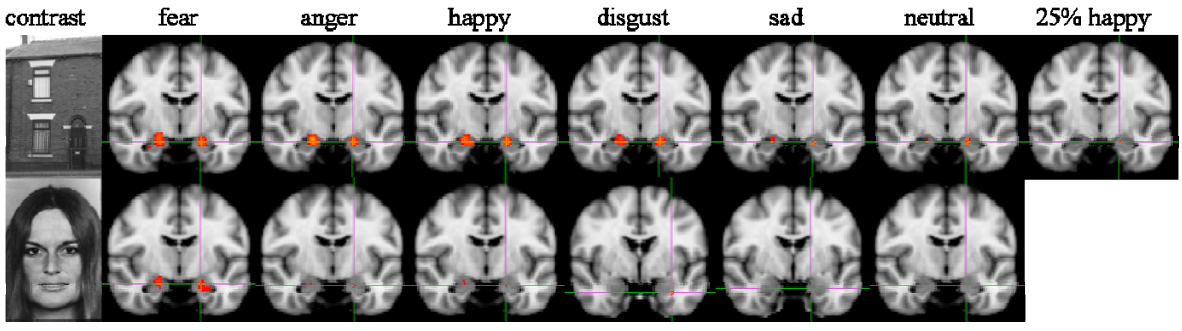

Figure 1.1. Statistical significance maps thresholded at $p < .001$ ($Z > 3.1$) depicting amygdala activation in each facial expression condition versus a buildings comparison condition (top line) and in each facial expression versus a 25% happy comparison condition (bottom line). The amygdala region is anatomically defined with the 5-100% masks from the Harvard-Oxford sub-cortical probability atlas. Images follow the radiological convention with the right hemisphere represented on the left side. For each contrast the cross cursor was positioned on the peak voxel in the left amygdala.

MNI coordinates of the response peaks and number of active voxels for each contrast are reported in Table 1.1 and Table 1.2. Analyses were carried out with two different statistical thresholds. Using a significance level of $p < .005$ ($Z > 2.6$) a large number of active voxels appeared in left and right amygdala when fearful, angry, happy and disgusted facial expressions were contrasted with buildings; significant activation was also observed for sad, neutral and 25% happy conditions but to a smaller extent. When a more stringent statistical criterion was applied, large activations remained significant for fearful, angry, happy and disgusted expressions, but only a reduced number of voxels crossed the $p < .001$ threshold ($Z > 3.1$) for sad, neutral and 25% happy conditions (see Table 1.2).

	condition	Left Amygdala				Right Amygdala			
		Z max	MNI (x y z)			Z max	MNI (x y z)		
building contrasts	fear	5.03	-20	-8	-18	4.64	20	-8	-18
	anger	5	-18	-10	-18	5.46	20	-8	-16
	happy	5.03	-18	-10	-18	5.27	20	-8	-16
	disgust	5.16	-18	-10	-18	4.45	16	-6	-16
	sad	3.64	-18	-10	-18	3.49	22	-10	-16
	neutral	3.61	-18	-10	-18	4.34	20	-6	-16
	25% happy	3.25	-18	-10	-18	3.35	18	-8	-18
25 happy contrasts	fear	3.73	-20	-8	-18	3.99	20	-8	-16
	anger	3.29	-20	-8	-20	3.25	24	-10	-16
	happy	3.24	-18	-8	-20	3.72	24	-10	-16
	disgust	3.17	-30	-2	-26	3.34	26	0	-26
	sad	-				-			
	neutral	-				-			
fear contrasts	anger	-				-			
	happy	2.67	-28	-12	-14	-			
	disgust	3.1	-28	-14	-16	2.69	20	-10	-12
	sad	3.35	-28	-8	-22	3.33	18	-8	-12
	neutral	3.24	-22	-8	-20	3.4	30	2	-24

Table 1.1. Z values above the 2.6 threshold and MNI coordinates of the activation peaks for the linear contrasts of interest in left and right amygdala.

	condition	threshold Z > 2.6			threshold Z > 3.1		
		Left	Right	R + L	Left	Right	R + L
building contrasts	fear	179	230	409	90	140	230
	anger	153	242	395	91	153	244
	happy	116	210	326	70	133	203
	disgust	96	187	283	64	103	167
	sad	33	37	70	8	14	22
	neutral	26	57	83	6	30	36
	25% happy	13	39	52	2	2	4
25 happy contrasts	fear	131	120	251	50	42	92
	anger	39	92	131	3	8	11
	happy	39	73	112	2	14	16
	disgust	35	37	72	2	2	4
	sad	0	0	0	0	0	0
	neutral	0	0	0	0	0	0
fear contrasts	anger	0	0	0	0	0	0
	happy	5	0	5	0	0	0
	disgust	22	1	23	0	0	0
	sad	94	67	161	16	10	26
	neutral	95	102	197	11	16	27

Table 1.2. Number of active voxels in left and right amygdala at significance levels of $p < .005$ ($Z > 2.6$, left columns) and $p < .001$ ($Z > 3.1$, right columns). Data refer to the 5-100% amygdala masks from the Harvard-Oxford sub-cortical probability atlas, which comprised a total of 830 voxels in the left amygdala and 950 voxels in the right amygdala.

Relative to the 25% happy face comparison, significant amygdala activation was observed for fearful, angry, happy and disgusted expressions when a $p < .005$ threshold ($Z > 2.6$) was applied. However, only fearful faces produced a consistent bilateral activation in the amygdala at the higher $p < .001$ threshold ($Z > 3.1$). Sad and neutral expressions did not show any significant activation in the amygdala when contrasted with the 25% happy expression (see Table 1.2).

In the contrasts of fear versus other expressions, significant activations appeared for sad and neutral faces, even though the number of active voxels was reduced at the $p < .001$ threshold ($Z > 3.1$). Twenty-two voxels in the left amygdala surpassed the lower $p < .005$ threshold ($Z > 2.6$) for the disgust versus fear condition. No significant amygdala differences were found for fear versus anger and fear versus happy contrasts.

The same analyses for the contrasts of interest, using statistical thresholds of $p < .005$ and $p < .001$, were performed with the 50-100% masks from the Harvard-Oxford sub-cortical probability atlas. The purpose was to use these masks to restrict the analysis to the amygdala region, allowing to evaluate whether activations from peri-amygdalar regions included within the 5-100% masks were distorting the overall pattern. Results for the 50-100% masks are reported in Table 1.3. Taking into account the reduced number of voxels included in these smaller masks (227 voxels in the left amygdala and 278 in the right amygdala, instead of 830 voxels in the left amygdala and 950 voxels in the right amygdala included in the 5-100% masks), the number of active voxels, in each contrast at the two significance levels applied, is consistent with the results reported above using the 5-100% masks. This confirmed that the observed activations actually occurred within the amygdala.

	condition	threshold Z > 2.6			threshold Z > 3.1		
		Left	Right	R + L	Left	Right	R + L
building contrasts	fear	77	120	197	44	84	128
	anger	85	129	214	49	93	142
	happy	56	128	184	34	85	119
	disgust	39	110	149	26	63	89
	sad	18	26	44	2	11	13
	neutral	9	46	55	2	27	29
	25% happy	3	33	36	0	2	2
25 happy contrasts	fear	50	64	114	17	26	43
	anger	30	40	70	1	1	2
	happy	25	38	63	1	2	3
	disgust	11	28	39	0	2	2
	sad	0	0	0	0	0	0
	neutral	0	0	0	0	0	0
fear contrasts	anger	0	0	0	0	0	0
	happy	0	0	0	0	0	0
	disgust	0	1	1	0	0	0
	sad	28	42	70	4	10	14
	neutral	44	52	96	7	10	17

Table 1.3. Number of active voxels in left and right amygdala at significance levels of $p < .005$ ($Z > 2.6$, left columns) and $p < .001$ ($Z > 3.1$, right columns). Data refer to the 50-100% amygdala masks from the Harvard-Oxford sub-cortical probability atlas, which comprised a total of 227 voxels in the left amygdala and 278 voxels in the right amygdala.

Next, the overall response of the amygdala to each condition was examined. Figure 1.2 shows the percent signal change in the left and right amygdala for each condition. The left amygdala showed positive activation for fearful, angry and happy expressions, whereas the right amygdala showed significant activation for fearful, angry, happy and disgusted expressions. Neither sad, neutral and 25% happy faces nor the building conditions gave any positive signal change in the amygdala. Finally, the response to fear was compared with each of the other emotions. There was no significant difference in the response to fear compared to happy, anger or disgust in either the left or right amygdala. However, there was a significantly greater response to fear compared to sadness and neutral in both the right and left amygdala.

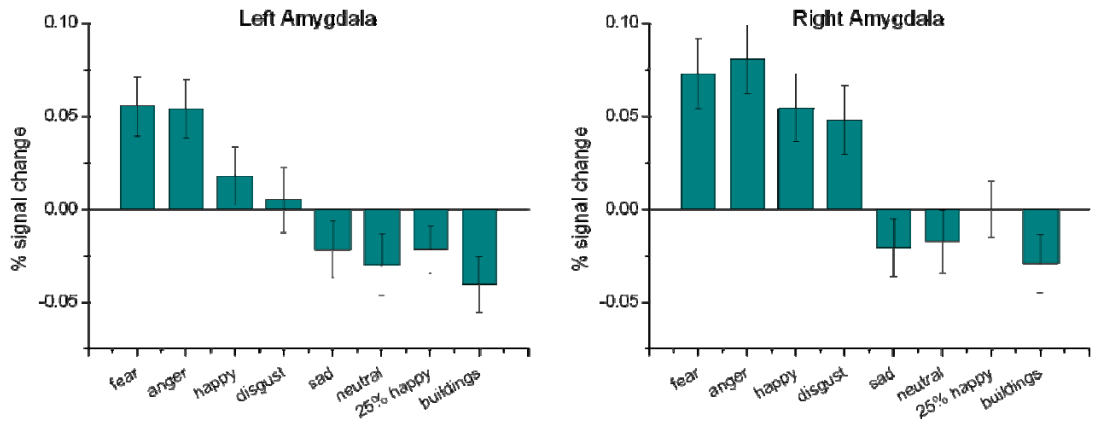

Figure 1.2. Percent signal change in left and right amygdala for each condition. The amygdala was defined by the 5-100% masks from the Harvard-Oxford sub-cortical probability atlas. Bars represent standard errors of the means.

Activations in brain regions other than the amygdala for each face expression versus baseline and versus 25% happy faces are reported in Table 1.4 and 1.5. In contrasts with buildings, extensive significant activations appeared in the lateral occipital cortex and precuneus for all face expressions (see Table 1.4). The fusiform cortex showed significant responses to all expressions apart from happy and activations in the frontal lobes were observed for all emotions but neutral faces. On the other hand, whole brain analysis for the within-category

contrasts (each face expression > 25% happy faces) showed no significant activation outside the amygdala for fearful expressions, as well as for sadness and neutral faces. Significant activation in posterior regions appeared for angry, disgusted and particularly for happy expressions (see Table 1.5).

expression	region	hemisphere	MNI (x y z)			n°voxels	Z score
fear	inferior lateral occipital cortex	R	48	-76	-8	1189	6.17
	precuneus	R	2	-60	36	533	4.77
	fusiform cortex	R	48	-52	-26	401	5.95
	inferior lateral occipital cortex	L	-54	-68	8	277	4.70
	occipital pole	L	-38	-92	-18	124	4.26
	frontal pole	L	-4	60	20	94	3.82
	inferior temporal gyrus	L	-46	-44	-28	66	4.04
	superior frontal gyrus	L	-6	50	35	49	3.96
	superior lateral occipital cortex	L	-46	-64	22	33	3.47
	fusiform cortex	L	-44	-58	-22	21	3.62
angry	fusiform cortex	R	46	-52	-26	1599	6.39
	lateral occipital cortex	L	-54	-76	2	383	4.74
	precuneus	R	6	-58	28	143	3.72
	inferior lateral occipital cortex	L	-38	-84	-14	126	4.38
	fusiform cortex	L	-44	-44	-22	119	4.02
	cingulate gyrus	L	-22	-48	14	75	3.66
	inferior frontal gyrus	R	56	26	-2	54	3.88
	supplementary motor cortex	L	-6	-4	56	43	3.57
	superior frontal gyrus	L	-2	14	58	35	3.88
	inferior lateral occipital cortex	R	38	-60	6	34	3.92
	frontal pole	L	-10	60	18	22	3.42
happy	inferior lateral occipital cortex	R	50	-76	-8	1193	6.27
	precuneus	R	4	-62	30	455	4.54
	inferior lateral occipital cortex	L	-40	-88	-18	208	4.50
	lateral occipital cortex	L	-54	-78	0	123	4.43
	medial frontal cortex	R	2	54	-18	91	4.34
	middle temporal gyrus	R	50	-44	2	84	4.30
	cingulate gyrus	L	-16	-46	16	26	3.5
disgust	inferior lateral occipital cortex	R	50	-76	-8	1440	6.54
	inferior lateral occipital cortex	L	-54	-70	10	490	4.59
	precuneus	R	2	-62	32	141	4.30
	inferior frontal gyrus	R	56	28	4	64	4.18
	fusiform cortex	L	-44	-44	-22	34	3.44
	medial frontal cortex	R	2	54	-20	25	3.78
sad	inferior lateral occipital cortex	R	48	-78	-8	638	5.74
	fusiform cortex	R	46	-50	-26	295	5.59
	precuneus	L	0	-60	38	230	4.42
	lateral occipital cortex	L	-54	-76	4	170	4.79
	inferior lateral occipital cortex	L	-44	-86	-14	129	4.20
	middle temporal gyrus	R	52	-44	2	94	4.37
	frontal pole	L	-4	58	20	23	3.62

neutral	inferior lateral occipital cortex	R	50	-76	-8	456	5.43
	fusiform cortex	R	46	-52	-26	256	5.74
	precuneus	R	6	-58	24	209	4.45
	lateral occipital cortex	L	-54	-70	8	125	4.25
	cingulate gyrus	L	-22	-44	12	84	4.04
	inferior lateral occipital cortex	L	-44	-86	-16	81	4.41
	middle temporal gyrus	R	38	-56	14	30	3.63
25% happy	inferior lateral occipital cortex	R	50	-76	-8	873	5.53
	Precuneus	R	2	-60	36	721	5.06
	frontal pole	R	6	62	14	397	4.19
	fusiform cortex	R	46	-52	-26	262	5.72
	lateral occipital cortex	L	-54	-68	6	180	4.20
	inferior lateral occipital cortex	L	-44	-88	-14	84	4.08
	middle temporal gyrus	R	52	-42	6	29	3.76

Table 1.4. Responses to facial expressions (each expressions > buildings) in all brain regions excluding the amygdala. Activations with more than 20 contiguous voxels which surpassed the $p < .001$ threshold ($Z > 3.1$, uncorrected) are reported. Cluster size is described by the total number of contiguous active voxels. Regions are labelled according to the Harvard-Oxford cortical atlas and MNI coordinates, laterality (R, right; L, left) and Z values of the peak voxels are reported.

expression	region	hemisphere	MNI (x y z)			n°voxels	Z score
angry	inferior fusiform cortex	L	-44	-48	-22	25	3.64
happy	occipital pole	R	34	-88	6	257	4.27
	occipital pole	L	-24	-96	4	108	3.67
	lateral occipital cortex	R	42	-74	-12	21	3.51
disgust	fusiform cortex	R	46	-48	-24	54	3.58
	lateral occipital cortex	R	34	-92	0	35	3.55

Table 1.5. Responses to emotion expressions (each expressions > 25% happy faces) in all brain regions excluding the amygdala. Activations with more than 20 contiguous voxels which surpassed the $p < .001$ threshold ($Z > 3.1$ uncorrected) are reported. Cluster size is described by the total number of contiguous active voxels. Regions are labelled according to the Harvard-Oxford cortical atlas and MNI coordinates, laterality (R, right; L, left) and Z values of the peak voxels are reported.

1.4. Discussion

The aim of this study was to determine whether activity in the amygdala is selective for emotional expressions in general or is only selective for particular expressions such as fear that signals the presence of potential threat. To address a potential source of conflicting conclusions from the previous literature (Sergerie et al., 2008), the responses to face images posing different emotional expressions were contrasted with face and non-face conditions.

Results showed that: (i) in contrasts of face conditions versus buildings, the amygdala responded to some extent to all expressions, however higher activations appeared for fear,

anger, happiness and disgust emotions; (ii) in within-category contrasts (each expression > 25% happy faces), fear produced higher peak values and more extended amygdala activation than other emotions, but significant activations appeared also for angry, happy and disgusted expressions; (iii) direct contrasts of fear versus other emotions and percent signal changes for each condition confirmed that activation was higher for the fearful expression as compared to sad and neutral, but not significantly higher for fear compared to angry, happy and disgusted faces. Consistent findings were observed using both masks narrowed to the amygdala nucleus and masks extended to the peri-amygdalar regions, confirming that the pattern of results was due to activations occurring within the amygdala.

Previous neuroimaging studies provided mixed evidence, being taken to support either a specific role for the amygdala in processing fearful faces (Calder et al., 2001) or a more general amygdala activation for several expressions (Fitzgerald et al., 2006; Winston et al., 2003). Present results show clearly why each position has some merits. The amygdala responded to all face expressions to some extent, as evidenced by contrasting each expression with buildings. However, this face versus non-face contrast does not rule out the possibility that the activation is to faces per se, rather than more specifically to facial expressions. The more easily interpreted contrasts are therefore those between facial expression conditions and the 25% happy face comparison conditions, since any differences found for these will reflect the processing of expression. In these results, contrasts between face expressions and the face control condition highlighted stronger activation for fearful expressions. Sergerie et al.'s (2008) meta-analysis has already evidenced the importance of the control condition, reporting stronger amygdala activations when a low-level baseline, such as scrambled-images or a grey screen, is used as compared with control conditions with neutral faces or other pictures. The present results support this conclusion and directly demonstrate that the use of a control condition with stimuli belonging to the face category (a

mildly happy expression) or a non-face category (e.g. buildings) can affect the results and therefore point towards different conclusions.

Although the amygdala was considered as a region of interest, results from the whole brain analysis were also informative. As expected, significant activations appeared in the face-selective regions in the occipital and temporal lobes (Haxby et al., 2000) when face conditions were contrasted with buildings. However, only a few activations were observed for contrasts of facial expressions versus 25% happy faces; in particular, any significant activation outside the amygdala was found for fearful expressions, whereas significant responses in the posterior regions appeared for angry, happy and disgusted faces. This confirms the amygdala special role in processing fear expression in faces.

Several studies in the affective neuroscience literature have identified the amygdala as a neural correlate for processing threatening stimuli (Dalgleish, 2004; Vytal and Hamann, 2010) and its response to fearful faces is consistent with this. On the other hand, Sander et al. (2003) pointed out that the amygdala involvement in processing fear-related stimuli does not necessary imply that its role is restricted to fear; instead they proposed a role for the amygdala in detecting relevant stimuli regardless of their valence.

To test whether the amygdala response is fear-specific, activation for fear was contrasted with each of the other facial expressions. This stringent comparison showed that responses to fear were higher than for some other expressions (e.g. sadness), which is inconsistent with the hypothesis that the amygdala is involved in processing all expressions. However, a more complex pattern was evident than a pure response to fear per se (see Figure 1.2 and Table 1.3). There are two main possible reasons why this might be the case, each worth taking seriously for further investigation. One (Aggleton and Young, 1999) is that only some of the multiple nuclei in the amygdala are involved in a specific fear evaluation mechanism and others serve more general social purposes. This is difficult to rule out with the current spatial

resolution of fMRI. The alternative is that the amygdala has a more general role in emotional appraisal for which fear is one of the most effective elicitors (Sander et al., 2003).

In summary, the present study helps to clarify conflicting results in the literature about amygdala responses to facial expressions. The findings show that the amygdala is highly responsive to fearful faces, but the activation is not specific to this emotion since increased fMRI signal was also observed to some extent for angry, happy and disgusted expressions. Because of this complex pattern, the analyses show that using different control conditions and applying different thresholds in the statistical analysis can lead to a pattern that emphasises a more general activation across emotions or alternatively a more selective response to fearful expressions. Such issues could have influenced previously reported patterns of findings and therefore should be taken into account in further studies intended to elucidate the profile of amygdala responses to faces and emotions.

2. Study 2: Response of face-selective brain regions to social traits in faces

2.1. Introduction

Following results of Study 1, which evidenced amygdala activation for fear expressions and other basic emotions, another interesting issue was to investigate amygdala responses to faces varying in level of trustworthiness. Indeed, the efficient perception of this facial dimension is critical for social interaction since, as basic emotion expression, it is a relevant cue to judge other people as approachable or source of potential threat (Adolphs et al., 1998). The neuroimaging literature concerning amygdala response to perceived trustworthiness has suggested different hypotheses for the role of the amygdala in face evaluation, supporting either the idea of the arousing signal for approach/avoidance behaviour in case of linear activation (Engell et al., 2007), or the function of salient stimuli detector in case of U-shaped activation (Said et al., 2008). The quadratic pattern is also consistent with the idea that faces are represented in a multidimensional space in which the origin represents the average face and more distinctive faces are represented away from the origin (Said et al., 2010; Valentine, 1991). From this perspective, the linear and nonlinear responses to trustworthiness in previous studies could be due to uncontrolled variation in the distinctiveness of faces (Said et al., 2011).

A key aim of the present study was therefore to address these different perspectives on the way that the amygdala represents information about faces by comparing the neural responses to trustworthiness and a control face dimension (male-female). To do this a novel set of stimuli was developed with naturalistic faces varying in perceived trustworthiness and along an orthogonal male-female dimension. Previous studies have used face photographs, which cannot vary relevant stimulus dimensions systematically, or computer-synthesised faces that, whilst useful, form highly constrained sets that may not utilise all of the cues that are naturally available to human observers. Differently, the new stimuli presented here were

derived from prototype images created with a photograph averaging technique, in order to maximise the presence of naturally occurring cues that underpin trustworthiness and gender judgments. These prototypes were then systematically manipulated through image-morphing to create independent dimensions of trustworthiness and gender.

Neural responses to these novel sets of stimuli were tested using a block design fMRI paradigm, to take advantage of its greater statistical power compared to event-related designs (Sergerie et al., 2008). If the social meaning of facial trustworthiness cues were crucial to determining the neural responses, we would expect the patterns of activation to vary with the trustworthiness of the faces, but not with changes in gender. On the other hand, if the distinctiveness of the face is important, then a similar pattern of activation should be evident for variation in both the social and control dimensions (Said et al., 2010; 2011).

A second aim of the study was to determine whether the pattern of response was specific to the amygdala or was evident in other face-responsive regions of the brain. Several previous studies have drawn conclusions based only on responses from the amygdala region itself, but it is crucial to correctly interpret these amygdala responses to know whether they are similar or different in form from the responses of other regions involved in face perception. Therefore analyses included responses from core face-selective regions of the occipital and temporal lobes (Haxby et al., 2000) as well as the amygdala itself.

2.2. Method

Participants

Twenty healthy volunteers (10 male, 10 female, mean age = 22.9 years, range 18-35) took part in the experiment. All participants were right-handed, with a western cultural background, and had normal or corrected to normal vision with no history of neurological illness. The study was approved and conducted following the guidelines of the Ethics

Committee of the York Neuroimaging Centre, University of York. All participants gave written consent prior to their participation.

Experiment stimuli

Figure 2.1 shows the complete matrix of images from which the stimuli used in the experiment were selected.

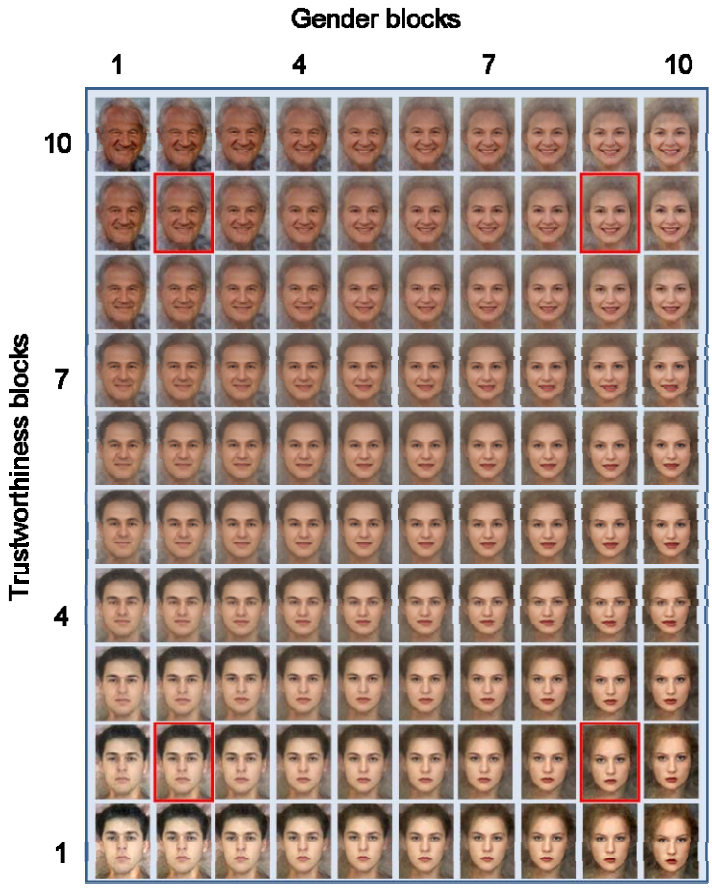

Figure 2.1 Matrix of faces created by computer image manipulation. Images in red squares represent the prototypes used to produce the matrix of 10 levels of face gender (rows) and ten levels of face trustworthiness (columns). Four trustworthiness conditions and four gender conditions were selected for the fMRI experiment, in order to cover the full range of each of the dimensions. Stimuli for the trustworthiness blocks were the rows labelled with numbers 1, 4, 7 and 10 of the matrix, thus including ten different face images with the same trustworthiness level but varying in terms of gender. Gender blocks consisted of columns 1, 4, 7 and 10, each with ten faces varying in trustworthiness but constant in terms of gender.

The matrix was created as follows. Photographs of 500 adult male and 500 adult female faces were collected from Internet. The photographs varied in pose, age and expression, to allow a range of cues as wide as possible in the images. However, photographs of famous

people were excluded, to eliminate potential influences of prior knowledge about the person. Moreover, only Caucasian adult faces were chosen, to reduce potential cultural influences. The 1,000 face photographs were rated for trustworthiness (using 1-7 scales) by six independent raters. From these ratings the 15 highest and 15 least trustworthy male faces and the 15 highest and 15 least trustworthy female faces were selected; constraints were that the photographs included no spectacles, were as close to frontal view as possible, showed no beards or moustaches, and no more than two faces with hats in each set were accepted. There was no matching on any other characteristic, with free variation of all other aspects. The faces in each set of 15 photographs were then averaged using PsychoMorph software (Tiddeman, et al., 2001) to create four prototypes (high and low trustworthy male, high and low trustworthy female). Image continua were then created for trustworthiness of male faces (from very high to very low trustworthiness) and for trustworthiness of female faces by caricaturing each prototype at two levels to increase its distance from the opposite prototype and by anti-caricaturing each prototype at two levels to decrease distance from the opposite prototype. For example, the highly trustworthy male prototype was caricatured to enhance its trustworthiness by increasing differences from the low trustworthy male prototype and it was anti-caricatured to diminish its trustworthiness by decreasing differences from the low trustworthy male prototype. In this way, a quasi-linear continuum of 10 male face-like images of varying trustworthiness was created, and a corresponding continuum of 10 female face-like images of varying trustworthiness.

These continua of 10 images were then presented in random order and rated for trustworthiness (on a 1-7 low-high trustworthy scale) by 10 raters (5 male, 5 female, mean age = 20.4 years, s.d. = 0.55) who did not otherwise participate in the study. The correlation between rated trustworthiness and position on the appropriate continuum was 0.94 for the male images and 0.95 for the female images, showing that the caricaturing and anti-

caricaturing manipulations were successful in creating continua varying systematically in perceived trustworthiness. However, it was also necessary to match continua needed for the present experiment so that the male and female prototype images were of equivalent high or equivalent low trustworthiness. We therefore selected a male and a female image that were rated equally low in trustworthiness, and a male and a female image that were rated equally high. These matched pairs of male and female images formed the four new prototypes used to generate Figure 2.1. They are shown at highlighted positions in Figure 2.1 corresponding to the intersections of the second and ninth rows with the second and ninth columns. The rest of the 10 x 10 matrix was generated by morphing the faces between the prototypes along the trustworthiness and the gender dimensions and adding a caricatured image in each of the four directions. If we consider the prototypes to represent 0% and 100% on each dimension, the manipulation used generated images with the following percentages along the gender (horizontal) and trustworthiness (vertical) axes of Figure 2.1: -15% 0% 15% 30% 45% 55% 70% 85% 100% 115%. On this scale, values falling outside the 0-100% range represent caricatures with respect to the opposite prototype.

Imaging parameters

Scanning was performed at the York Neuroimaging Centre at the University of York with a 3 Tesla HD MRI system with an eight channels phased array head coil (GE Signa Excite 3.0 T, High resolution brain array, MRI Devices Corp., Gainesville, FL). Axial images were acquired for functional and structural MRI scans. For fMRI scanning, echo-planar images were acquired using a T2*-weighted gradient echo sequence with blood oxygen level-dependent (BOLD) contrast (TR = 3 sec, TE = 32.7 msec, flip-angle = 90°, acquisition matrix 128 x 128, field of view = 288 mm x 288 mm). Whole head volumes were acquired with 38 contiguous axial slices, each with an in-plane resolution of 2.25 mm x 2.25 mm and

a slice thickness of 3 mm. The slices were positioned for each participant to ensure optimal imaging of the temporal lobe regions, where the amygdala is situated. T1-weighted images were acquired for each participant to provide high-resolution structural images using an Inversion Recovery (IR = 450 msec) prepared 3D-FSPGR (Fast Spoiled Gradient Echo) pulse sequence (TR = 7.8 sec, TE = 3 msec, flip-angle = 20°, acquisition matrix = 256 x 256, field of view = 290 mm x 290 mm, in-plane resolution = 1.1 mm x 1.1 mm, slice thickness = 1 mm). To improve co-registration between fMRI and the 3D-FSPGR structural a high resolution T1 FLAIR was acquired using the same physical dimensions as the fMRI protocol (TR = 2850 msec, TE = 10 msec, acquisition matrix 256 x 224 interpolated to 512 giving effective in plain resolution of 0.56 mm).

Localiser scan

In order to identify brain regions responding selectively to faces, participants performed a separate localiser scan (see Andrews et al., 2010). Twenty blocks with 10 images were run, using Neurobehavioural System Presentation 13.0 software. Each block contained images from one of five different categories: faces, bodies, objects, places or Fourier-scrambled images derived from the previous categories. Face images were taken from the Psychological Image Collection at Stirling (PICS; http://pics.psych.stir.ac.uk/) and bodies were selected from a body images collection at Bangor (http://pages.bangor.ac.uk/~pss811/page7/page7.html). Images of other categories were taken from website sources. Each image was presented for 700 msec followed by a 200 msec fixation cross, giving a block duration of 9 seconds for the 10 images. Stimulus blocks were interleaved with resting periods of 9 seconds with a fixation cross superimposed on a grey screen. The five conditions were repeated four times in a counterbalanced order.

Trustworthy/Gender scan

The experiment aimed to test whether the response patterns in the amygdala and face-selective regions are specific to the trustworthiness dimension or if similar patterns appear for faces varying along an independent and orthogonal male-female dimension. A block design was used with eight conditions divided into four trustworthiness conditions and four gender conditions. Each of the four trustworthiness blocks comprised the images from a row of the stimulus matrix shown in Figure 2.1 (rows labelled as 1, 4, 7 and 10 were selected) and therefore involved faces varying in terms of gender but with the same trustworthiness level. Each of the four gender blocks consisted of a column from the stimulus matrix (columns 1, 4, 7 and 10 were selected) and therefore involved faces varying in level of trustworthiness but not in terms of gender. Consequently, the eight conditions presented sampled the full range of each of the two orthogonal dimensions. The blocks for each condition were repeated five times in a counterbalanced order. Within each block the 10 images were presented in a pseudorandom order for 1 second each followed by a 200 msec fixation cross, giving a total block duration of 12 seconds; blocks were interleaved with a 12 seconds fixation cross on a grey screen. To monitor attention during the scan session a red spot detection task was used. In one or two images per block a small red spot appeared; subjects were instructed to look at the stimuli and press with the right index finger a response button whenever they saw the red spot. Subjects responded correctly to the majority of the red spot trials (mean accuracy = 98.6%, s.d. = 2.87).

After the fMRI scan a behavioural task was run to check how each participant perceived the stimuli. Participants were asked to rate on a seven-point scale the trustworthiness (1 = very untrustworthy, 7 = very trustworthy) and the masculinity-femininity (1 = high masculine, 7 = high feminine) of the images used in the experiment. These two sets of ratings were completed separately in a counterbalanced order.

fMRI data analysis

Image analyses were performed by means of FEAT (FMRI Expert Analysis Tool), part of FSL (http://www.fmrib.ox.ac.uk/fsl). For each participant the following pre-statistic processing was applied: motion correction using MCFLIRT (Jenkinson et al., 2002), slice-timing correction using Fourier-space time-series phase-shifting, non-brain removal using BET (Smith, 2002), spatial smoothing using a Gaussian kernel (FWHM 5mm in the localiser scan and 6mm in the main experiment), grand-mean intensity normalisation of the entire 4D dataset by a single multiplicative factor; high-pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 60 sec in the localiser scan and sigma= 120 sec in the main experiment).

Face-selective regions comprising the core components identified by Haxby et al. (2000) were individually defined in each participant's brain using the localiser scan by averaging the four contrasts faces > bodies, faces > objects, faces > places and faces > scrambled images. The average of these four contrasts in each participant was thresholded at $Z > 2.6$ ($p < .005$, uncorrected). In this way, the fusiform face area (FFA), occipital face area (OFA) and right posterior superior temporal sulcus (pSTS) could be identified at the level of each single participant. These regions of interest (ROIs) were defined from the thresholded statistical images (see Andrews et al., 2010). The FFA, OFA and pSTS each appeared as a contiguous cluster of voxels in each participant located respectively in the inferior fusiform gyrus, in the posterior occipital cortex and in the superior temporal lobe. A different approach had to be taken to define the amygdala, which is not reliably identified through a functional localiser at the individual level. A face-responsive ROI in the amygdala was therefore defined by considering the statistical map of amygdala activation at the group level, resulting from the four contrasts averaged and thresholded at $Z = 3$ ($p \leq .001$, uncorrected), which was back-transformed into the individual MRI space for each participant.

Within these functionally identified face-selective regions (amygdala, OFA, FFA, pSTS) derived from the functional localiser scan, data from the main experiment were analysed by extracting the time-course of the filtered MR data as percent signal change in each voxel and then averaging the voxels within each ROI for each participant. The average time-course for the different conditions was calculated and data were normalised relative to the zero time point for that stimulus block. The peak of activation, considered as the average of the response between 9 and 15 seconds after block onset, was used for the analyses.

For each ROI the following analyses were performed to test the linear and quadratic responses. First, a linear regression and a second-order polynomial were fitted to the responses at group level in order to investigate the activation pattern in each region. Second, a linear regression and a second-order polynomial were fitted to each individual participant's responses and paired t-tests were used to test differences between the R-squared of the two fitted equations in each ROI. Finally, paired sample t-tests were performed to compare the linear and quadratic regressions for the gender and trustworthiness dimensions.

2.3. Results

Behavioural data

The post-scan behavioural ratings were analysed to check that the participants in the fMRI experiment rated the stimuli in line with what was intended. The trustworthiness and gender ratings of each participant were correlated with the four trustworthiness and the four gender levels included in the fMRI scan. One participant was excluded from the following analyses because of a very low correlation score for the trustworthiness rating ($r = 0.01$), whereas for the remaining participants the correlations were always > 0.8 for both dimensions (mean $r = 0.96$, for trustworthiness rating; mean $r = 0.98$, for gender rating). In this post-scan behavioural task, participants rated the stimuli on a 7-point scale separately for the

trustworthiness and gender dimensions. Responses at stimuli included in the fMRI experiment were analysed, hence levels 1, 4, 7 and 10 of each dimension (see Figure 2.1). The mean rating for trustworthiness level one was 2.36 (s.d. = 1.27), 3.94 (s.d. = 0.77) for level four, 4.87 (s.d. = 0.84) for level seven, and 6.14 (s.d. = 1.14) for level ten. The mean rating for gender level one was 1.53 (s.d. = 0.38), 2.85 (s.d. = 0.67) for level four, 5.26 (s.d. = 0.72) for level seven, and 6.37 (s.d. = 0.91) for level ten. The mean rating for both dimensions significantly correlated with the trustworthiness ($r = .995$, $p = .005$) and gender levels ($r = .99$, $p = .01$).

Localiser scan

Figure 2.2 shows the location of regions within the amygdala, the occipital and temporal lobes (FFA, OFA, pSTS) that showed face-selective activity from a whole-brain group analysis of the localiser scan data. Mean MNI coordinates and size of each region across participants are reported in Table 2.1. The FFA and OFA were identified in all of the 19 participants and right pSTS in 18 participants.

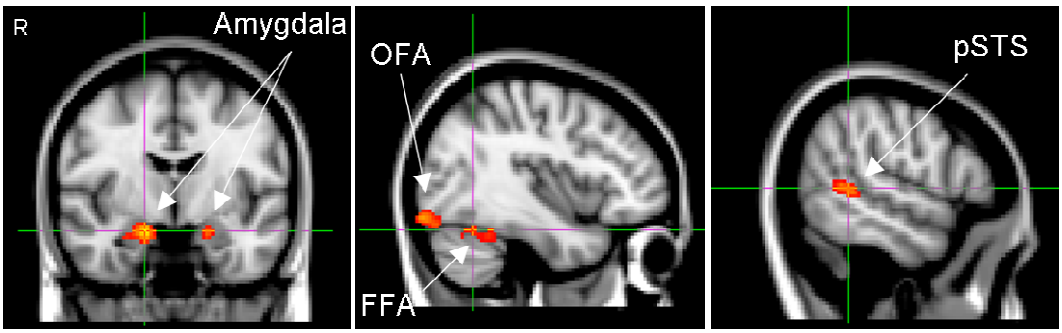

Figure 2.2. Location of the face-selective regions (amygdala, FFA, OFA, pSTS) in a whole-brain group analysis of the localiser scan. Statistical parametrical maps thresholded at $Z = 3$ ($p \leq .001$, uncorrected) resulting from the average of four contrasts (faces > bodies, faces > objects, faces > places and faces > scrambled images) are represented. Images follow the radiological convention, with the right hemisphere represented on the left side.

Region		N	MNI coordinates (x, y, z)			Size (cm^3)
Amygdala	R	19	18	-6	-18	4.44
	L	19	-18	-10	-18	1.24
FFA	R	19	42 (4)	-56 (8)	-23 (5)	2.23 (1.48)
	L	18	-42 (4)	-58 (7)	-23 (4)	1.35 (1.06)
OFA	R	19	39 (6)	-81 (9)	-14 (5)	2.19 (1.93)
	L	18	-37 (5)	-83 (5)	-18 (5)	1.42 (1.25)
pSTS	R	18	50 (8)	-53 (8)	5 (6)	0.82 (0.79)

Table 2.1. MNI coordinates and size of face-selective regions. The left and right amygdala were defined at the group level. FFA, OFA and pSTS were defined in each participant; values represent the mean (s.d.) across all 19 participants.

Trustworthy/Gender scan

Figure 2.3 shows the peak response in each ROI for faces varying along the trustworthiness and gender dimensions. Since both hemispheres showed similar response patterns in FFA, OFA and amygdala, the responses in the right and left hemispheres were combined for these regions. In contrast, the pSTS region could only be reliably identified in the right hemisphere. For the trustworthiness dimension, results at group level showed bigger R-squared values for the quadratic polynomial than for the linear regression in all the face-selective regions. The same pattern of greater overall quadratic than linear responses for all regions was also seen for the gender dimension (Table 2.2).

Quadratic and linear regressions were then fitted to the individual responses in each ROI and paired sample t-tests confirmed that the R-squared values for the quadratic polynomial were significantly higher than the R-squared for the linear regression for both the dimensions in all the regions (amygdala: trustworthiness [$t(18) = 4.97$, $p < .001$], gender [$t(18) = 6.33$, $p < .001$]; FFA: trustworthiness [$t(18) = 5.07$, $p < .001$], gender [$t(18) = 5.1$, $p < .001$]; OFA: trustworthiness [$t(18) = 6.12$, $p < .001$], gender [$t(18) = 4.12$, $p = .001$]; right pSTS: trustworthiness [$t(17) = 4.15$, $p = .001$], gender [$t(17) = 5.27$, $p < .001$]).

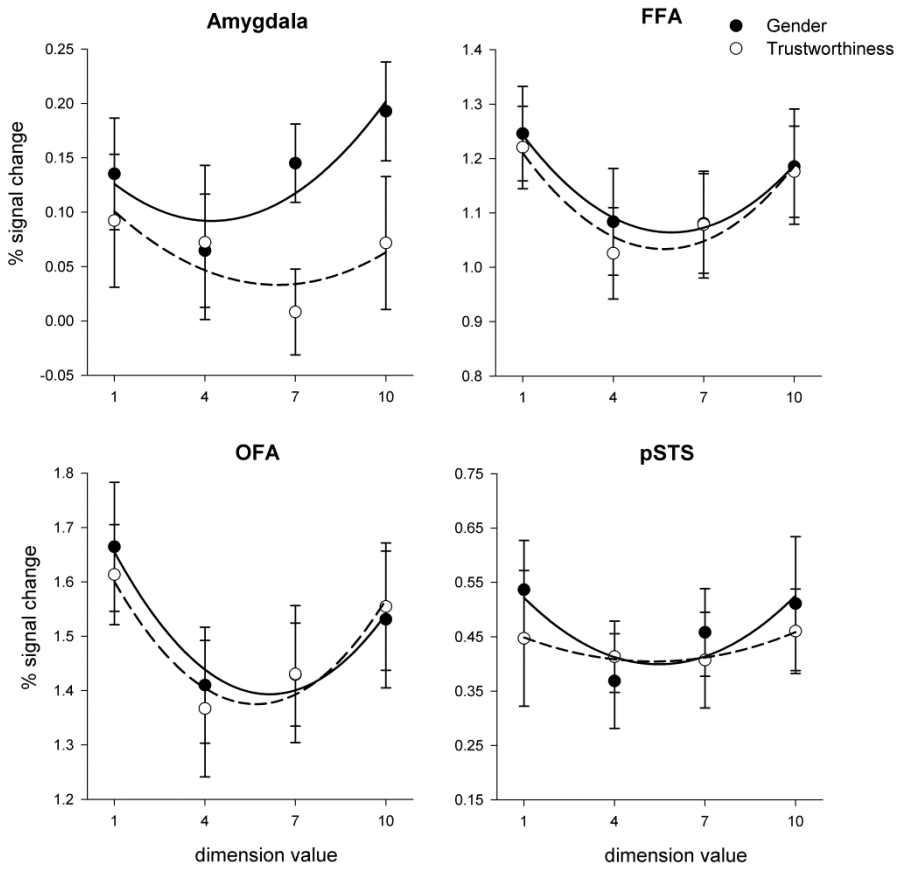

Figure 2.3. Response to trustworthiness and gender dimensions in the four ROIs defined by the localiser scan. U-shaped lines represent the quadratic polynomial that best fitted the data in the trustworthiness (dotted line) and gender dimensions (full line); bars represent standard errors of the mean.

Having established the general pattern of quadratic rather than linear response in all ROIs, it is of interest to verify whether the quadratic component was more pronounced for one dimension than the other. However, there were no significant differences comparing quadratic R-squared between the two dimensions of gender and trustworthiness in any region (paired sample t-tests, $p > .05$).

	Trustworthiness		gender	
	quadratic	linear	quadratic	linear
Amygdala	0.63	0.20	0.80	0.38
FFA	0.91	0.01	0.99	0.09
OFA	0.92	0.02	0.95	0.18
pSTS	0.97	0.03	0.74	0

Table 2.2. R-squared values for the quadratic polynomial and linear regressions for the two dimensions.

2.4. Discussion

This study investigated the response pattern in the amygdala and the core face-selective brain regions to faces varying in a social (trustworthiness) and a control (male-female) gender dimension. To do this a novel set of stimuli was created that consisted of naturalistic face images, systematically varied in perceived trustworthiness and gender. Results of behavioural ratings showed that the stimuli comprised the multiplicity of cues that are used to evaluate variations in trustworthiness. There were high correlations between rated trustworthiness and vertical position of the images shown in Figure 2.1. Since the essence of the method used to derive the prototype images was simply averaging face photographs rated as high or low in trustworthiness, the continued presence of high and low trustworthiness in the averaged prototype images shows that the cues that convey these impressions must have been reasonably consistently present in the original photographs. Inspection of Figure 2.1 suggests that the trustworthiness dimension involves cues that include a combination of age, skin colour and hostile expression. Oosterhof and Todorov (2008) have already shown by means of computer models that trustworthiness evaluation is sensitive to emotional expression. In the same study maturity cues did not correlate with trustworthiness evaluation when internal features, linked with the trustworthiness features, were masked. However, Oosterhof and Todorov's (2008) stimuli were synthesised computer images with a limited range of ages and smoothing of texture cues such as wrinkles that can signify a loss of elasticity in the skin. In contrast, the present stimuli were created with a data-driven approach without any a priori constraint, thus taking into account the multiplicity of naturally occurring cues, which influence trustworthiness judgments, and age seems to be part of this evaluation.

To determine how the brain responded to the stimulus set, a localiser scan was used to functionally define the amygdala and the core face-selective regions (OFA, FFA, pSTS) in

the occipital and temporal lobes (Haxby et al., 2000). The main findings from this analysis were: (i) the amygdala responded to varied trustworthiness with a U-shaped quadratic function; (ii) the amygdala also showed a U-shaped pattern of response to changes in gender; (iii) FFA, OFA and right pSTS showed a similar U-shaped pattern for both the trustworthiness and gender dimensions.

Previous neuroimaging studies have reported different patterns of response in the amygdala to variations in trustworthiness. Some studies have shown a greater response in the amygdala for untrustworthy as compared to trustworthy faces (Winston et al., 2002) with a linear trend in amygdala activation for increasing untrustworthiness (Engell et al., 2007), whereas other reports have found U-shaped quadratic responses in the amygdala (Todorov et al., 2008; Said et al., 2008). The present results provide support for a U-shaped quadratic response to trustworthiness in the amygdala. Critically, this finding was replicated for a functionally defined face-selective region within the amygdala using a novel set of naturalistic face images which varied systematically in perceived trustworthiness and using the higher statistical power gained from a fMRI block design.

The specificity of amygdala responses to social cues in faces has remained an issue of debate since previous studies have reported U-shaped amygdala responses to face dimensions different from trustworthiness. For example, Winston et al. (2007) found greater activation in the right amygdala when highly attractive or unattractive faces were presented compared to moderately attractive faces; although a correlation between attractiveness and face valence could potentially have influenced this result (Todorov and Engell, 2008). Another fMRI study showed a quadratic response for computer-generated faces that varied along a dimension orthogonal to trustworthiness with lower social relevance (Said et al., 2010). Results from the present study confirm these previous findings since non-linear activations were found in the amygdala for realistic faces morphed along a continuum of perceived

trustworthiness and a comparison gender dimension. Notably, the entirely data-driven approach used to create the present stimuli offers independent confirmation that findings with more artificial stimulus sets can be considered reliable, and of course enhances the ecological validity of results by allowing to systematically manipulate the dimensions of interest without constraints on the range of cues naturally available in face perception.

As well as the amygdala, it was of interest to clarify also the response pattern in face-selective regions in the occipital and temporal lobes. Therefore, a functional localiser scan was used to define bilateral FFA, bilateral OFA and right pSTS in each participant, and then extract the percent signal change during the main experiment within each ROI. These face-selective regions form Haxby et al.'s (2000) core system for face perception, and all of them showed U-shaped activations; again with no significant difference between the two dimensions. Previous studies mostly focused on the activation within the amygdala (Engell et al., 2007; Todorov et al., 2008) and hypothesised that the activity in the posterior face-selective regions was influenced by the amygdala (Todorov and Engell, 2008). Only Said et al. (2010) used a separate localiser scan to functionally define the face-selective regions on an individual subject level and, similarly to present results, they reported quadratic activations in FFA for their social and non-social dimensions. However, responses in OFA and pSTS were less clear in Said et al.'s (2010) study, showing a non-significant quadratic trend in OFA and a quadratic effect in pSTS for the social dimension but not for the non-social dimension. In contrast, the common quadratic pattern found in all the face-selective regions in the present experiment might be taken to suggest that these areas are equally important for the perceptual analysis of the stimuli. The different experimental designs used here and in Said et al.'s (2010) study might potentially account for the different effects found in OFA and pSTS. Beside this, though, the features of the control dimension could have a key role in understanding activations in these regions. The stimuli created for this

experiment varied along two dimensions that are both well recognisable as face categories, trustworthiness and gender, whereas Said et al. (2010) used computer modelling to generate a control dimension orthogonal to the social dimension but not definable as a specific face category. Therefore, it may prove to be the case that ability to identify face variations as ecologically relevant dimensions is important to eliciting quadratic responses in OFA and pSTS.

Overall, the present findings can be interpreted in line with the concept that faces are represented by a multidimensional space in which each face represents a particular location (Valentine, 1991). The origin of the face space reflects the average face and as the distance from the origin increases as faces become more distinctive. Loffler et al. (2005) provided neuroimaging support for this perspective, since they showed that the response of face-selective regions increases with the geometric distance from the average face. Therefore, the U-shaped function shown here and in other studies could be considered coherent with the idea that responses from the amygdala and other face-selective regions are at least in part driven by coding the difference between the presented faces and an average face, regardless of the specific social meaning of the stimuli (Said et al., 2010; 2011).

An alternative explanation for these results might be that trustworthiness and gender are both important dimensions, which in light of their social properties require specific coding mediated by U-shaped activations in the amygdala and face-selective regions. However, the hypothesis of a multidimensional representation for face stimuli at present seems more likely in light of previous findings of increased fMRI signal for increasing distinctiveness in face geometry (Loffler et al., 2005) and reports of quadratic activations for different face dimensions manipulated both with computer models and with photographs (Said et al., 2010; Winston et al., 2007). Although the stimuli were not explicitly controlled for distinctiveness, it is likely that the way they were generated would lead to images that lie closer to the centre

of Figure 2.1 being closer to an average face (more 'typical' in appearance) and those falling toward the periphery of Figure 2.1 being more distinctive. This was checked by asking to a separate group of ten participants to rate the images included in the matrix along the distinctiveness-typicality dimension. Results of this rating confirmed that perceived face distinctiveness increased moving from the centre to the edges of the matrix along both the dimensions. Indeed, rated distinctiveness was highly correlated with the U-shaped regressor for both the trustworthiness ($r = 0.88$, $p = .001$) and the gender dimensions ($r = 0.92$, $p < .001$).

Previous studies have interpreted non-linear responses to trustworthiness in the amygdala in terms of detecting and evaluating socially salient stimuli that are relevant for guiding approach and avoidance behaviour (Sander et al., 2003; Todorov, 2008; Vuilleumier, 2005). The concept of face distinctiveness is not in conflict with the idea that the amygdala is involved in evaluating and directing attention toward relevant stimuli. Instead, it suggests that the approach/avoidance system is not in itself sufficient to explain how multiple facial cues are processed by the brain, whereas the distance from an average face in terms of distinctiveness could be a simple and efficient property for highlighting stimuli that require additional evaluation (Said et al., 2010). The present results add support to this view. In particular, a common response was found in the amygdala and posterior face-selective regions to orthogonal dimensions with different social content, suggesting that all these areas are involved in coding face stimuli in terms of their distinctiveness as well as the social cues conveyed by facial features. Nonetheless, the theoretical explanation of why these regions are sensitive to this feature and the mechanisms underlying face evaluation remain difficult issues. Face distinctiveness could be considered an important cue per se; indeed it is spontaneously encoded from faces and less typical faces are better recognized (Santos and Young, 2005; Valentine, 1991). Therefore, these results could be interpreted by considering

that faces at the extremes of the stimuli matrix were processed as perceptually salient because of their distinctiveness, independently of their being varied along the trustworthiness or gender dimensions. This could have driven the quadratic response in the amygdala, because of its sensibility to the personal impact of the stimuli (Ewbank et al., 2009). This hypothesis is in line with the idea of the amygdala as detector of relevant events (Sander et al., 2003) and can account for different effects reported in previous fMRI studies, such as increased amygdala response when participants received increasing reward or punishment in a competitive game (Zalla et al., 2000), or quadratic amygdala activation when socio-biological facial features like self-resemblance and race were varied (Platek and Krill, 2009). On the other hand, amygdala activation is reported to increase linearly when modulated by the intensity of gustative or olfactory stimuli (Anderson et al., 2003; Small et al., 2003), or by the rated intensity of emotional faces (Sato et al., 2004) and socially relevant concepts (Cunningham et al., 2004). Further studies could investigate whether the effects in the posterior face-selective regions are due to a modulatory influence from the amygdala (Vuilleumier et al., 2004) or directly depend on the distance of faces from the average face (Loffler et al., 2005).

In summary, the results from this study could help to clarify how different face-selective brain regions respond to face stimuli in order to code cues that can be socially relevant. In line with the idea of the amygdala as a salient stimuli detector (Sander et al., 2003), previous findings of quadratic responses to face trustworthiness have been replicated (Said et al., 2008; Todorov et al., 2008). However, U-shaped activation pattern was not specific for this social dimension. Indeed, similar responses were observed in the amygdala and posterior face-selective regions (OFA, FFA, right pSTS) for faces varying along a gender dimension, suggesting that the images may be processed in terms of their distinctiveness from an average face. Future studies could explore this possibility by asking whether the average face

against which the images are coded as more or less distinctive is represented by a general average of the faces seen in a population or by the average of the faces presented in a specific context. This should be possible by creating an average face for the experiment that differs from the general population average of faces encountered in daily life.

3. Study 3: Emotion representation in a distributed cortical system

3.1. Introduction

Affective neuroscience research consistently points to the amygdala as a critical component of the system mediating emotion recognition (Adolphs, 1999; Calder et al., 2001). Results of Study 1 and Study 2 confirm this evidence, but, together with the amygdala, different cortical areas participate in expression discrimination (Adolphs et al., 1996; Phan et al., 2002). Among these, the medial prefrontal and right somatosensory cortices are considered two important structures part of the extended system for emotion processing (Adolphs, 1999; Haxby et al., 2000). However, it is still not clear how different emotions are represented in these areas and previous literature provides inconsistent results concerning the possibility of segregated circuits for distinct emotions in the distributed cortical network (Calder et al., 2001; Phan et al., 2002). In light of this, the present study aimed at clarifying whether the activation of the mPFC and rSC in emotion processing is specific for the type of emotion. To address this issue, an experiment was carried out using state-dependent TMS (Silvanto et al., 2008; Silvanto and Pascual-Leone, 2008), which is based on the assumption that TMS effects depend on the pre-existing activation state of the targeted neural population. This method allows perturbing specific neural populations, which could show distinct functional properties although spatially overlapped (Silvanto et al., 2008). Previous studies investigating emotion representation in mPFC and rSC reported mixed evidence concerning the type of emotion, which activates these regions or resulted to be affected by TMS with a classical virtual lesion approach (Kesler/West et al., 2001; Harmer et al., 2001; Pitcher et al., 2008; Pourtois et al., 2004). Possible explanations for these inconsistent results could be that different emotions produce different level of activity in the same cortical area or are represented by distinct neural populations included in the same region, therefore a TMS state-dependent paradigm could disentangle this issue.

Specifically, it was used a TMS-priming paradigm (Cattaneo et al., 2008) in which participants were primed with a word related to happiness or anger, and were then asked to indicate whether a face following the prime word was happy or angry (Carroll and Young, 2005). A TMS pulse was delivered before target onset. According to TMS state-dependent view (Silvanto et al., 2008; Silvanto and Pascual-Leone, 2008), combining TMS with a priming paradigm enables to assess the existence of possible functionally distinct neural representations for different emotions within the stimulated cortical area. In particular, a different TMS effect on primed or unprimed targets can reveal that the area contains neural populations that were selectively activated by the different primes processing (Cattaneo et al., 2008; Silvanto et al., 2010). On the contrary, if this is not the case and there are no distinct representations within the area, all the neural populations should respond equally regardless of prime category and no interaction between TMS and prime type would appear. Therefore, if the stimulated cortical region contains distinct neural representations for happiness and anger, priming to either one should differentially modulate the initial activation state of these populations, and TMS should interact with the priming effect. Specifically, TMS should have a different effect on emotion recognition depending on whether the prime word and the target face refer to same (congruent trials) or different (incongruent trials) emotions.

In line with previous evidence, the prediction was that TMS over the mPFC would have differentially affected performance depending on prime type (see Harmer et al., 2001; Phillips et al., 1998). Predictions for the TMS effects over the rSC were less straightforward: according to Pitcher et al. (2008), TMS should affect emotion discrimination regardless of the emotion type. However, according to other studies (Pourtois et al., 2004; Oberman et al., 2007), rSC recruitment in emotion recognition may vary depending on the extent of facial mimicry induced by different facial emotions presented. If this is the case and the intensity

of the single–pulse TMS is sufficient to disrupt embodied representations in rSC, TMS should affect emotion discrimination regardless of prime type, since the area is likely to be activated by faces processing but not by emotional words.

3.2. Method

Participants

Twenty healthy volunteers participated in the experiment. All subjects (8 male, 12 female, mean age = 22.3 years, s.d. = 2.4) were University students and gave written consent prior to their participation. All participants had normal or corrected to normal vision and no history of mental or neurological illness or other specific contraindications to TMS. The experiments took place in the TMS laboratory of the University Milano-Bicocca with the approval of the local Ethic Committee.

Material

Eight Italian words were used in the experiment as prime words (see Table 3.1). All words were chosen from the Corpus and Frequency Lexicon of Written Italian (COLFIS, see http://www.istc.cnr.it/material/database/colfis/index_eng.shtml), and included four "anger-related" words (rage, ire, aggressiveness, violence) and four "happy-related" words (joy, gaiety, happiness, good cheer). Prime words were chosen on the basis of a preliminary questionnaire administered to 14 undergraduate students (7 male, 7 female, mean age = 24.6), different from those participating in the TMS experiment. Subjects were asked to rate on a four-point Likert scale (1 = "not at all", 2 = "a little", 3 = "enough", 4 = "a lot") the relatedness of each word with anger and happiness. The questionnaire included 16 words for the two emotions; the four words of each emotion with the higher relatedness score for the target emotion and a score < 1.5 for the opposite emotion were selected. Independent sample

t-test on the selected words confirmed that there were no significant differences between the two categories in word length ($p = 0.75$) and total written frequency ($p = 0.64$). The neutral prime consisted of a string of eight "#" (where eight corresponded to the mean number of letters of the emotional prime words). A non-word string was used as neutral prime in line with a previous study (Cattaneo et al., 2010a), in order to avoid accidental systematic associations between non-emotional words and individual positive or negative feelings. The face stimuli were coloured photographs of 8 different unknown individuals, 4 male and 4 female, with either an angry or happy face. Photographs were chosen from the Bosphorus 3D Database (Savran et al., 2008) on the basis of a preliminary study with 20 additional students (10 male, 10 female, mean age = 24.1). Photographs were presented on a computer screen, displaying seven different expressions (neutral, anger, disgust, fear, happiness, sadness, surprise) together with the name of the seven emotions; participants were required to match each facial expression with the corresponding name. The photographs of the four individuals whose angry and happy expressions were more consistently identified (mean accuracy score > 75 %) were selected as stimuli for the main experiment.

Angry words		Happy words	
Italian	English	Italian	English
Gioia	Joy	Aggressività	Aggressiveness
Allegria	Gaiety	Collera	Rage
Contentezza	Happiness	Ira	Ire
Buonumore	Good cheer	Violenza	Violence

Table 3.1. Words related to anger and happiness used as primes.

Experimental Procedure

Figure 3.1 depicts the timeline of an experimental trial. Subjects were asked to judge as fast and as accurately as possible whether the target face expressed anger or happiness, pressing one of two buttons with the index and middle right hand fingers; response-button correspondence was randomized across subjects. Each experimental trial started with a

fixation point in the middle of the screen lasting for 500 msec and followed by a blank screen for 300 msec. Then, the prime word appeared for 250 msec, followed again by a blank screen (300 msec) and the target face stimulus, which remained on the screen until the subject responded. The experimental procedure included eight blocks, two for each stimulation site (mPFC, rSC, Vertex) and two for the baseline no-TMS condition, with a total of 192 trials for each experimental condition. Blocks order was counterbalanced across subjects.

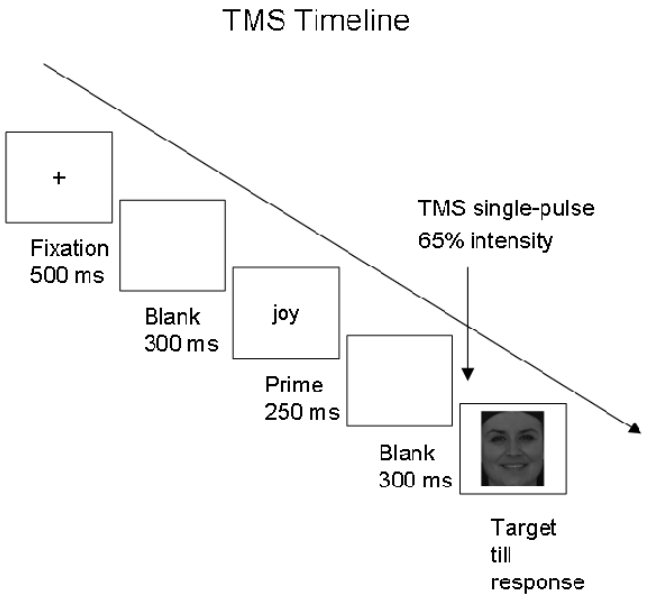

Figure 3.1. Timeline of a TMS experimental trial of the main Experiment. In each trial, the prime was a happiness-related word, an anger-related word, or a neutral non-word. The target was a face with either a happy or an angry expression. In the TMS conditions, a single-pulse TMS was applied over the medial prefrontal cortex (mPFC), the right somatosensory cortex (rSC) or Vertex at target onset.

TMS

In the TMS trials a single-pulse TMS was delivered immediately before the target onset by means of a Magstim Standard Rapid magnetic stimulator (Magstim, Whitland, UK) with a figure-of-eight coil (70 mm diameter) at 65% intensity of the maximum stimulator output. A fixed intensity was chosen on the basis of previous studies (e.g. Campana et al., 2002; Cattaneo et al., 2010a). The stimulated areas were the mPFC, rSC and the Vertex (control

site), in addition to a baseline condition without TMS. The face area of rSC was localized using the SofTaxic Evolution Navigator System (E.M.S., Bologna, Italy). This system allows the co-registration of the coil and subject's head positions and the localization on the scalp of the position corresponding to the cortex area of interest on the basis of the subject's MRI. Four subjects had their own T1-weighted structural MRI. When an individual MRI is not available, Softaxic allows computing an estimate MRI volume on the basis of a set of points registered from the subject's scalp. Talairach's coordinates for rSC (x = 51, y = -13, z = 29) were individualized on the basis of a previous fMRI study (Drevets et al., 2005). For mPFC stimulation, the coil was positioned on the scalp at one-third of the distance between the nasion and the inion on the midline between the left and right periauricular points (see Figure 3.2) (see Harmer et al., 2001, for similar procedure). The Vertex was localised as the point falling half the distance between the nasion and the inion on the same midline (Pitcher et al., 2008).

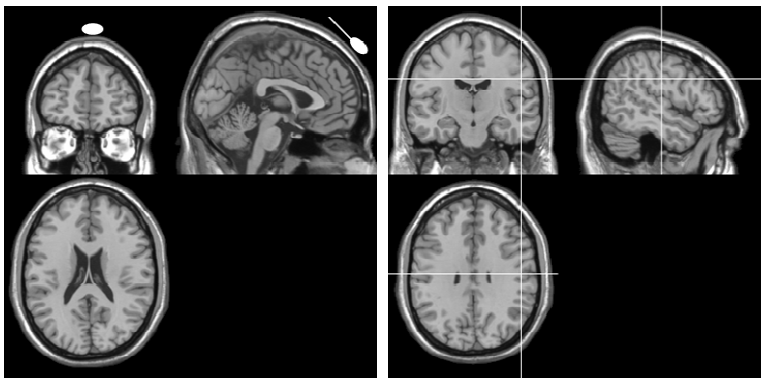

Figure 3.2. Normalized locations of mPFC and rSC. mPFC was localized at one-third of the distance from nasion to inion; the coil position is reported (left side). Localization of the face area in the rSC was based on Talairach coordinates 51, -13, 29 (right side).

3.3. Results

Trials were classified as "congruent" when the prime and the target face referred to the same emotion (e.g. joy and happy face), as "incongruent" when the prime and the target face referred to a different emotion (e.g. joy and angry face), and "neutral" when the prime was

neutral. Mean percentage accuracy for congruent trials was 95.1% in the baseline, 95.9% in the Vertex, 94.3% in the mPFC and 95.7% in the rSC condition. For incongruent trials, mean percentage accuracy was 95.6% in the baseline, 93.8% in the Vertex, 95.7% in the mPFC and 93.7% in the rSC condition. For neutral trials, mean percentage accuracy was 95.6% in the baseline, 95.5% in the Vertex, 96.6% in the mPFC and 95.5% in the rSC condition.

In order to consider possible TMS effects on both RTs and accuracy these two measures were incorporated in a single analysis by dividing RTs by the proportion of correct responses. This is a standard measure, which allows combining RTs and accuracy in a single performance score controlling for any potential speed-accuracy trade-offs across participants and conditions (Kiss et al., 2009; Brozzoli et al., 2008; Igarashi et al., 2007; Mevorach et al., 2006). One subject was excluded because his RTs (adjusted for accuracy) were > 2 SD the participants' mean. Hence, all the analyses were carried out on 19 subjects. In the baseline condition the mean adjusted RTs were faster in the congruent trials (532.36 msec) compared to the incongruent (535.14 msec) and the neutral trials (543.82 msec). However, a repeated measures ANOVA with Prime (three levels: congruent, incongruent, neutral) as within-subjects variable revealed that the effect of Prime was not significant [$F(2, 36) = 1.74$, $p = .19$][1].

To verify that the Vertex could be considered as a control condition, baseline and Vertex stimulation were compared by means of pairwise t-test for each prime type. Vertex and baseline did not differ in any experimental condition (congruent primes [$t(18) = 0.2$, $p = .84$], incongruent primes [$t(18) = -1.33$, $p = .2$], neutral primes [$t(18) = -0.007$, $p = .99$]).

[11] Critically, the same pattern of results was reported in a control behavioural experiment, carried out on 12 new subjects (5 male, 7 female; mean age = 23, S.D. = 2.98) using the Italian word "neutrale" ("neutral" in English) as neutral prime. Mean RTs (adjusted for accuracy) were faster for congruent trials (568.82 ms) than for incongruent (581.93 ms) and neutral trials (588.55 ms). A repeated measures ANOVA with Prime (three levels: congruent, incongruent, neutral) as within-subjects variable revealed that the effect of Prime was not significant [$F (2, 22) = .82$, $p = .45$]. These data rule out the hypothesis that the effects we reported in the baseline condition of our experiment depend on the use of a non-word neutral prime.

The effect of TMS on priming

Since baseline and Vertex did not show any difference, Vertex was used as unique control condition. To investigate whether TMS interfered with the priming effect, the difference between congruent and incongruent trials, considered as a measure of the congruent primes facilitatory effect, was compared among the three TMS conditions. Figure 3.3 shows the effect of TMS on the priming benefit.

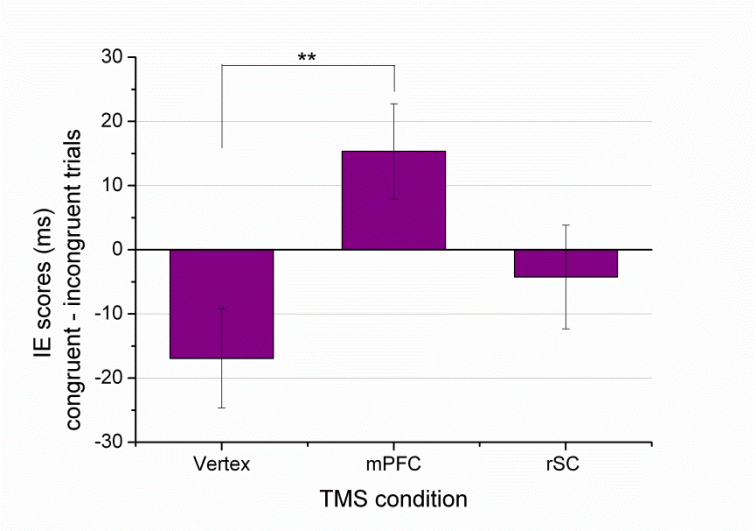

Figure 3.3. Priming effect expressed in msec (i.e. difference between RTs adjusted for accuracy in congruent and incongruent trials) in the three TMS conditions. Negative values indicate that target discrimination was faster on congruent trials than on incongruent trials. TMS over mPFC abolished the benefit of the congruent prime; the double asterisk indicates a significant effect ($p < .001$), error bars represent Standard error of the means.

A 3x2 repeated measures ANOVA on the priming effect with TMS (three levels: Vertex, mPFC, rSC) and Emotion (two levels: anger, happiness) as within-subjects variables showed a significant main effect of TMS [$F(2,36) = 4.45$, $p = .019$], while neither the effect of Emotion [$F(1,18) = 0.19$, $p = .67$] nor the interaction [$F(2,36) = 1.71$, $p = .20$] were significant. Bonferroni post hoc analysis showed a significant difference between Vertex and mPFC conditions [$t(18) = -5.06$, $p < .001$], whereas the difference between Vertex and rSC was not significant [$t(18) = -1.05$, $p = .92$].

To investigate whether the effect of TMS on priming benefit differently affected congruent or incongruent trials, a repeated measures ANOVA with TMS (three levels: Vertex, mPFC,

rSC), Emotion (two levels: anger, happiness) and Prime (two levels: congruent, incongruent) as within-subjects variables was performed on the mean response latencies adjusted for accuracy. The analysis showed a significant main effect of TMS [$F(2,36) = 3.57$, $p = .038$] and a significant interaction TMS x Prime [$F(2,36) = 4.45$, $p = .019$]. The main effect of Emotion [$F(1,18) = .02$, $p = .89$] and Prime [$F(1,18) = .18$, $p = .68$] and the remaining interactions were not significant. Pairwise comparisons on the TMS main effect revealed that adjusted RTs were longer when TMS was applied over the mPFC and the rSC as compared to the Vertex, but such difference was not significant when correcting for multiple-comparisons (mPFC-Vertex [$t(18) = -2.12$, $p = .048$], rSC-Vertex [$t(18) = -2.17$, $p = .043$], significance level < .025, according to Bonferroni correction).

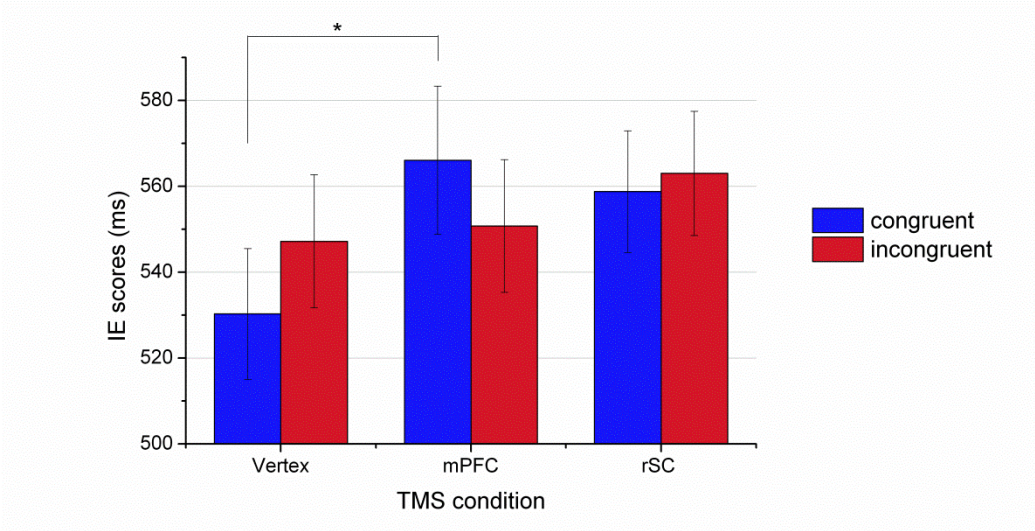

Figure 3.4. Effects of prime type on facial expression recognition for the three stimulated sites. TMS over mPFC interfered with facial expression discrimination when targets were preceded by congruent primes; an asterisk indicates a significant effect ($p < .05$), error bars represent Standard error of the means.

To further investigate the significant interaction TMS x Prime, simple main effect analyses of TMS for each prime were carried out collapsing together the two emotions. TMS was found to significantly affect congruent trials [$F(2,36) = 4.57$, $p = .017$] but not incongruent trials [$F(2,36) = 2.03$, $p = .15$] (see Figure 3.4). Pairwise t-tests showed that adjusted RTs increased for congruent trials when TMS was applied over mPFC as compared to the Vertex

[t(18) = -3.27, p = .004]. Critically, as shown in Figure 3.4, TMS over rSC seems to have an unspecific effect by overall increasing response latency regardless of prime type, although not to a significant extent ([t(18) = -2.01, p = .06] for congruent trials and [t(18) = -1.78, p = .092] for incongruent trials).

3.4. Control experiment

In order to exclude the possibility that results were due to TMS interfering with priming per se rather than specifically with emotional priming, a control experiment was carried out, in which a gender priming task was used and the same cortical sites as in the previous experiment were stimulated.

3.5. Method

Participants

Thirteen subjects participated in this experiment (5 male, 8 female, mean age = 23, s.d. = 2.8) in the TMS laboratory of the University of Milano-Bicocca. All subjects had normal or corrected to normal vision, no specific contraindication for TMS and they gave written consent to their participation.

Material and procedure

The same paradigm, TMS sites and procedure of the previous experiment were used, but the task was to judge as fast and accurately as possible whether the target face was a male or female. The Italian words “maschio” (male) and “femmina” (female) were used as prime words together with the word “vivente” (alive) as neutral prime. This word was chosen because in Italian it has the same number of letters as the other prime words and carries no

gender information. The face stimuli were the same eight individuals' photographs from the Bosphorus 3D Database (Savran et al., 2008), but with a neutral expression.

3.6. Results

Data were analysed using the same measures, namely RTs adjusted for accuracy, and steps of the affective priming experiment. Trials were classified as "congruent" when the prime and the target face referred to the same gender (e.g. male and male face), as "incongruent" when the prime and the target referred to different gender (e.g. male and female face), and "neutral" when the prime was neutral. Mean percentage accuracy for congruent trials was 97.4% at the baseline, 97.2% during stimulation of the Vertex, 96.4% during stimulation of the mPFC and 97% during stimulation of the rSC. For incongruent trials, mean percentage accuracy was 96.4% at the baseline, 96.4% in the Vertex, 96.9% in the mPFC and 95.8% in the rSC condition. For neutral trials, mean percentage accuracy was 97.8% at the baseline, 96.8% in the Vertex, 97.3% in the mPFC and 96.7% in the rSC condition.

At the baseline with no-TMS, adjusted RTs were 476.70 msec for congruent trials, 489.46 msec for neutral trials, and 491.26 msec for incongruent trials. A repeated measures ANOVA with Prime (three levels: congruent, incongruent, neutral) as within-subjects variable revealed that the effect of Prime was not significant [$F(2,24) = 1.95$, $p = .16$]. Baseline and Vertex conditions, compared by means of pairwise t-tests for each prime type, did not differ in any experimental condition (congruent primes [$t(12) = 0.34$, $p = .74$], incongruent primes [$t(12) = -0.02$, $p = .98$], neutral primes [$t(12) = 1.24$, $p = .24$]); hence, Vertex was used as unique control condition in the following analyses. The potential effect of TMS on priming benefit (difference between congruent and incongruent trials) was compared in the three TMS conditions. A 3x2 repeated measures ANOVA on the priming effect with TMS (three levels: Vertex, mPFC, rSC) and target Gender (two levels: female,

male) as within-subjects variables did not lead to any significant effect (TMS [$F(2,24) = 1.16$, $p = .34$], Gender [$F(1,12) = .20$, $p = .66$], TMS x target Gender interaction [$F(2,24) = .06$, $p = .94$]). As for the main Experiment, a further repeated measures ANOVA was carried out on the mean adjusted RTs with TMS (three levels: Vertex, mPFC, rSC), target Gender (two levels: male, female) and Prime (two levels: congruent, incongruent) as within-subjects variables. The analysis showed a significant main effect of Prime [$F(1,12) = 5.57$, $p = .036$], indicating that adjusted RTs were faster in congruent than in incongruent trials. Neither the effect of TMS [$F(2,24) = 1.23$, $p = .31$], nor of target Gender [$F(1,12) = 2.82$, $p = .12$] was significant. None of the interactions reached significance.

3.7. Discussion

This study investigated the role of the mPFC and the rSC in discriminating happy and angry expressions by using a TMS-priming paradigm. Results showed that TMS delivered over the mPFC at target onset significantly affected the priming effect, compared to the Vertex. In particular, stimulation of the mPFC selectively interfered with the discrimination of both angry and happy expressions when the prime was congruent with the target, but not when it was incongruent. According to the state-dependent view of TMS (Silvanto et al., 2008), this TMS-prime interaction suggests that the stimulated area contains distinct neural representations for the emotions of anger and happiness that were selectively activated by the prime. In other words, the presentation of a specific prime induced an activation imbalance between specific neural populations mediating the representation of the corresponding emotion within the targeted region, and TMS selectively interacted with this activity imbalance (Silvanto et al., 2008). Conversely, if the mPFC contained a common representation for both emotions, TMS would have affected response latencies regardless of the type of prime. This was indeed the case for rSC stimulation that did not significantly

interfere with the priming effect. As shown in Figure 3.4, TMS over this region led to a trend for a general impairment compared to the Vertex condition, regardless of the prime used. Critically, the effects we reported proved to be specific for emotion processing, since TMS did not affect a gender priming task in which the same faces used in the main experiment (but with a neutral expressions) were presented as targets (see Control Experiment).

These data suggest that the mPFC contains different representations for different emotions, and that these representations can be activated by an emotional word (and not only by presentation of an emotional face). This supports the hypothesis that the role of the prefrontal cortex in emotion recognition may be related to lexical knowledge of the facial expression. Accordingly, Adolphs et al. (2000) reported that patients with frontal damage failed in a verbal categorization task, in which emotion expressions had to be matched to the correct name. The results are also in line with Phillips et al. (1998b)'s fMRI study that found activation in the middle frontal gyrus (BA 32) during happy face perception and with Kesler/West et al. (2001), who found activation in the mPFC for both angry and happy faces processing. More specifically, Kesler/West et al. (2001) reported activation of BA 32 and 10 for happiness and superiorly in BA 9 for anger. In the present experiment, the coil was positioned over the mPFC; based on individual MRI available this area seems to correspond approximately to BA 32. However, considering TMS spatial resolution (Walsh and Cowey, 2000), we cannot exclude that BA 9 and 10 were also affected by stimulation. These findings extend previous neuroimaging evidence by showing that mPFC plays a causal role in emotion recognition and contains different neural representations for the emotions of anger and happiness.

In a previous TMS study, stimulation of the mPFC was found to selectively affect the processing of angry expressions, whereas happy expressions were unaffected (Harmer et al., 2001). Conversely, present findings suggest that mPFC encodes both happy and angry faces,

with different neural representations associated with the two emotions. It is possible that the different methodologies used here and in Harmer et al.'s (2001) study may account for the different outcomes. In particular, Harmer et al. (2001) assessed the effect of TMS in discriminating anger and happiness from neutral expressions in two separate tasks. They presented morphed faces with increasing expression intensity and they selected for the TMS experiment only those stimuli close to the recognition threshold of each subject. This task was probably more difficult than the one used in the present experiment, thus possibly leading to a difference in the discrimination between happy and angry faces that did not emerge in this study (i.e. the main effect of target emotion was not significant). In addition, Harmer et al. used repetitive TMS (four pulses applied at target offset) with a classical "virtual lesion" approach (Walsh and Cowey, 2000), while the TMS state-dependent paradigm assessed both emotional expressions in a single task. This approach allows a higher functional resolution relative to "virtual lesion" TMS, because one can control which neural population within the stimulated area is facilitated or inhibited by TMS (Silvanto et al., 2008).

Analysis of rSC-TMS showed that the stimulation over the rSC did not modulate the priming effect, although a trend for an overall impairment on emotion recognition was observed (regardless of the prime used and of the target emotion presented). This finding is in line with the hypothesis that rSC plays a role in face expressions recognition but is not differentially activated by different emotions (see Pitcher et al., 2008). The fact that TMS over rSC did not result into a clear impairment (as in the case of Pitcher et al., 2008) may be due to the specific TMS timing and parameters: as already mentioned, a single pulse of TMS was delivered at target onset, whereas Pitcher et al. (2008) used repetitive TMS. If, as suggested, the rSC plays a role in emotion recognition through mimicry simulation and the reactivation of the somatovisceral sensations linked to the perceived emotion (Niedenthal,

2007; Oberman et al., 2007; Adolphs et al., 2000; Adolphs, 1999), it is likely that stimulation in the present experiment occurred too early to interfere with this processing. Similarly, the early timing of TMS stimulation may have prevented possible differences related to the level of simulation induced by different emotions to emerge (see Pourtois et al., 2004). Conversely, present results clearly indicate that the rSC was not differently activated by words conveying an emotional meaning. Thus, it seems that the role played by the rSC in emotion recognition is specific for visual facial expressions and does not extend to more abstract concepts.

Previous TMS studies, which used state-dependent paradigms, have consistently found that, following adaptation, TMS facilitates the detection of adapted stimuli (e.g. Silvanto et al., 2007; Cattaneo et al., 2009). This suggested that TMS preferentially stimulated the less active neural populations relative to more active ones (Silvanto et al., 2008). However, the evidence on how TMS interacts with priming is less clear, with reports of facilitation of unprimed trials (Cattaneo et al., 2010a) as well as impairment of primed trials (Silvanto et al., 2010). The results of the present study support the latter solution, since TMS over mPFC affected trials with a congruent prime.

Another issue concerns the baseline no-TMS condition in which the priming task did not show an evident behavioural effect, since response latencies in congruent and incongruent trials were not significantly different. This may appear at first puzzling. However, similar results were obtained in a previous behavioural study (Carroll and Young, 2005) with affective priming tasks. Indeed, Carroll and Young (2005) found evidence of a facilitation of the congruent condition relative to the neutral condition, but no inhibition in the incongruent condition compared to either the neutral or congruent condition. The authors hypothesized that this was due to a "leakage" effect between emotions, since emotional categories in part overlap. In other words, an incongruent emotional prime may still prime the following target

by activating an "emotional" network, whereas this does not happen in the case of a neutral prime. There may be some degree of priming between emotional prime and incongruent target since face expressions and emotional words are never completely unrelated. Nonetheless, we were interested in assessing whether the mPFC contains separate neural representations for different emotions. Accordingly, analyses were performed on the priming effect (i.e. the difference between congruent and incongruent trials). TMS over the mPFC was found to differently affect emotional processing depending on the prime used, suggesting that anger and happiness words activated at least partially segregated neural circuits within the stimulated region. Notably, similar results have been found in previous studies, which have reported state-dependent TMS effects in spite of weak behavioural effect (Cattaneo et al., 2010b; Cohen Kadosh et al., 2010). Therefore, the present findings confirm the importance of the TMS state-dependent paradigm (Silvanto et al., 2008) as a tool to study high cognitive functions. However, the physiological basis of the state-dependent effects is not completely clear, and future studies should directly investigate how TMS interacts with neural mechanisms involved in affective priming combining TMS with neuroimaging or electrophysiological techniques such as fMRI or electroencephalography (EEG).

Finally, these results could also shed light on the debate concerning priming mechanisms with affective stimuli. The priming effect in this type of task has been interpreted as due to spreading activation between related concepts in a semantic network (Fazio et al., 1986; Fazio, 2001). According to this explanation related stimuli share some features, so that the congruent prime facilitates target recognition by activating these common features (Masson, 1995). The alternative hypothesis posits that affective priming is due to processes occurring at the response selection stage rather than at the semantic encoding stage, so that the longer latencies in the incongruent trials are due to a Stroop-like response conflict mechanism

(Wentura, 1999; De Houwer et al., 2002). Previous studies demonstrated that the mPFC is involved in response inhibition and control stimulus-response contingencies (Picton et al., 2007) and it is activated by the Stroop effect (Liu et al., 2004). According to the response-selection account, mPFC-TMS should have increased the priming effect, resulting into a larger difference between congruent and incongruent trials. On the contrary, mPFC-TMS selective affected congruent trials with longer response latencies when prime and target referred to the same emotion, while incongruent trials were unaffected. This suggests that the priming effect likely depends on spreading activation among related concepts (Masson, 1995).

In summary, the present data contribute to clarify the role of the mPFC and of the rSC in the network underlying affective processes. The mPFC was found to contain selective representations for angry and happy emotions supporting previous evidence indicating that the mPFC implements both negative and positive emotions (Davidson and Irwin, 1999). On the other side, the rSC did not seem to be involved in representing emotional concepts.

4. Study 4: Cortical responsiveness in face processing: a TMS-EEG study

4.1. Introduction

In light of the results from Study 3, which demonstrated the critical role of mPFC in expression discrimination, the next interesting step is to understand how this area is linked to other components of the face processing network. Indeed, despite a wide literature on the neural correlates of face perception and the functional selective role of the regions within the core system (Calder and Young, 2005), the interactions among different areas have only recently become a question for research. Neuroimaging studies have investigated by means of DCM analyses the functional organization in the distributed network of face perception showing that both feed-forward and top-down connections could be modulated by the type of stimuli and task (Summerfield et al., 2006; Fairhall and Ishai, 2007). This evidence has suggested that the connectivity between the core and the extended system is probably more complex than previously thought (Haxby et al., 2000) and that the coupling between areas could have a relevant role in face processing (Ishai, 2008). Thus, new studies point to understand which variables modulate the neural coupling and the temporal dynamic of the connections within the network (Barbeau et al., 2008).

The temporal dynamics can be investigated non-invasively thanks to the recent development of the combined TMS-EEG. This technique allows a direct measurement of the excitability and effective connectivity of the human cerebral cortex combining together the advantage of the TMS to directly manipulate the cortical activity, and the high temporal resolution of the EEG recording (Taylor et al., 2008). The analysis of the TMS-evoked potentials (TEPs) can provide information about the timing in which different regions are involved in a behavioural task and how the neural signal is distributed and modulated during cognitive processing (Miniussi and Thut, 2010). The TMS-EEG has been used to show that visual attention for specific features of the stimuli modulate the spreading of activation from

anterior towards posterior regions and that cortical reactivity to the TMS perturbation is task-dependent (Morishima et al., 2009; Johnson et al., 2012). Moreover, in the face processing domain, Sadeh et al. (2011) used TMS-EEG to demonstrate a causal link between the activity in the occipital face area and the amplitude of the face-specific N170 component recorded in the temporo-occipital electrodes.

In light of these data, the present study aimed at measuring local cortical excitability and long-range connectivity within the face processing network by means of combined TMS-EEG. TMS was applied over the medial prefrontal cortex (mPFC) 100 msec after face stimulus onset during a face identity or a face expression matching task, while continuous EEG was recorded using a 60-channel TMS-compatible amplifier. Temporo-occipital ERPs recorded in separate blocks with and without stimulation were compared in order to investigate TMS effects on face-related components. Moreover, task-dependent modulation of local and distributed cortical excitability was examined by analyzing TEPs amplitude in the frontal electrodes near the TMS site and in temporal and occipital areas during the face tasks and a passive point fixation.

4.2. Method

Participants

Twelve healthy volunteers (6 male, 6 female, mean age = 31.4 years, s.d. = 8.4) participated in the study. One participant was excluded from the analyses because of a high number of trials rejected due to signal noise in the EEG registration. All participants gave written informed consent prior to their participation. The study took place in the TMS-EEG laboratory of the University of Milano-Bicocca with the approval of the local Ethic Committee.

Procedure

Stimuli consisted in face photographs from the Ekman series (Ekman and Friesen, 1976): three different female individuals posing three different expressions (happy, fear, neutral) were selected. Stimuli were presented in the centre of a computer screen covering a visual angle of ~8° x 11° for 700 msec, interleaved with a fixation cross which remained on the screen for a randomized interval between 1200 and 1400 msec. Participants were asked to maintain the central fixation during the experiment. Each block consisted in 180 presentations of the face stimuli. In the face expression task the same expression was repeated consecutively twice in the 15% of trials and participants were instructed to respond with a right-hand button when the expression repetition occurred. In the face identity task participants were instructed to respond for repeated identity (15% of trials). Stimulus order was controlled to avoid repetition of identical stimuli (same identity and same expression). In the TMS condition a single TMS pulse was delivered over the mPFC 100 msec after face stimulus onset. To ensure a sufficient number of good trials in the TMS condition both face tasks were repeated twice during the experimental session. In the no task condition TMS was applied during a passive point fixation; pulses were separated by 1900-2100 msec in order to maintain the same pulse-interval as in the face task conditions. Therefore, for each subject the experiment consisted in 6 blocks: one expression task and one identity task with only ERPs recording, two TMS expression tasks, two TMS identity tasks, one TMS no task block. The order of the ERPs-task, TMS-task and TMS no task conditions and the order the two face tasks within each condition was counterbalanced across subjects.

TMS stimulation

TMS was delivered with an Eximia TMS stimulator (Nextim, Helsinki, Finland) using a focal bi-pulse, figure of eight 70-mm coil. The coil was positioned between AFZ-FZ

electrodes targeting the first medial prefrontal gyrus in the right hemisphere. TMS target was identified in each subject using a Navigated Brain Stimulation (NBS) system (Nextim, Helsinki, Finland) that uses infrared-based frameless stereotaxy to map the position of the coil and subject's head within the reference space of the individual's high resolution MRI space. The NBS system estimates the electrical field induced by TMS taking into account head shape, distance from scalp, coil position and orientation. TMS was delivered at an estimated mean intensity of 101 ± 6 V/m (62 ± 3 % of the stimulator output). The TMS click sound was covered by playing a masking noise reproducing the time-varying frequency components of the TMS click into earplugs worn by the subjects during the experimental sessions (Massimini et al., 2005; Rosanova et al., 2009).

EEG recording and analysis

EEG was recorded with a 60-channel TMS compatible amplifier (Nextim; Helsinki, Finland), which uses a sample-and-hold system to hold the amplifier output constant from 100 μsec pre- to 2 msec post-TMS pulse avoiding amplifier saturation (Virtanen et al., 1999). Two electrodes placed over the frontal sinuses were used as reference and ground, and eye movements were recorded with two additional electrodes placed near the eyes. Electrodes impedance was kept below 5 kΩ and data were recorded with a rate acquisition of 1450 Hz. Data were pre-processed using SSP Biomedical Data Analysis Package, Version 1.12 (SiSyPhus Software, 2010) running in Matlab R2011b (Mathworks, Natick, MA, USA). Data were down-sampled to 725 Hz, continuous signal was split in trials starting 800 msec pre- and ending 800 msec post-TMS pulse, trials with excessive artefacts were removed by visual inspection and a band-pass filter between 2-80 Hz was applied. Data were then re-referenced to a common baseline between -300 and -80 msec before TMS pulse. The effect of TMS on temporo-occipital responses was examined considering the averaged signal from

contiguous electrodes in the occipital (left: PO3 - O1, midline: POZ - OZ, right: PO4 - O2) and temporal areas (left: TP9 - TP7, right: TP8 - TP10). The peak-to-peak amplitude of the P1-N1 and N1-P2 components was measured in each subject in order to compare the EEG signal in the expression and identity task in the TMS and no-TMS condition by means of repeated measures ANOVA TMS (yes/no) x task (expression/identity) x side (left/right/midline).

Further analyses were performed in order to examine task-specific TMS effects by subtracting ERPs on no-TMS conditions from those in the TMS conditions for each face task (Morishima et al., 2009). This allowed comparing TEPs in the TMS expression, TMS identity and TMS no task conditions. Since these analyses included electrodes near the stimulator, individual electrodes with excessive noise were interpolated using spherical spline interpolation (Perrin et al., 1998), then independent component analysis (ICA) was used to identify and remove muscle and residual TMS related artefacts (Korhonen et al., 2011; Johnson et al., 2012). Analyses included electrodes located in sites of interest, namely the prefrontal region below the stimulator, the temporal and occipital electrodes where face-specific ERPs components could be recorded (Bentin et al., 1996). Signal from contiguous sensors in the frontal (left: AF1- F1, midline: AFZ - FZ, right: AF2 - F2) temporal (left: TP9 - TP7, right: TP8 - TP10) and occipital regions (left: PO3 - O1, midline: POZ - OZ, right: PO4 - O2) was averaged and four time windows were defined for each region based on visual inspection of EEG components. Statistical analyses were then conducted with SPSS software considering the mean signal within each time window as dependent variable in a repeated measures ANOVA, condition (TMS expression/TMS identity/TMS no task) x side (left/right/midline).

4.3. Results

Behavioural performance

Participants detected 35.5 % of repetitions of the same expression (mean RT 580.4 msec) and 52.3 % of identity repetitions (mean RT 555.5 msec) in the face tasks when no TMS was delivered. In the TMS blocks detection rate was 39.1 % in the expression task (mean RT 578 msec) and 49.9 % in the identity task (mean RT 552.1 msec). Reaction times and accuracy were collected only during the 700 msec of face presentation whereas participants' responses given after stimulus offset were not recorded for the analysis; this likely explains the low level of accuracy in task performances. Repeated measures 2x2 ANOVA with TMS and task as within factors revealed a significant main effect of task both for accuracy [$F(1,10) = 6.7$, $P = .027$] and RT [$F(1,10) = 16.7$, $p = .002$] but no significant effect of TMS or interactions. As previously reported (Campbell at al., 1996; Münte et al., 1998), participants were more accurate and faster in detecting identity than expression repetitions, whereas TMS did not significantly affect the tasks. The absence of TMS effects on behavioural performance is likely due to the use of a single pulse paradigm instead of repetitive TMS (Harmer et al., 2001).

EEG results

Face stimuli presentation during EEG recording typically produces temporo-occipital evoked responses, which are considered face-specific because components are larger for processing faces than other categories of objects, in particular the negative component between 130 and 200 msec (Rossion and Jaques, 2008). In this experiment (see Figure 4.1), face presentation in both the expression and identity tasks elicited a first positive component (P1) at 100 msec in the temporo-occipital electrodes followed by a negative deflection at 150 msec (N1) and a second positive component after 200 msec (P2). The effect of mPFC-TMS on these posterior

components was examined by comparing the peak-to-peak amplitude of P1-N1 and N1-P2 in the expression and identity task in the TMS and no-TMS conditions.

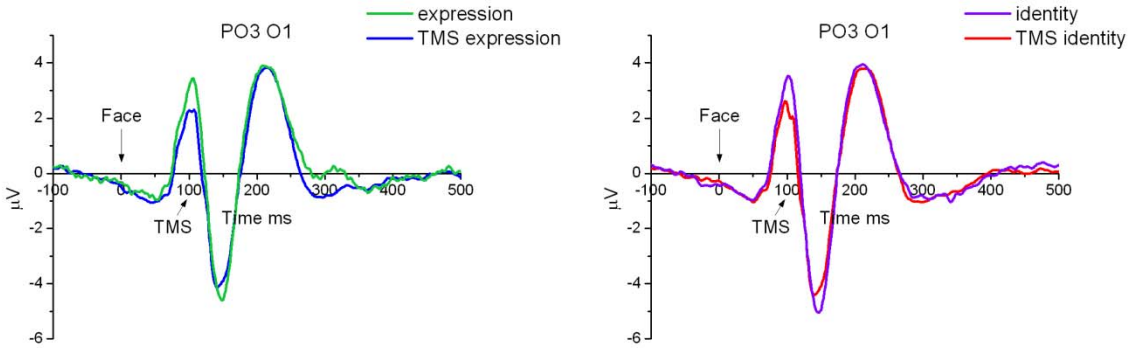

Figure 4.1. Scalp potentials recorded in the occipital electrodes during the expression behavioural task in the TMS and no-TMS condition (left) and the identity behavioural task in the TMS and no-TMS condition (right). Figure depicts signal recorded in the left occipital electrodes.

In the occipital electrodes a repeated measures ANOVA on P1-N1 amplitude with factors TMS (yes, no), task (expression, identity) and side (left, midline, right) revealed a significant effect of TMS [$F(1,10) = 8.2$, $p = .017$] and side [$F(2,20) = 18.01$, $p < .001$]. The P1-N1 amplitude was larger in the left and right electrodes than in the midline ($p < .001$ and $p = .001$ respectively, Bonferroni corrected) and it was significantly reduced in the TMS condition with no differences between expression and identity task. The ANOVA on N1-P2 amplitude revealed only a significant effect of side [$F(2,20) = 13.62$, $p < .001$] due to larger responses in the left and right electrodes than in the midline ($p < .001$ and $p = .005$), but no other significant effects. The same analyses were performed on the left (TP9, TP7) and right temporal (TP8, TP10) electrodes but no significant results were found (all $p > .05$).

Briefly, the main result of these analyses was that the amplitude of the first occipital component P1-N1 was reduced when TMS was applied over the mPFC 100 msec after face onset.

Task-specific TMS effects

Figure 4.2 depicts scalp potentials recorded during the different experimental conditions in the electrodes below the stimulator. As the figure shows, the frontal ERPs time-course of the face tasks is characterized by a first negativity at 100 msec, followed by a positivity at 150 msec and a second negativity about 220 msec after face onset (Eimer and Holmes, 2002).

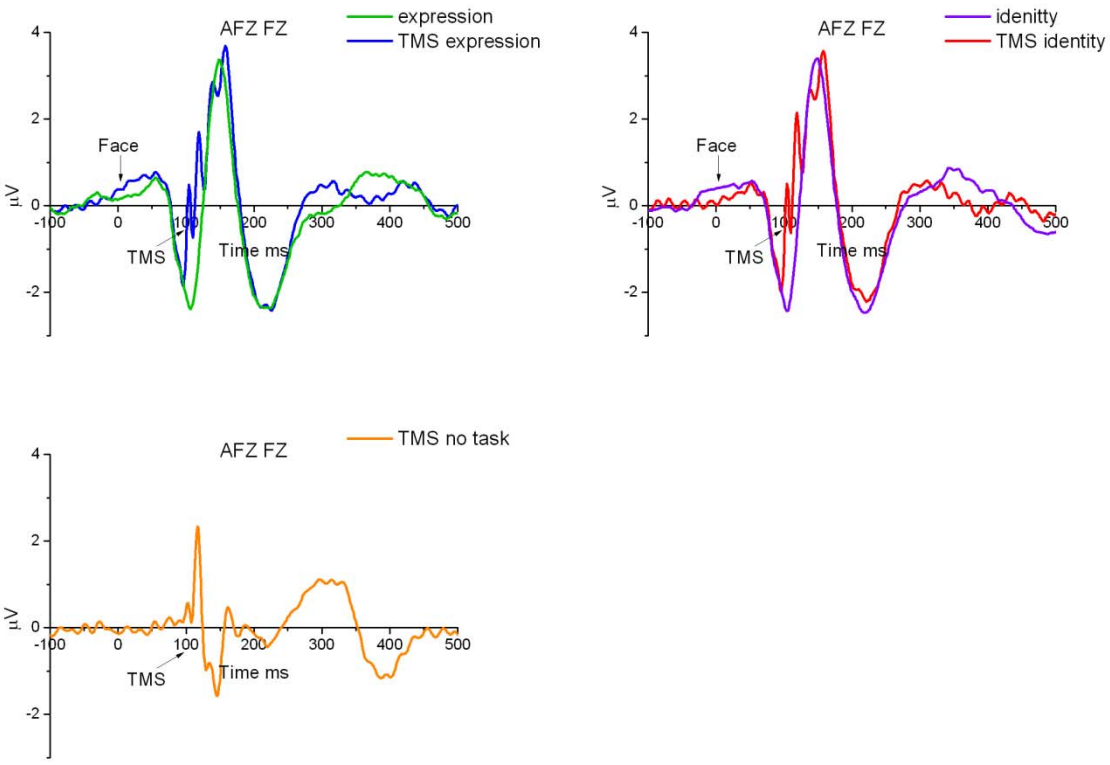

Figure 4.2. Scalp potentials recorded in the frontal electrodes during the expression task in the TMS and no-TMS condition (top left), the identity task in the TMS and no-TMS condition (top right) and in the TMS condition during the point fixation (bottom).

In reporting results $msec_{TMS}$ is used to specify time from the TMS pulse. In the no task condition TMS produced a TEP with a positive peak at 20 $msec_{TMS}$ followed by a negative deflection and oscillation lasting until 400 $msec_{TMS}$ after the pulse. In the TMS face tasks condition the TEPs waveform is partially overlapped to the ERPs waveform. Therefore, task-specific TMS effects were compared by subtracting the ERPs in the no-TMS condition from those in the TMS condition for each face task (Morishima et al., 2009); then the mean signal

in the time windows of interest was analysed by means of repeated measures ANOVA with condition (TMS expression, TMS identity, TMS no task) and side as factors (see method section for details).

Frontal TMS-evoked potentials

In the frontal electrodes the four time windows identified were T1: 0-25 $msec_{TMS}$, T2: 25-55 $msec_{TMS}$, T3: 55-140 $msec_{TMS}$, T4: 140-250 $msec_{TMS}$ (Figure 4.3).

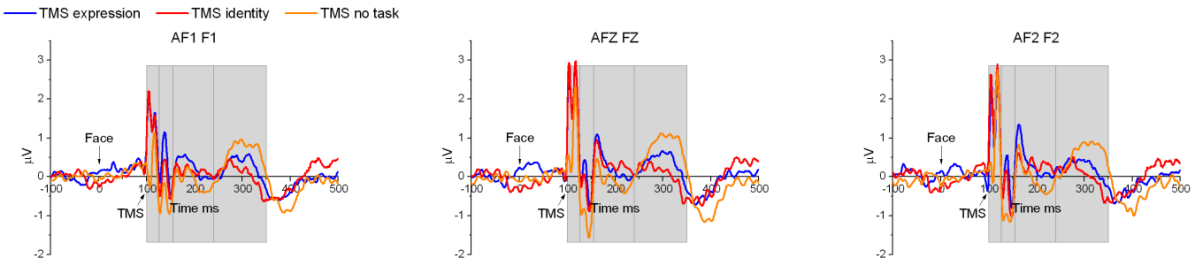

Figure 4.3. Mean TEPs in the frontal electrodes for each experimental condition. Blue and red lines represent TEPs in the TMS expression and TMS identity conditions after the subtraction of the respective ERPs in the no-TMS blocks. Orange lines represent TEPs in the no task condition. The shaded areas represent the four time windows considered in the analyses.

Within 0-25 $msec_{TMS}$ a 3x3 repeated measures ANOVA revealed a significant effect of condition [$F(2,20) = 9.05$, $p = .002$] and side [$F(2,20) = 6.16$, $p = .008$]. Post hoc tests (Bonferroni correction) showed that TMS no task significantly differed from TMS expression ($p =. 001$) and TMS identity ($p = .036$). Moreover, TEPs in the midline electrodes were significantly larger than TEPs in the left electrodes ($p = .043$). The main effect of side was marginally significant also at T2 (25-55 $msec_{TMS}$) [$F(2,20) = 3.52$, $p = .049$], whereas there were no significant results at T3. The effect of condition was significant also at T4 [$F(2,20) = 7.93$, $p = .003$], since TMS no task significantly differed from the TMS identity condition ($p = .02$).

In summary, TMS applied during the face tasks produced larger frontal TEPs in the first time window after stimulation (0-25 $msec_{TMS}$) than during point fixation, whereas between 140-250 $msec_{TMS}$ TEPs were larger in the TMS no task condition.

Temporal TMS-evoked potentials

The same time windows as in the frontal electrodes were identified in the left and right temporal electrodes (Figure 4.4).

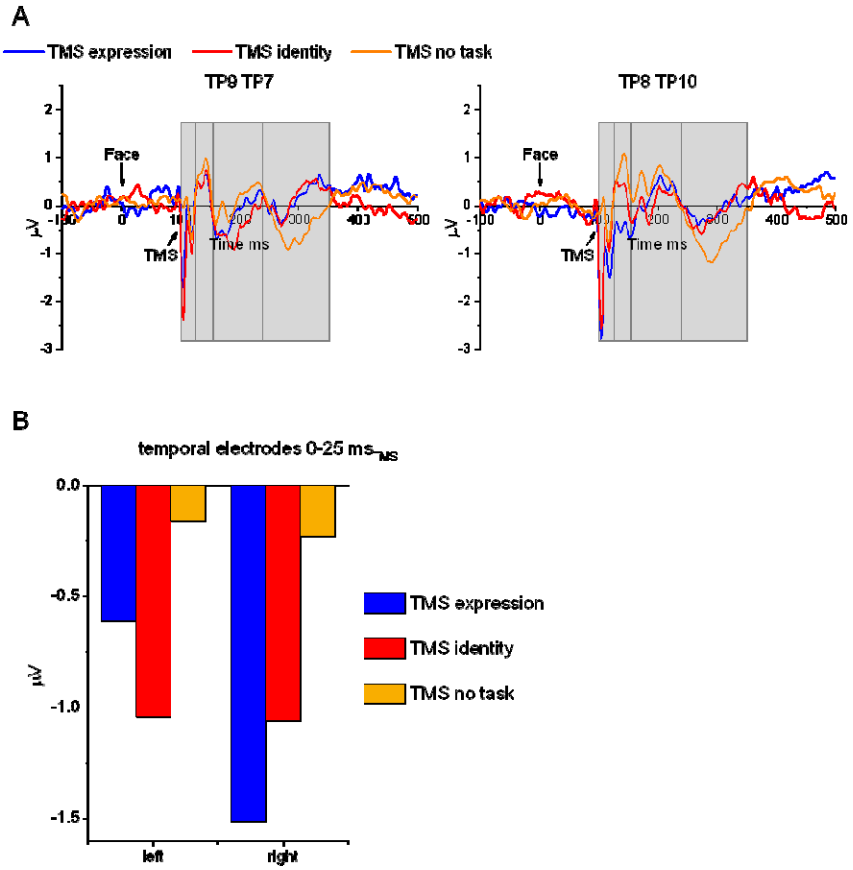

Figure 4.4. A) Mean TEPs in the temporal electrodes for each experimental condition. Blue and red lines represent TEPs in the TMS expression and TMS identity conditions after the subtraction of the respective ERPs in the no-TMS blocks. Orange lines represent TEPs in the no task condition. The shaded areas represent the four time windows considered in the analyses. B) Averaged signal in the first time window in the left and right temporal electrodes for the TMS expression, TMS identity and TMS no task conditions.

In this region within 0-25 $msec_{TMS}$ the 3x2 repeated measures ANOVA showed a significant effect of condition [$F(2,20) = 7.99$, $p = .003$] due to more negative TEPs in the TMS

expression than in the TMS no task condition (p = .008, Bonferroni correction). Notably, temporal electrodes in this early time window showed also a significant two-way interaction condition x side [F(2,20) = 3.62, p = .045]; planned t-test comparisons of left and right TEPs for each condition revealed significantly larger TEPs in the right than in the left electrodes only for the TMS expression condition [t(10) = 2.53, p = .03], but not for the TMS identity [t(10) = .06, p = .95] and no task condition [t(10) = .5, p = .63].

No significant results were found at T2 and T3, whereas in the later time window (140-250 $msec_{TMS}$) the main effect of condition [F(2,20) = 14.7, p < .001] and side [F(1,10) = 7.9, p = 018] were significant. TEPs were more negative in the right side electrodes and larger in the TMS no task than in the TMS expression (p = .004) and TMS identity (p = .005) conditions.

In brief, temporal electrodes showed greater TEPs in the TMS expression condition with a specific increase in the right electrodes between 0-25 $msec_{TMS}$. Then, at 140-250 $msec_{TMS}$ TMS had greater effects in the no task condition and TEPs resulted overall larger on the right side.

Occipital TMS-evoked potentials

In the occipital electrodes the time windows identified were T1: 0-28 $msec_{TMS}$, T2: 28-80 $msec_{TMS}$, T3: 80-150 $msec_{TMS}$, T4: 150-260 $msec_{TMS}$ (Figure 4.5).

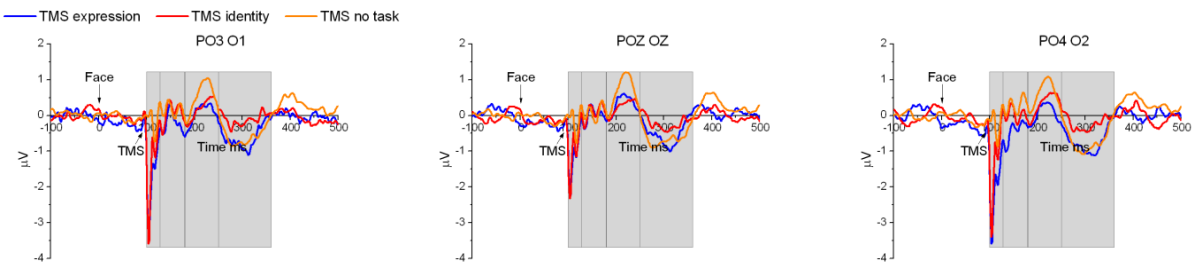

Figure 4.5. Mean TEPs in the occipital electrodes for each experimental condition. Blue and red lines represent TEPs in the TMS expression and TMS identity conditions after the subtraction of the respective ERPs in the no-TMS blocks. Orange lines represent TEPs in the no task condition. The shaded areas represent the four time windows considered in the analyses.

In the earlier time window a 3x3 repeated measures ANOVA showed a significant effect of condition [$F(2,20) = 6.85$, $p = .005$] and side [$F(2,20) = 10.54$, $p = .001$]; also the interaction condition x side was significant [$F(4,40) = 3.72$, $p = .011$]. TEPs were larger in the TMS expression and TMS identity compared with the TMS no task condition ($p = .013$ and $p = .039$, respectively) and midline electrodes significantly differed from left ($p = .008$) and right ($p = .020$) electrodes. To further investigate the significant interaction condition x side, simple main effect analyses of side for each condition were carried out. Side effect was significant in the TMS expression condition [$F(2,20) = 10.77$, $p = .001$] due to smaller TEPs in the midline than in the left ($p = .039$) and right electrodes ($p = 0.14$). Differently, there were no effects of side in the TMS identity [$F(2,20) = 2.27$, $p > .05$] and TMS no task [$F(2,20) = 0.84$, $p > .05$] conditions. Analyses at T2 showed no significant main effect of condition or side ($p > .05$), but a significant interaction condition x side [$F(4,40) = 2.89$, $p = .034$]. Simple main effect analyses of side for each condition revealed a significant side effect only in the TMS no task condition [$F(2,20) = 4.85$, $p = .019$], due to larger TEPs in the right than midline electrodes ($p = .043$). On the contrary, between 28-80 $msec_{TMS}$ there were no differences among left, midline and right TEPs in the TMS expression or identity conditions ($p > .05$). Finally, in the occipital electrodes no significant effects were found in the later T3 and T4 time windows.

To conclude, as in the previously described frontal and temporal regions, TEPs recorded at occipital electrodes were larger during the face task conditions in the first time window. Moreover, the significant interaction revealed larger TEPs in the left and right than in the midline electrodes during the expression task, whereas between 28-80 $msec_{TMS}$ TEPs were larger in the right electrodes in the case of point fixation.

4.4. Discussion

Neuroimaging studies have shown that a distributed cortical system mediates face processing (Haxby et al., 2000; Ishai et al., 2005), but the interactions among areas remained an open question (Fairhall and Ishai, 2007; Davies-Thomson and Andrews, 2012). This study directly assessed by means of TMS-EEG the cortical excitability and neural connections between the extended and the core system part of the face processing network. Results showed that TMS delivered over the mPFC 100 msec after face onset produced a specific reduction in the P1-N1 amplitude recorded at occipital electrodes. Moreover, behavioural tasks requiring face expression or face identity processing modulated the TEPs amplitude in the frontal electrodes near the stimulation site, as TEPs amplitude recorded in the temporal and occipital electrodes in the very early time window after the TMS pulse with differential effects of hemisphere in the case of the expression task.

The N1 component found in this study likely corresponds to the face-specific N170 reported in the EEG literature. This component is traditionally considered linked to the structural encoding of faces (Bentin et al., 1996) and not influenced by attentional and cognitive modulations, which are instead reflected in later components originated in frontal regions (Eimer and Holmes, 2002; Ashley et al., 2004). Conversely, the present results suggest an early top-down modulation during face processing by prefrontal regions. In line with these data, recent studies have shown that the amplitude of the face-selective N170 could be affected by emotional expressions (Wronka and Walentowska, 2011) and by the valence of the context in which faces are presented (Galli et al., 2006), supporting the hypothesis of a top-down modulation from associative areas at an early stage of face processing. The role of the top-down influence between the frontal and the face-sensitive visual areas has been shown also in fMRI studies which reported increased connectivity from the prefrontal cortex towards posterior face-responsive regions during mental imagery of faces (Mechelli et al.,

2004) or perceptual decision about faces (Summerfield et al., 2006). These latter studies highlighted the crucial role of the long-range connections and the importance of considering a complex cortical network in studying face processing (Ishai, 2008), taking advantage from the good spatial resolution of the fMRI technique. The results of the present study confirm long-range effects during face processing. Moreover, thanks to the higher temporal resolution of the TMS-EEG technique, which allowed recording EEG signal since 2 msec after TMS pulse, these data shed light on the temporal dynamics of neural transmission: indeed, the reduced amplitude of P1-N1 following mPFC-TMS demonstrates a causal link between activity in the mPFC and modulation of the early ERPs component in the posterior occipital area.

Analyses of the task-specific TMS effects showed that, as expected, frontal TEPs increased during face behavioural tasks, regardless the type of task, and were larger in the electrodes below the stimulator than in the opposite hemisphere. These local effects within 25 msec from the TMS pulse confirmed that task performance modulates cortical excitability (Johnson et al., 2012). Interestingly, analyses of the temporal electrodes in the same early time window revealed selective effects for the expression task, which showed larger TEPs compared to the no task condition and specific increase in the right hemisphere. This could be interpreted as an increased responsiveness to the mPFC stimulation of neurons in the right temporal region critical for explicit emotion discrimination. Since the modulation of mPFC excitability was unspecific for type of task, whereas the larger TEPs effect in the right hemisphere was specific for the expression task, an increased responsiveness of temporal neurons rather than enhanced neural output from the prefrontal cortex can be hypothesized. Enhanced connectivity between the temporal fusiform gyrus and the amygdala during view of emotional faces has been previously demonstrated by means of fMRI (Fairhall and Ishai, 2007). The present results confirm changes in connectivity between the core and the

extended system of face processing depending on the encoding of different facial features. Furthermore, this TMS-EEG study allowed identifying a causal link between activity in mPFC and temporal regions, showing that functional coupling between these areas occurred at an early stage of face processing (the TMS pulse was delivered 100 msec after face onset) and was modulated by the type of behavioural task requiring explicit processing of face expression rather than face identity. This clarifies the dynamics of the cortical connectivity within the face processing network, showing modulation of neural transmission within 25 msec after TMS pulse and a selective increase of connectivity within the right hemisphere during face expression discrimination.

Similarly to what detected in the temporal region, TEPs recorded within 28 msec from TMS in the occipital electrodes showed increased responsiveness to mPFC stimulation during the face tasks as compared to the no task condition. Following anatomical fronto-occipital connections, which are symmetrically distributed in the two hemispheres (Gschwind et al., 2012), TEPs were larger in the left and right electrodes than in the midline. Crucially, the symmetrical TEPs increase in the right and left electrodes was specific for the TMS expression condition, suggesting that TEPs propagation in the two hemispheres towards the occipital cortex was enhanced during the face expression task. On the contrary, in the absence of face task performance TEPs propagation in the occipital region remained within the same hemisphere where stimulation occurred, with larger signal recorded between 28-80 msec in the right electrodes than in the midline. Cortical connections in the occipito-temporal face network have been shown in a previous TMS-EEG study (Sadeh et al., 2011), which demonstrated a causal link between stimulation of face category-selective occipital cortex and increasing of the correspondent category-specific ERPs component. More long-range connectivity between prefrontal and posterior visual areas has been assessed (Morishima et al., 2009) looking at TEPs transmission in different neural networks

depending on the type of attended stimulus. Here the same face stimuli were used in two separate behavioural tasks in order to investigate neural transmission in a specific cortical circuit; results showed that explicit processing of different facial features modulated long-range effects within the network. This also brings out the sensitivity of the TMS-EEG as a technique to probe individual cortical networks that are involved in specific task performance.

It is worth noting that in frontal and temporal areas there were also later effects at 140-250 msec in the no task condition. Temporal electrodes showed larger negativity while frontal electrodes showed increased positivity during point fixation than during the face tasks. Even though analyses of the time-frequency domain were not carried out, this seems coherent with the long-lasting oscillations produced by TMS, which have been shown to persist until 300 msec after stimulation (Ferrarelli et al., 2010; Massimini et al., 2005). In the case of face task performance these long-lasting oscillations are reduced. One possibility to be addressed in the future is that the cognitive task interfered with the natural temporal development of the oscillations since these areas are involved in processing the face stimuli.

A wide ERPs literature have identified temporo-occipital and frontal components specific for face perception and modulated by emotional expressions (Bentin et al., 1996; Eimer and Holmes, 2007), while TMS studies have demonstrated that mPFC contributes to facial expressions processing (Harmer et al., 2001). However, interactions between frontal and posterior regions and timing in which emotional encoding occurs were still opened questions (Ashley et al., 2004; Vuilleumier and Pourtois, 2007; Wronka and Walentowska, 2011). In this study, TMS was applied on mPFC 100 msec after face onset, which corresponded to the timing of the first frontal negativity (see Fig. 4.2 and Wronka and Walentowska (2011) for previous example). The significant effect on the amplitude of the occipital P1-N1 component supports the hypothesis of a top-down regulation already at this early stage of stimulus

perception. Moreover, specific effects on neural transmission in the right temporal and occipital electrodes during the face expression task suggest that the mPFC is involved in face tasks by modulating cortical activity in posterior regions deputed to process facial features related to emotion discrimination (Pessoa et al., 2002). These results might be relevant in studying anomalous functional activity in clinical populations like autism spectrum or mood disorders which show impairments in interpreting facial emotions and abnormal activations in the emotion-related brain circuit (Wang et al., 2004; Leppänen, 2006).

In summary, by means of combined TMS-EEG it has been shown that perturbation of mPFC in the early stage of face processing affects the activity recorded at the occipital electrodes suggesting an immediate top-down modulation within the face perception circuit. Cortical excitability in the fronto-temporo-occipital network was also affected by the type of task, with changes in the neural transmission from mPFC towards the posterior regions within 30 msec time-range. In particular, the results suggest that explicit processing of facial expression is associated with both an increase of the functional coupling between prefrontal and right temporal regions and an enhancement of fronto-occipital connectivity.

General Discussion

People continuously interact with each other, thus the efficient perception of expressions and personality traits in faces of others is crucial to maintain an adaptive behaviour in the social environment. In the past years the face perception system has been a very active research field in neuroscience and different models have been proposed, moving from hypotheses of functional distinctive brain areas with selective role in face processing (Haxby et al., 2000; Calder and Young, 2005), towards the idea of a network with more distributed and interactive mechanisms (Ishai et al., 2005; Ishai, 2008). Taking advantage from the use of different neuroimaging and electrophysiological techniques, the studies presented in this thesis have investigated the neural correlates of face expressions processing. In particular, the specific contribute of distinct regions part of the extended system has been addressed, taking into account also the role of interactions among different areas and connections between the core and the extended system (Haxby et al., 2000).

Study 1 and Study 2 aimed at clarifying, by means of fMRI, amygdala responses to basic emotions and trustworthiness traits. Indeed, it is well reported in neuroimaging and neuropsychological studies that the amygdala is involved in face expression and traits recognition, but questions concerning whether its role is specific for type of stimuli, in particular threatening and negative valence faces, remained unsolved; alternatively, the amygdala could be involved in the encoding of more general socially relevant cues (Adolphs et al., 2002; Cristinzio et al., 2007; Sergerie et al., 2008). Results of Study 1 revealed that the amygdala was highly responsive to fearful faces but was also activated by angry, happy and disgusted expressions. In Study 2 the amygdala showed similar U-shaped activations for faces morphed along trustworthiness and gender dimensions with increased response at the extremes of both variables. These results support the hypothesis that the amygdala mediates social evaluation and emotional appraising of faces in order to detect stimuli which could be

relevant to guide behaviour in the social environment (Sander et al., 2003). In this view amygdala responded to fearful faces because they can be signal of danger, as other negative expressions like anger and disgust, but also happy faces elicited amygdala activation since they convey important cues to adaptive interactions (Canli et al., 2002). U-shaped responses to trustworthiness and gender suggested that the amygdala was also sensitive to face distinctiveness. Even though this feature does not refer to a specific category as basic emotion expressions, it is an important characteristic in classifying faces (Valentine, 1991) and it represents a simple cue to encode face information (Lee et al., 2000; Leopold et al., 2006). Therefore, face distinctiveness can be considered a socially relevant facial features as expressions and other face dimensions and this can account for the U-shaped amygdala activation (Said et al., 2010).

Study 2 also showed that posterior face-selective regions (OFA, FFA, and STS) responded to trustworthiness and gender dimensions with the same U-shaped patterns revealed in the amygdala suggesting a similar sensitivity to face distinctiveness. This highlighted the distributed nature of the face perception system showing that activity in the face-selective regions in the occipital and temporal lobes can be modulated by variation along different face dimensions. One possible interpretation is that these regions are not only involved in the perceptual analysis of faces but they also contribute to social and emotional evaluation of faces (Allison et al., 2000; Surgulatze et al., 2003). Results of Study 2 do not allow determining whether the U-shaped activation in posterior face-selective regions depends on feed-forward mechanisms in the core system or top-down influence from the extended system. However, these data are in line with the hypothesis that the analysis of the multiplicity of facial features is mediated by the synchronized activity of different regions included in the face processing network, rather than an independent coding of distinct characteristics supported by dissociable areas (Calder and Young, 2005; Ishai, 2008). Future

studies could clarify interconnections between these areas and whether the response pattern in posterior regions is stimuli driven or depends on top-down influence from other areas (Baseler et al., 2012; Vuilleumier and Driver, 2007).

To further examine this issue, in Study 3 and Study 4 I used TMS and a combined TMS-EEG experiment to investigate the role of different cortical areas part of the extended system and analyse possible interactions between regions. Results of Study 3 shed light on the debate concerning the representation of different emotions in the cerebral cortex, showing that mPFC is involved in the semantic knowledge of emotions and contains distinct representations for angry and happy expressions. This confirms that regions part of the extended system involved in processing emotional information and control behaviour also contribute to distinguish specific facial expressions (Haxby et al., 2002).

Finally, by means of TMS-EEG Study 4 revealed interesting evidence about top-down influence in the face processing network and modulation of cortical responsiveness depending on the type of task. In particular, results demonstrated that the activation of the mPFC can causally interfere with the activity generated by the occipital cortex at an early stage of face processing; moreover, tasks requiring discrimination of expressions rather than identity of faces differentially modulated the neural transmission from mPFC towards the right and left temporal region. These final results confirm that face perception relies on a complex cortical network (Ishai, 2008). Within this network prefrontal regions have a critical role in mediating emotional evaluation (Davidson and Irwin, 1999) and likely participate in different stages of face processing throughout the modulation of activity in other structures of the circuit (Nomura et al., 2004; Summerfield et al., 2006).

Therefore, the neural system proposed by Haxby et al. (2000) remains a binding model for research on face perception; however, the results presented above and most recent studies point to the presence of interactive mechanisms among multiple areas rather than a

hierarchical information processing throughout regions with functional selective roles (Ishai, 2008). In particular, connections between different regions seem to be relevant from an early stage of face perception both within the core system and between the core and the extended system (Beseler et al., 2012; Fairhall and Ishai, 2007).

In summary, this thesis examined the neural correlates of face expressions processing considering the role of the amygdala, the face-selective regions in the occipital and temporal lobes, the somatosensory and the prefrontal cortices. Taken together, the results from the four studies show that posterior areas together with the amygdala and prefrontal cortex contribute in creating a complex representation of faces, which comprises information about the emotional expression and social cues useful for human interaction. The amygdala proved to act as a relevant stimuli detector coding different facial expressions and variations in face distinctiveness. Posterior face-selective regions were modulated as the amygdala by variations along different facial dimensions. Furthermore, the medial prefrontal cortex is involved in the semantic encoding of different emotions. In this streaming of information it is possible to hypothesise that both feed-forward and top-down mechanisms interact; anyhow, the interconnections and the functional coupling between areas are crucial to provide the final representation of the face stimuli.

Acknowledgements

My first thanks is for my supervisor Costanza Papagno, she initiated me at research and supported me in the four years of PhD with her great expertise and passion for this work.

I would like to thank Andy Young, who supervised me during my visiting at the University of York, Tim Andrews and Aziz Asghar, who shared with me their fine competence and gave me the opportunity to collaborate in their research group.

I am also grateful to the researchers who collaborated in the different projects part of my thesis, in particular Leonor Romero, Mario Rosanova and Zaira Cattaneo: thank you for the precious assistance and all the nice hours spent together in lab.

Finally, thanks to all the students at the University of Milano-Bicocca and University of York who participated at the different stages of the experiments.

References

Adolphs, R. (2002). Neural system for recognizing emotion. *Current Opinion in Neurobiology,* 12, 169-177.

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Neurosciences*, 3 (12), 469-479.

Adolphs, R., Baron-Cohen, S., Tranel, D. (2002). Impaired recognition of social emotion following amygdala damage. *Journal of Cognitive Neuroscience*, 14, 1264-1274.

Adolphs, R., Damasio, H., Tranel, D., Cooper, G., Damasio, A. R. (2000). A role for somatosensory cortices in the visual recognition of emotion as revealed by three-dimensional lesion mapping. *The Journal of Neuroscience*, 20 (7), 2683-2690.

Adolphs, R., Tranel D., Hamann, S., Young, A. W., Calder, A. J., Phelps, E. A., Anderson, A., Lee, G. P., Damasio, A. R. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia,* 37, 1111-1117.

Adolphs, R., Tranel, D., Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, 393, 470-474.

Adolphs, R., Damasio, H., Tranel, D., Damasio, A. R. (1996). Cortical system for the recognition of emotion in facial expression. *The Journal of Neuroscience*, 16 (23), 7678-7687.

Adolphs, R., Tranel, D., Damasio, H., Damasio, A. R. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, 372, 669-672.

Aggleton, J.P., Young, A.W. (1999). The enigma of the amygdala: on its contribution to human emotion. In L. Nadel, R. D. Lane (Eds.), *Emotion and cognitive neuroscience* (pp. 106-128). New York: Oxford University Press.

Allison, T., Puce, A., McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences,* 4 (7), 267-278.

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D. E., Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, 6, 196-202.

Anderson, A. K., Spencer, D. D., Fulbright, R. K., Phelps, E. A. (2000). Contribution of the anteromedial temporal lobes to the evaluation of facial emotion. *Neuropsychology,* 14, 526-536.

Andrews, T. J., Davies-Thompson, J., Kingstone, A., Young, A. W. (2010). Internal and external features of the face are represented holistically in face-selective regions of visual cortex. *The Journal of Neuroscience*, 30 (9), 3544-3552.

Andrews, T. J., Ewbank, M. P. (2004). Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. *NeuroImage,* 23, 905-913.

Ashley, V., Vuilleumier, P., Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *NeuroReport*, 15, 211-216.

Bar, M., Neta, M., Linz, H. (2006). Very first impressions. *Emotion*, 6, 269-278.

Barbeau, E. J., Taylor, M. J., Regis, J., Marquis, P., Chauvel, P., Ligeois-Chauvel, C. (2008). Spatio temporal dynamic of face recognition. *Cerebral Cortex,* 18, 997-1009.

Baseler, H. A., Harris, R. J., Young, A. W., Andrews, T. J. (2012). Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. *Cerebral Cortex,* in Press.

Beckmann, C., Jenkinson, M., Smith, S.M. (2003). General multi-level linear modelling for group analysis in FMRI. *NeuroImage,* 20, 1052-1063.

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.

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain,* 122, 883-893.

Breiter, H. C., Etcoff, N. L., Whalen, P. J., Kennedy, W. A., Rauch, S. L., Buckner, R. L., Strauss, M. M., Hyman, S. E., Rosen, B. R. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron,* 17, 875-887.

Broks, P., Young, A. W., Maratos, E. J., Coffey, P. J., Calder, A. J., Isaac, C. L., Mayes, A. R., Hodges, J. R., Montaldi, D., Cezayirli, E., Roberts N., Hadley, D. (1998). Face processing impairments after encephalitis: amygdala damage and recognition of fear. *Neuropsychologia*, 36 (1), 59-70.

Bruce, V., Young, A. (2012). *Face perception.* Hove, East Sussex: Psychology Press.

Bruce, V., Young, A. (1986). Understanding face recognition. *British Journal of Psychology,* 77, 305-327.

Calder, A. J., Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience,* 6, 641-651.

Calder, A. J., Lawrence, A. D., Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Review Neuroscience,* 2, 352-362.

Calder, A. J., Young, A.W., Rowland, D., Perrett, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society (B),* 264, 919-925.

Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., Etcoff, N. L. (1996). Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. *Cognitive Neuropsychology,* 13, 699-745.

Campana, G., Cowey, A., Walsh, V. (2002). Priming of motion direction and area V5/MT: a test of perceptual memory. *Cerebral Cortex*, 12, 663-669.

Campbell, R., Brooks, B., de Haan, E., Roberts, T. (1996). Dissociating face processing skills: decisions about lip-read speech, expression, and identity. *The Quarterly Journal of Experimental Psychology,* 49, 259-314.

Canli, T., Sivers, H., Whitfield, S. L., Gotlib, H., Gabrieli, J. D. E. (2002). Amygdala response to happy faces as a function of extroversion. *Science,* 296, 2191.

Carroll, N. C., Young, A. W. (2005). Priming of emotion recognition. *The Quarterly Journal of Experimental Psychology*, 58 (7), 1173-1197.

Cattaneo, Z., Devlin, J. T., Salvini, F., Vecchi, T., Silvanto J. (2010a).The causal role of category-specific neuronal representation in the left ventral premotor cortex (PMv) in semantic processing. *NeuroImage*, 49, 2728-2734.

Cattaneo, L., Sandrini. M., Schwarzbach, J. (2010b). State-dependent TMS reveals a hierarchical representation of observed acts in the temporal, parietal and premotor cortices. *Cerebral Cortex*, 20 (9), 2252-2258.

Cattaneo, Z., Rota, F., Walsh, V., Vecchi, T., Silvanto, J. (2009). TMS-adaptation reveals abstract letter selectivity in the left posterior parietal cortex. *Cerebral Cortex*, 19, 2321-2325.

Cattaneo, Z., Rota, F., Vecchi, T. and Silvanto, J. (2008). Using state-dependency of transcranial magnetic stimulation (TMS) to investigate letter selectivity in the left posterior parietal cortex: a comparison of TMS-priming and TMS-adaptation paradigms. *European Journal of Neuroscience,* 28, 1924-1929.

Cohen Kadosh, R., Muggleton, N., Silvanto, J., Walsh, V. (2010). Double dissociation of format-dependent and number-specific neurons in human parietal cortex. *Cerebral Cortex*, 20, 2166-2171.

Cristinzio, C., Sander, D., Vuilleumier, P. (2007). Recognition of emotional face expressions and amygdala pathology. *Epileptologie*, 24, 130-138.

Cunningham, W. A., Raye, C. L., Johnson, M. K. (2004). Implicit and explicit evaluation: fMRI correlates of valence, emotional intensity, and control in processing of attitudes. *Journal of Cognitive Neuroscience*, 16, 1-13.

Dalgleish, T. (2004). The emotional brain. *Nature Review Neuroscience,* 5, 582-589.

Damasio, A. R, (1994). *Descartes' error: emotion, reason, and the human brain.* Grosset/Putnam.

Davidson, R. J., Irwin, W. (1999). The functional neuroanatomy of emotion and affective style. *Trends in Cognitive Sciences*, 3 (1), 11-21.

Davies-Thomson, J., Andrews, T. J. (2012) Intra- and inter-hemispheric connectivity between face-selective regions in the human brain. *Journal of Neurophysiology*, in Press.

De Houwer, J., Hermans, D., Rothermund, K., Wentura, D. (2002). Affective priming of semantic categorization responses. *Cognition and Emotion*, 16 (5), 643-666.

Dolan, R. J., Fletcher, P., Morris, J., Kapur, N., Deakin, J. F. W., Frith, C. D. (1996). Neural activation during covert processing of positive emotional facial expressions. *NeuroImage*, 4, 194-200.

Drevets, W. C., Burton, H., Videen, T. O., Snyder, A. Z., Simpson, J. R., Raichle, M. E. (2005). Blood flow changes in human somatosensory cortex during anticipated stimulation. *Nature*, 373 (19), 249-252.

Edwards, K. (1998). The face of time: temporal cues in facial expression of emotion. *Psychological Science,* 9, 270-276.

Eimer, M., Holmes, A. (2007). Event-related brain potential correlates of emotional face processing. *Neuropsychologia,* 45, 15-31.

Eimer, M., Holmes, M. (2002). An ERP study on the timecourse of emotional face processing. *Neuroreport*, 13, 427-431.

Ekman, P., Rosenberg, E. (1997). *What the face reveals.* New York: Oxford University Press.

Ekman, P., Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, California: Consulting Psychologists Press.

Engell, A. D., Haxby, J. V., Todorov, A. (2007). Implicit trustworthiness decision: automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, 19, 1508-1519.

Ewbank, M. P., Barnard, P. J., Croucher, C. J., Ramponi, C., Calder, A. J. (2009). The amygdala response to images with impact. *Scan,* 4, 127-133.

Fairhall, S. L., Ishai, A. (2007). Effective connectivity within the distributed cortical network for face perception. *Cerebral Cortex,* 17, 2400-2406.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: an overview. *Cognition and Emotion*, 15 (2), 115-141.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50 (2), 229-238.

Feinstein, J. S., Adolphs, R., Damasio, A., Tranel, D. (2010). The human amygdala and the induction and experience of fear. *Current Biology,* 21, 34-38.

Ferrarelli, F., Massimini, M., Sarasso, S., Casali, A., Riedner, B. A., Angelini, G., Tononi, G., Pearce, R. A. (2010). Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. *PNAS,* 107, 2681-2686.

Fiske, S. T., Cuddy, A. J. C., Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Science,* 11, 77-83.

Fitzgerald, D. A., Angstadt, M., Jelsone, L. M., Nathan, P. J., Phan, K. L. (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. *NeuroImage,* 30, 1441-1448.

Friston, K. J., Harrison, L., Penny, W. (2003). Dynamic causal modelling. *NeuroImage,* 19, 1273-1302.

Galli, G., Feurra, M., Viggiano, M. P. (2006). "Did you see him in the newspaper?" Electrophysiological correlates of context and valence in face processing. *Brain Research,* 1119, 190-202.

Gläsher, J., Adolphs, R. (2003). Processing of the arousal subliminal and supraliminal emotional stimuli by the human amygdala. *The Journal of Neuroscience*, 23 (32), 10274-10282.

Grill-Spector, K., Knouf, N., Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience,* 7, 555-562.

Gschwind, M., Pourtois, G., Schwarts, S., Van De Ville, D., Vuilleumier, P. (2012). White-matter connectivity between face-responsiveness regions in the human brain. *Cerebral Cortex*, 22, 1564-1576.

Hariri, A. R., Bookheimer, S. Y., Mazziotta, J. C. (2000). Modulating emotional responses: effects of a neocortical network on the limbic system. *NeuroReport*, 11, 43-48.

Harmer, C. J., Thilo, K. V., Rothwell, J. C., Goodwin, G. M. (2001). Transcranial Magnetic Stimulation of medial prefrontal cortex impairs processing of angry facial expression. *Nature Neuroscience*, 4 (1), 17-18.

Haxby, J. V., Hoffman, E. A., Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry,* 51, 59-67.

Haxby, J. V., Hoffman, E. A., Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4 (6), 223-233.

Heberlein, A. S., Padon, A. A., Gillihan, S. J., Farah, M. J., Fellows, L. K. (2008). Ventromedial frontal lobe plays a critical role in facial emotion recognition. *Journal of Cognitive Neuroscience*, 20 (4), 721-733.

Hornak, J., Bramham, J., Rolls, E. T., Morris, R. G., O'Doherty, J., Bullock, P. R., Polkey, C. E. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulated cortices. *Brain,* 126, 1691-1712.

Hornak, J., Rolls, E. T., Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioral changes following ventral frontal lobe damage. *Neuropsychologia*, 34 (4), 247-261.

Hussey, E., Safford, A. (2009). Perception of facial expression in somatosensory cortex support simulationist models. *The Journal of Neuroscience*, 29 (2), 301-302.

Igarashi, Y., Kitagawa, N., Spence, C., Ichihara, S. (2007). Assessing the influence of schematic drawings of body parts on tactile discrimination performance using the crossmodal congruency task. *Acta Psychologica*, 124, 190-208.

Ishai, A. (2008). Let's face it: it's a cortical network. *NeuroImage,* 40, 415-419.

Ishai, A., Schmidt, C. F., Boesinger, P. (2005). Face perception is mediated by a distributed cortical network. *Brain Research Bulletin,* 67, 87-93.

Jenkinson, M., Bannister, P., Brady M., Smith, S. (2002). Improved optimisation for the robust and accurate linear registration and motion correction of brain images. *NeuroImage,* 17 (2), 825-841.

Jenkinson, M., Smith, S. M. (2001). A Global Optimisation Method for Robust Affine Registration of Brain Images. *Medical Image Analysis,* 5, 143-156.

Johnson, J. S., Kundu, B., Casali, A. G., Postle, B. R. (2012). Task-dependent changes in cortical excitability and effective connectivity: a combined TMS-EEG study. *Journal of Neurophysiology,* 107, 2383-2392.

Kesler/West, M. L., Andersen, A. H., Smith, C. D., Avison, M. J., Davis, C. E., Kryscio, R. J., Blonder, L. X. (2001). Neural substrates of facial emotion processing using fMRI. *Cognitive Brain Research*, 11, 213-226.

Kirouac, G., Doré, F. Y. (1984). Judgement of facial expression of emotion as a function of exposure time. *Perceptual and Motor Skills,* 59,147–150.

Kiss, M., Driver, J., Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science*, 20 (2), 245-251.

Korhonen, R. J., Hernandez-Pavon, J. C., Metsomaa, J., Mäki. H., Ilmoniemi, r. J., and Sarvas, J. (2011). Removal of large muscle artifacts from transcarnial magnetic stimulation-evoked EEG by independent component analysis. *Medical and Biological Engineering and Computing,* 49, 397-407.

Lane, R. D., Reiman, E. M., Bradley, M. M., Lang, P. J., Ahern, G. L., Davidson, R. J., Schwartz, G. E. (1997). Neuroanatomical correlates of pleasant and unpleasant emotion. *Neuropsychologia*, 35 (11), 1437-1444.

Lee, K., Byatt, G., Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: testing the face-space framework. *Psychological Science,* 11(5), 379-385.

Leopold, D. A., Bondar, I. V., Giese, M. A. (2006). Norm-based face encoding by single neurons in the monkey inferotemporal cortex. *Nature,* 442, 572-575.

Leppänen, J. M. (2006). Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. *Current Opinion in Psychiatry,* 19, 34-39.

Liu, X., Banich, M. T., Jacobson, B. L., Tanabe, J. L. (2004). Common and distinct neural substrates of attentional control in an integrated Simon and spatial Stroop task as assessed by event-related fMRI. *NeuroImage*, 22, 1097-1106.

Loffler, G., Yourganov, G., Wilkinson, F., Wilson, H. R. (2005). fMRI evidence for the neural representation of faces. *Nature Neuroscience,* 8, 1386-1390.

Mah, L. W. Y., Arnold, M. C., Grafman, J. (2005). Deficits in social knowledge following damage to ventromedial prefrontal cortex. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 17 (1), 66-74.

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science,* 309, 2228-2232.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology*, 21 (1), 3-23.

Mevorach, C., Humphreys, G. W., Shavel, L. (2006). Opposite biases in salience-based selection for the left and right posterior parietal cortex. *Nature Neuroscience*, 9 (6), 740-742.

Miniussi, C., Thut, G. (2010). Combining TMS and EEG offers new prespects in cognitive neuroscience. *Brain Topography,* 22, 249-256.

Mechelli, A., Price, C. J., Friston, K. J., Ishai, A. (2004). Where bottom-up meets top-down: neural interactions during perception and imagery. *Cerebral Cortex,* 14, 1256-1265.

Morishima, Y., Akaishi, R., Yamada, Y., Okuda, J., Toma, K., Sakai, K. (2009). Task-specific signal transmission from prefrontal cortex in visual selective attention. *Nature Neuroscience,* 12, 85-91.

Morris, J. S., Friston, K. J., Büchel, C., Frith, C. D, Young, A. W., Calder, A. J., Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain,* 121, 47-57.

Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A.J., Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expression. *Nature,* 383, 812-815.

Münte, T. F., Brack, M., Grootheer, O., Wieringa, B. M., Matzke, M., Johannes, S. (1998). Brain potentials reveal the timing of face identity and face expression judgments. *Neuroscinece Research,* 30, 25-34.

Niedenthal, P. M., 2007. Embodying emotion. *Science,* 316, 1002-1005.

Nomura, M., Ohira, H., Haneda, K., Iidaka, T., Sadato, N., Okada, T., Yonekura, Y. (2004). Functional association of the amygdale and ventral prefrontal cortex during cognitive evaluation of facial expressions primed by masked angry faces: an event-related fMRI study. *NeuroImage*, 21, 352-363.

Oberman, L. M., Winkielman, P., Ramachandran, V. S. (2007). Face to face: blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience*, 2, 167-178.

Oosterhof, N. N., Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America,* 105, 11087-11092.

Perrin, F., Pernier, J., Bertrand, O., Friston, K. J. (1998). Spherical splines for scalp potential and current density mapping. Electroencephalography and Clinical *Neurophisiology,* 72, 184-187.

Pessoa, L, Adolphs, R. (2010). Emotion processing and the amygdala: from a 'low road' to 'many roads' of evaluating biological significance. *Nature Review Neuroscience,* 11, 773-782.

Pessoa, L., McKenna, M., Gutierre, E., Ungerleider, L. G. (2002). Neural processing of emotional faces require attention. *PNAS,* 99, 11458-11463.

Phan, K. L., Wager, T., Taylor, S. F., Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage,* 16, 331-348.

Philippi, C. L., Mehta, S., Gabrowski, T., Adolphs, R., Rudrauf, D. (2009). Damage to association fiber tracts impairs recognition of the facial expression of emotion. *The Journal of Neuroscience*, 29 (48), 15089-15099.

Phillips, M. L., Williams, L., Senior, C., Bullmore, E. T., Brammer, M. J., Andrew, C., Williams, S. C. R., David, A. S. (1999). A differential neural response to threatening and non-threatening negative facial expressions in paranoid and non-paranoid schizophrenics. *Psychiatry Research: Neuroimaging Section,* 92, 11-31.

Phillips, M. L., Young, A. L., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., Williams, S. C. R., Bullmore, E. T., Brammer, M., Gray, J. A. (1998a). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of the Royal Society (B),* 265, 1809-1817.

Phillips, M. L., Bullmore, E. T., Howard, R., Woodruff, P. W. R., Wright, I. W., Williams, S. C. R., Simmons, A., Andrew, C., Brammer, M., David, A. S. (1998b). Investigation of facial recognition memory and happy and sad facial expression perception: an fMRI study. *Psychiatry Research: Neuroimaging section*, 83, 127-138.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C. R., Gray, j. A., Davis, A. S. (1997). A specific substrate for perceiving facial expressions of disgust. *Nature*, 389, 495-498.

Picton, T. W., Stuss, D. T., Alexander, M. P., Shallice, T., Binns, M. A., Gillingham, S. (2007). Effects of focal frontal lesions on response inhibition. *Cerebral Cortex*, 17 (4), 826-838.

Pitcher, D., Garrido, L., Walsh, V., Duchaine, C. (2008). Transcranial Magnetic Stimulation disrupts the perception and embodiment of facial expression. *The Journal of Neuroscience*, 28 (36), 8929-8933.

Platek, S. M., Krill, A. L. (2009). Self-face resemblance attenuates other-race face effect in the amygdala. *Brain Research*, 1284, 156-160.

Pourtois, G., Sander, D., Andres, M., Grandjean, D., Reveret, L., Olivier, E., Vuilleumier, P., (2004). Dissociable roles of the human somatosensory and superior temporal cortices for processing social face signal. *European Journal of Neuroscience*, 20, 3507-3515.

Rosanova, M., Casarotto, S., Pigorini, A., Canali, P., Casali, A. G., Massimini, M. (2012). Combining transcranial magnetic stimulation with electroencephalography to study human cortical excitability and effective connectivity. *Neuromethods,* 67, 435-457.

Rosanova, M., Casali, A., Bellina, V., Retsa, F., Mariotti, M., Massimini, M. (2009). Natural frequencies of the human corticothalamic circuits. *The Journal of Neuroscience,* 29, 7679-7685.

Rossion, B., Jacques, C. (2008). Does physical interstimulus variance account for early electrophysiological face sensitive response in the human brain? Ten lessons on the N170. *NeuroImage,* 39, 1959-1979.

Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., Mayer, E. (2003). A network of occipito-temporal face-selective areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain,* 126, 2381-2395.

Sadeh, B., Pitcher, D., Brandman, T., Eisen, A., Thaler, A., Yovel, G. (2011). Stimulation of category-selective brain areas modulates ERP to their preferred categories. *Current Biology,* 21, 1894-1899.

Said, C. P., Haxby, J. V., Todorov, A. (2011). Brain system for assessing the affective value of faces. *Philosophical Transaction of the Royal Society*, 366, 1660-1670.

Said, C., Dotsch, R., Todorov, A. (2010). The amygdala and FFA track both social and non-social face dimension. *Neuropsychologia*, 48, 3596-3605.

Said, C. P., Baron, S. G., Todorov, A. (2008). Nonlinear amygdala response to face trustworthiness: contribution of high and low spatial frequency information. *Journal of Cognitive Neuroscience*, 21 (3), 519-528.

Sander, D., Jordan, G., Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Reviews in the Neuroscience*, 14, 303-316.

Santos, I. M., Young, A. W. (2008). Effect of inversion and negation on social inferences from faces. *Perception,* 37, 1061-1078.

Santos, I. M., Young, A. W. (2005). Exploring the perception of social characteristics in faces using the isolation effect. *Visual Cognition,* 12, 213-247.

Sato, W., Yoshikawa, S., Kochiyama, T., Matsumara, M. (2004). The amygdala processes the emoziona significante of facial expressions: an fMRI investigation using the interaction between expression and face direction. *NeuroImage*, 22, 1006-1013.

Savran, N., Alyüz, H., Dibeklioğlu, O., Çeliktutan, B., Gökberk, B., Sankur, L., Akarun, Bosphorus Database for 3D Face Analysis. The First COST 2101 Workshop on Biometrics and Identity Management (BIOID 2008), Roskilde University, Denmark, May 2008.

Sergerie, K., Chochol, C., Armony, J. L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neuroscience and Behavioural Reviews,* 32, 811-830.

Silvanto, J., Schwarzkopf, D.S., Gilaie-Dotan, S., Rees, G. (2010). Differing causal roles for Lateral Occipital cortex and Occipital Face Area in invariant shape recognition. *European Journal of Neuroscience*, 32 (1), 165-171.

Silvanto, J., Muggleton, N., Walsh, V. (2008). State-dependency in brain stimulation studies of perception and cognition. *Trends in Cognitive Sciences*, 12 (12), 447-454.

Silvanto, J., Pascual-Leone, A. (2008). State-dependency of Transcranial Magnetic Stimulation. *Brain Topography*, 21 (1), 1-10.

Silavanto, J., Muggleton, N. G., Cowey, A., Walsh, V. (2007). Neural adaptation reveals state-dependent effects of transcranial magnetic stimulation. *European Journal of Neuroscience*, 25 (6), 1874-1881.

Small, D. M., Gregory, M. D., Mak, Y. E.,. Gitelman, D., Mesulam, M. M., Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gestation. *Neuron*, 39, 701-711.

Smith, S. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143-155.

Sprengelmeyer, R., Young, A. W., Schroeder, U., Grossenbacher, P. G., Federlein, J., Büttner, T., Przuntek, H. (1999). Knowing no fear. *Proceedings of the Royal Society: Biological Sciences,* 266, 2451-2456.

Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., Hirsh, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science,* 314, 1311-1314.

Surguladze, S. A., Brammer, M. J., Young, A. W., Andrew, C., Travis, M. J., Williams, S. C. R., Phillips, M. L. (2003). A preferential increase in the extrastriate response to signals of danger. *NeuroImage,* 19, 1317-1328.

Taylor, P. J., Walsh, V., Eimer, M. (2008). Combining TMS and EEG to study cognitive function and cortico-cortico interactions. *Behavioural Brain Research,* 191, 141-147.

Tiddeman, B., Burt, D. M., Perrett, D. I. (2001). Computer graphics in facial perception research. *IEEE Computer Graphics and Applications*, 21, 42-50.

Todorov, A. (2008). Evaluating faces on trustworthiness. An extension of system for recognition of emotions signalling approach/avoidance behaviors. *Annals of the New York Academy of Science*, 1124, 208-224.

Todorov, A., Baron, S. G., Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Scan*, 3, 119-127.

Todorov, A., Engell A. D. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Scan*, 3, 303-312.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology Section A*, 43, 161-204.

Virtanen, J., Ruohonen, J., Naatanen, R., Ilmoniemi, R. (1999). Instrumentation for the measurement of electrical brain responses to transcranial magnetic stimulation. *Medical and Biological engineering and computing,* 37, 322-326.

Vuilleumier, P. (2005). How brain beware: neural mechanisms of emotional attention. *Trends in Cognitive Science*, 9, 585-594.

Vuilleumier, P., Driver, J. (2007). Modulation of visual processing by attention and emotion: windows on causal interactions between human brain regions. *Philosophical Transactions of the Royal Society,* 362, 837-855.

Vuilleumier, P., Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*, 45, 174-194.

Vuilleumier, P., Richardson, M. P., Armony, J. L., Driver, J., Dolan, R. J. (2004). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, 7, 1271-1278.

Vytal, K., Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotion: a voxel-based meta-analysis. *Journal of Cognitive Neuroscience,* 22, 2864-2885.

Walsh, V., Cowey, A. (2000). Transcranial magnetic stimulation and cognitive neuroscience. *Nature Neuroscience*, 1, 73-79.

Wang, A. T., Dapretto, M., Hariri, A. R., Sigman, M., Bookheimer, S. Y. (2004). Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry,* 43, 481-490.

Wentura, D. (1999). Activation and inhibition of affective information: evidence for negative priming in the evaluation task. *Cognition and Emotion*, 13 (1), 65-91.

Whalen, P. J., Shin, L. M., McInerney, S. C., Fischer, H., Wright, C. I., Rauch, S. L. (2001). A functional MRI study of human amygdala responses to facial expressions of fear versus anger. *Emotion,* 1, 70-83.

Willis, J., Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17, 592-598.

Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., Dolan, R. J. (2007). Brain system for assessing facial attractiveness. *Neuropsychologia*, 45, 195-206.

Winston, J. S., O'Doherty, J., Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *NeuroImage,* 20, 84-97.

Winston, J. S., Strange, B. A., O'Doherty, J., Dolan, R. J. (2002) Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, 5, 277-283.

Woolrich, M. W. (2008). Robust Group Analysis Using Outlier Inference. *NeuroImage,* 41, 286-301.

Woolrich, M. W., Behrens, T. E. J., Beckmann, C. F., Jenkinson, M., Smith, S. M. (2004). Multi-level linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage,* 21, 1732-1747.

Wronka, E., Walentowska, W. (2011). Attention modulated emotional expression processing. *Psychophysiology,* 48, 1047-1056.

Young, A. W., Perrett, D. I., Calder, A. J., Sprengelmeyer, R., Ekman, P. (2002). *Facial expressions of emotion: Stimuli and test (FEEST).* Bury St. Edmunds: Thames Valley test Company.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., Perrett, D.I. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition,* 63, 271-313.

Young, A. W., Hellawell, D. J., Van De Wal, C., Johnson, M. (1996). Facial expression processing after amygdalotomy. *Neuropsychologia,* 34, 31-39.

Young, A. W., Aggleton, J. P., Hellawell, D. J., Johnson, M., Broks, P., Hanley, J.R. (1995). Face processing impairments after amygdalotomy. *Brain,* 118, 15-24.

Zalla, T., Koechlin, E., Pietrini, P., Basso, G., Aquino, P., Sirigu, A., Grafman, J. (2000). Differential amygdala responses to inning and loosing: a functional magnetic resonance imaging study in humans. *European Journal of Neuroscience*, 12, 1764-1770.