ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 22 (2004) 372–380

# Dissociable amygdala and orbitofrontal responses during reversal fear conditioning

J.S. Morris[a] and R.J. Dolan[b,*]

[a] *Institute of Child Health, London WC1N 1EH, UK*
[b] *Wellcome Department of Cognitive Neurology, London WC1N 3BG, UK*

Received 25 September 2003; revised 17 December 2003; accepted 16 January 2004

**The neural mechanisms underlying the persistence and plasticity of human emotional learning are unknown. Here we describe dissociable neural responses in amygdala and orbitofrontal cortex during acquisition and reversal of discriminatory fear conditioning. During acquisition, increased responses in bilateral amygdala were elicited by a face stimulus (A = CS+) predictive of an aversive noise compared to another nonpredictive face (B = CS−). With subsequent reversal of the conditioning contingency, face B (new CS+) elicited enhanced responses in right orbitofrontal cortex, while face A (old CS+) continued to evoke increased responses in right ventral amygdala. Thus, while orbitofrontal cortex exhibited rapid reversal of acquired fear responses, ventral amygdala showed a persistent, nonreversing "memory" for previous fear-related stimulus associations.**


*Keywords:* Amygdala; Orbitofrontal responses; Fear conditioning

## Introduction

Memories of highly charged emotional events (e.g., one's own wedding day) are typically vivid and long-lasting (Cahill and McGaugh, 1998). Similarly, as exemplified by the adage "once bitten, twice shy", learned emotional responses are rapidly acquired and resistant to modification (LeDoux, 1996). Indeed, it is suggested that fear learning "is particularly resilient, and in fact may represent an indelible form of learning" (LeDoux, 1996). Stability of emotional responding is usually adaptive, facilitating the avoidance of common dangers and the pursuit of habitual rewards. However, if emotional responses are exaggerated or misplaced, resistance to modification can become pathological, as, for example, in phobias. The neural mechanisms mediating "indelibility" of fear learning, therefore, are a likely candidate mechanism for the perpetuation of phobic behaviours. Evidence from both animal and human studies strongly implicates the amygdala in the acquisition and persistence of emotional memories and behaviours (Bechara et al., 1995; Cahill et al., 1995, 1996; LeDoux, 1995, 1996; LeDoux et al., 1988).

Despite a tendency to persist, emotional behaviours are capable of modification under the right conditions. Most phobic patients, for example, can learn to control their avoidant responses (towards spiders, social situations, etc) with appropriate behaviour therapy (Marks, 1987). The ability to modify emotional responses is particularly important for animals with complex social relationships such as primates, where rapid and flexible readjustment of behaviour is likely to be highly adaptive. Studies of both monkeys and humans performing learning tasks involving reversal of reward contingencies strongly implicate orbitofrontal cortex in the modification of reinforced behaviours (Fellows and Farah, 2003; Jones and Mishkin, 1972; Rolls et al., 1996; Thorpe et al., 1983). Monkeys with orbitofrontal lesions continue to respond to formerly rewarded objects on object reversal tasks, rather than switching to currently rewarded stimuli (Jones and Mishkin, 1972). Electrophysiological studies in monkeys have shown that orbitofrontal cells exhibit very rapid response switches (e.g., after a single trial) following reversal of reinforcement contingencies (Rolls et al., 1996; Thorpe et al., 1983). By contrast, many amygdala cells do not show switches of activity following reversal; that is, these amygdala cells continue to respond according to the original reinforcement association (Sanghera et al., 1979). These animal data suggest, therefore, that orbitofrontal cortex and amygdala have distinct roles in emotional learning and behaviour. Whereas emotional responses are rapidly modified by orbitofrontal cortex in accordance with changing external contingencies, amygdala-dependent "memories" of established emotional associations continue to be maintained in the brain.

Studies investigating amygdala involvement in emotional learning have typically used classical (Pavlovian) conditioning paradigms (e.g., Buchel et al., 1998; LaBar et al., 1998; LeDoux et al., 1988; Morris et al., 1998b, 2001a). By contrast, previous studies showing orbitofrontal involvement in reversal learning have employed operant (instrumental) paradigms (e.g., Fellows and Farah, 2003; Jones and Mishkin, 1972; Rolls et al., 1996; Thorpe et al., 1983). Whereas amygdala is implicated in the acquisition of both classically conditioned and operant responses via dissociable neural pathways (Amorapanth et al., 2000), there are no previous data regarding orbitofrontal involvement in the reversal of classi-

* Corresponding author. Wellcome Department of Cognitive Neurology, 12 Queen Square, London WC1N 3BG, UK. Fax: +44-207-8131420.
*E-mail address:* r.dolan@fil.ion.ucl.ac.uk (R.J. Dolan).
**Available online on ScienceDirect (www.sciencedirect.com.)**

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.01.012

cally conditioned responses. In the present study, we decided to investigate acquisition and reversal of classical (Pavlovian) conditioning without an operant (instrumental) component. In addition to allowing direct comparison of acquisition and reversal responses using previous fear learning paradigms (Buchel et al., 1998; Morris et al., 2001a), Pavlovian conditioning avoids confounding responses reflecting stimulus association learning with instrumental response acquisition. Subjects viewed two neutral faces, one of which (A = CS+) was initially paired with an aversive noise on 33% of trials, while the other face (B = CS−) was never paired. In a subsequent reversal session, the conditioning contingency was changed so that face B (old CS−) was paired with the noise to become the new CS+, while face A (old CS+) was no longer paired with the noise to become the new CS−. The specific aim was to determine the nature of human amygdala and orbitofrontal cortex responses to conditioned stimuli during acquisition and reversal of Pavlovian fear learning.

## Materials and methods

### *Subjects*

Twelve right-handed volunteer subjects were recruited by advertisement. None of the subjects had any history of neurological or psychiatric disorder and all were medication-free at the time of experiment. All subjects gave informed consent to the study that was approved by the Joint Ethics Committee of the National Hospital for Neurology and Neurosurgery and Institute of Neurology.

### *Experimental design*

During scanning, subjects viewed greyscale pictures of two faces (one male, one female, both with neutral expressions) shown singly in a pseudorandomized order. The interstimulus interval was 6 s. Each picture was projected for 500 ms onto a screen above the head volume coil in the scanner. Subjects were required to indicate, via a button press, the sex of each presented face. Every scanning session was divided into four seamless phases: (1) acquisition of conditioning, (2) repeat conditioning, (3) reversal conditioning, and (4) new conditioning. Subjects were not informed about the separate phases in the experiment. Thirty-six presentations of both faces (i.e., total of 36 × 2 = 72 presentations) were made in each of the four phases. During acquisition of conditioning (1), one of the faces (A = CS+) was immediately followed by a 500 ms 100 dB white noise burst (UCS) on 12 of the 36 presentations (in a pseudorandomized order). The other face (B = CS−) was never followed by the noise. In six subjects, the male face was used as the CS+, in the other six subjects the female face was the CS+. In the repeat conditioning phase (2), the same stimulus contingency applied as in the previous phase (i.e., face A was followed by noise on 12 presentations). In the reversal conditioning phase (3), face B (the old CS−) was immediately followed by the noise on 12 presentations, while face A (old CS+) was never paired with the noise. Finally, in the new conditioning phase (4), two new faces (one male, one female, both with neutral expressions) were presented, one of which (C = CS+) was paired with the noise on 12 occasions, while the other (D = CS−) was never paired. The order of phases 2–4 was counterbalanced across the 12 subjects. This experimental design ensured that unpaired face A and face B conditions in phases 2 and 3 were physically and temporally identical, differing only in the conditioning contingency concurrent during each phase; that is, on paired trials. It should be noted that the counterbalanced design created an order effect such that the reversal session occurred before the repeat session in an "early" subgroup of subjects, but after the repeat session in a "late" subgroup. In the "early" reversal subgroup, the conditioning contingency in the subsequent repeat

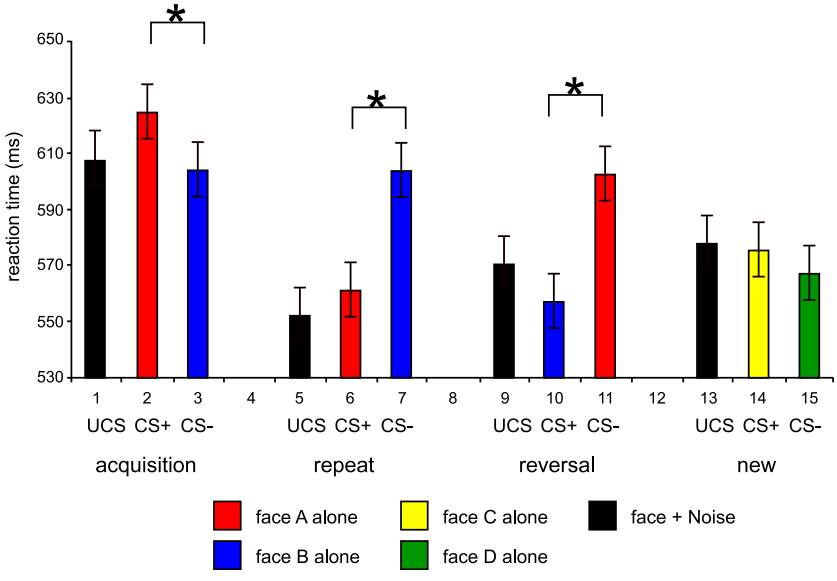


Fig. 1. Reaction times. A graphical display of mean response latencies in each stimulus condition in all three phases. Stimulus contingencies in each phase are indicated by the labels "UCS", "CS+", and "CS−". Stimulus identities, i.e., face A, face B, etc., which remain constant across different phases are indicated by the colour of the histogram columns. Bars on the histogram represent 2 standard errors. Significant ($P < 0.05$) pairwise differences between the mean reaction times for two conditions are indicated by asterisks (*).

session was the same as in the preceding reversal session; that is, no "double" reversal of contingencies occurred. The new conditioning phase (4) provided an additional control for the specificity of changes observed in the reversal phase; that is, to exclude differences in neural response resulting simply from nonspecific novelty of the reversal phase relative to the repeat phase.

*Data acquisition*

Neuroimaging data were acquired with a 2-T Magnetom VISION whole-body MRI system equipped with a head volume coil. Contiguous multislice T2*-weighted echoplanar images were obtained using a sequence that enhanced blood oxygenation level dependent (BOLD) contrast (FOV = 192 mm, TE = 40 ms, flip angle = 90°).Volumes covering the whole brain (48 slices, slice thickness 2 mm) were obtained every 4.1 s (TR). A T1-weighted anatomical MRI was also acquired for each subject. BOLD signal dropout was noted at sinus cavity susceptibility boundaries resulting in reduced coverage of orbitofrontal cortex, one of the regions of interest. The spatial extent of usable orbitofrontal signal ranged approximately from $x = 0$, $y = 23$, $z = -23$ posteriorly to $x = 50$, $y = 50$, $z = -18$ laterally to $x = 0$, $y = 62$, $z = -23$ anteriorly. All reported orbitofrontal activations lie within these boundaries. We are unable to determine from our present neuroimaging data whether additional orbitofrontal activations occurred in regions affected by signal dropout. Reaction times in the sex decision task were recorded for each presented face and used as an index of conditioning (Critchley et al., 2002; Gottfried et al., 2002; McIntosh et al., 2003). Changes in skin conductance were also recorded, but unfortunately a scanner-related artefact in these measurements made them unusable.

*Data analysis*

The fMRI data were analysed using statistical parametric mapping (SPM99) (Friston et al., 1995; see also http://www.fil.ion.ucl.ac.uk/spm). Following realignment to the first volume, the functional (T2*-weighted) scans were spatially normalized to a standard template. The structural (T1-weighted) MRIs were coregistered to the functional scans and transformed into the same standard space. The functional data were smoothed using a

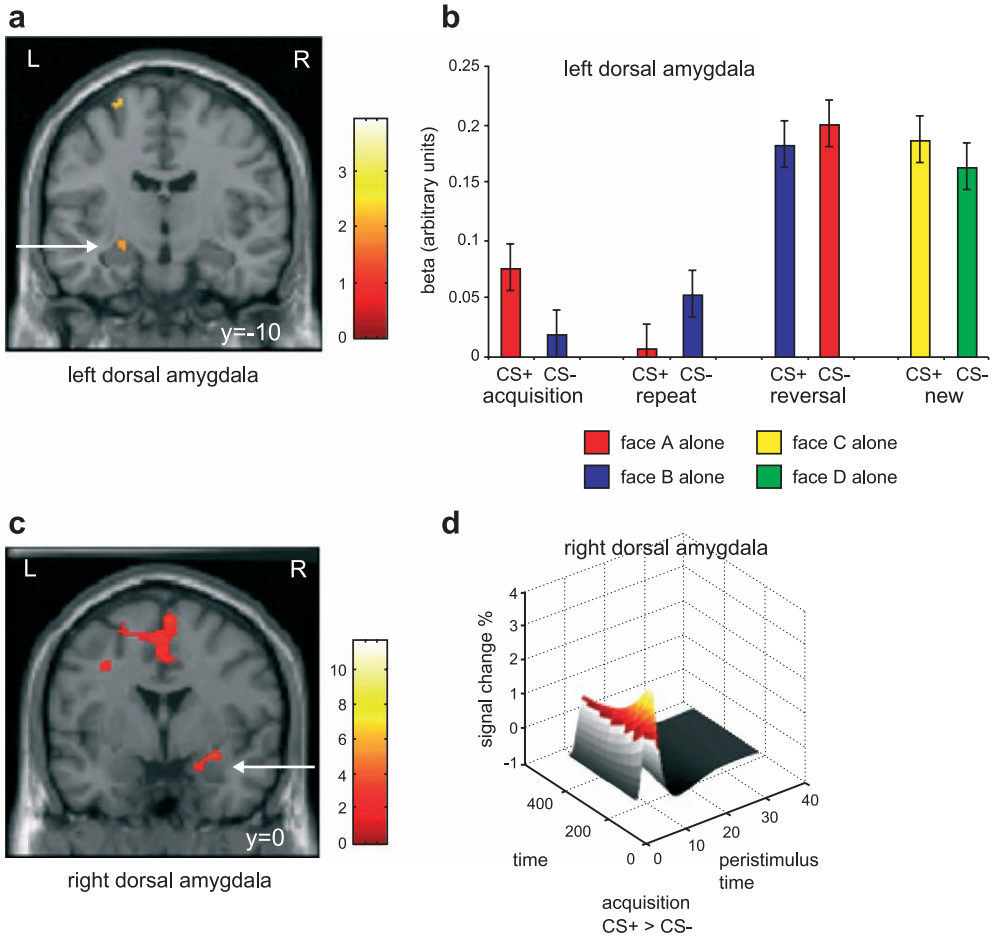


Fig. 2. Acquisition CS+ responses. (a) Statistical parametric map (SPM) showing dorsal region of left amygdala (indicated by arrow) with greater mean response to face A (CS+) than face B (CS−) in the acquisition phase. (b) Graphical representation of parameter estimates for response in the maximal left amygdala voxel in a. (c) Statistical parametric map (SPM) showing dorsal region of right amygdala (indicated by arrow) where responses to face A (CS+) exhibit significant time-dependent decreases in amplitude [relative to face B (CS−)] in the acquisition phase. In a and c, activations are displayed on coronal sections of a canonical MRI. (d) 3-D graphical representation of the event-related activity in the maximal right amygdala voxel in c.

6 mm (full width at half maximum) isotropic Gaussian kernel to allow for corrected statistical inference. The evoked haemodynamic responses for the different stimulus events were modelled as delta functions convolved with a synthetic haemodynamic response function and its temporal derivative (Josephs et al., 1997).

Specific effects (e.g., acquisition CS+ > acquisition CS−) were tested by applying linear contrasts to the parameter estimates for each event. Contrasts were specified for individual subjects, and the resulting contrast images were entered into a second-level random effects analysis to produce group results. The resulting $t$ statistic at every voxel constitutes an SPM. A time by event interaction analysis was performed by multiplying the regressors for the stimulus events with a mean corrected exponential function having a time constant one quarter of the session length. The value of this time constant was chosen from previous time-dependent analyses of fear-conditioned amygdala responses (Buchel et al., 1998; Morris et al., 2001a). Contrasts tested for the difference of the interaction terms between the acquisition CS+ and acquisition CS− conditions. Reported $P$ values were corrected for the number of comparisons made within each a priori region of interest (Worsley et al., 1996). Maxima in amygdala [$x = 18$, $y = -2$, $z = -28$, radius of volume of interest (VOI) = 8 mm], medial and lateral orbitofrontal cortex ($x = 12$, $y = 50$, $z = -12$ and $x = 42$, $y = 42$, $z = -6$, VOI radius = 20 mm), fusiform gyrus ($x = 36$, $y = -62$, $z = -20$, VOI radius = 20 mm), anterior cingulate ($x = 3$, $y = -4$, $z = 40$, VOI radius = 20 mm), insula ($x = 48$, $y = 12$, $z = -6$, VOI radius = 20 mm) and posterior thalamus (18, −28, 12, VOI radius = 20 mm) were used to define search volumes from previously reported activations in studies of face processing and fear conditioning (Blair et al., 1999; LaBar et al., 1998; Morris et al., 1997, 1998a,b, 2001a). All activations in these regions with a significance level of $P < 0.05$ (corrected) are reported.

## Results

### *Behavioural data*

Reaction times for making a sex decision were significantly slower ($P < 0.01$) for the CS+ face than the CS− face in the acquisition phase (Fig. 1). In the subsequent repeat and reversal phases, reaction times (RTs) were significantly faster ($P < 0.001$) for the CS+ face (Fig. 1). In the new phase, RTs to CS+ and CS− faces did not differ significantly, although there was trend for RTs to the CS+ to be slower (Fig. 1). It is notable that RTs in all conditions were faster in the new phase than the initial acquisition phase, suggestive of a nonspecific practice effect (Fig. 1). There were no significant differences in RTs between subjects with "early" and "late" reversal sessions (early reversal mean RTs: UCS = 547.6 ms, CS+ = 551.4 ms, CS− = 598.1 ms; late reversal mean RTs: UCS = 581.6, CS+ = 560.1 ms, CS− = 597.2 ms).

### *Neuroimaging data*

#### *Acquisition phase*

We compared event-related neural responses evoked by the CS+ and CS− faces during the acquisition phase. A dorsal region of left amygdala showed a greater mean response to the CS+ face (Figs. 2a,b; Table 1a). Other regions showing a significantly greater mean response to the CS+ face included right insula and bilateral posterior thalamus (Table 1a). Previous studies (Buchel et al., 1998, LaBar et al., 1998) have shown the importance of time-dependent, learning-related responses in amygdala, and therefore we also tested for time × condition interactions during the acquisition phase. This analysis revealed a dorsal region of right amygdala in which responses to the CS+ face were initially increased but then rapidly declined during the remainder of the phase compared to the CS− face (Figs. 2c,d; Table 1b).

#### *Repeat and reversal phases*

To identify reversal of learning-related differential responses, we compared event-related neural responses evoked by CS+ and CS− faces in repeat and reversal phases. Right ventromedial prefrontal and lateral orbitofrontal cortex showed increased CS+ responses in both repeat and reversal phases; that is, face A (CS+) elicited increased responses in the repeat phase compared to B (CS−), while face B (new CS+) evoked increased responses in the reversal phase compared to A (old CS+) (Figs. 3a,b; Table 1c). Regions exhibiting similar reversal of learning-related responses included right medial amygdala, right anterior cingulate, and bilateral posterior fusiform gyrus (Fig. 3; Table 1c). By contrast, several regions with increased responses to face A (CS+) in the repeat phase continued to show greater responses to face A (old CS+) than face B (new CS+) in the reversal phase. Regions showing this nonreversing pattern of learning-related response included right ventral amygdala, right anterior cingulate (lateral to the "reversing" cingulate region), right anterior fusiform gyrus, and right posterior thalamus (Fig. 4; Table 1d). There were no significant differences between "ear-

Table 1
Brain regions activated during reversal conditioning

| Brain area | Coordinates ($x$, $y$, $z$) | $P$ value |
|---|---|---|
| *(a) Acquisition CS+ > acquisition CS−* | | |
| Left dorsal amygdala | −24, −10, −12 | <0.05 |
| Right insula | 42, 30, 0 | <0.05 |
| Right posterior thalamus | 34, −28, 0 | <0.01 |
| Left posterior thalamus | −26, −22, 8 | <0.01 |
| *(b) (Acquisition CS+ > acquisition CS−)* × time interaction | | |
| Right dorsal amygdala | 30, 0, −12 | <0.05 |
| Right anterior cingulate | 0, 4, 58 | <0.05 |
| *(c) (Repeat CS+ and reversal CS+) > (repeat CS− and reversal CS−)* | | |
| Right medial orbitofrontal cortex | 4, 38, −14 | <0.05 |
| Right lateral orbitofrontal cortex | 42, 44, −16 | <0.05 |
| Right medial amygdala | 18, 2, −24 | <0.05 |
| Right anterior cingulate | 12, 4, 56 | <0.05 |
| Right posterior fusiform gyrus | 28, −74, −8 | <0.05 |
| Left posterior fusiform gyrus | −34, −72, −14 | <0.05 |
| *(d) (Repeat CS+ and reversal CS−) > (repeat CS− and reversal CS+)* | | |
| Right ventral amygdala | 30, 2, −26 | <0.01 |
| Right anterior cingulate | 18, 0, 40 | <0.05 |
| Right anterior fusiform gyrus | 28, −50, −12 | <0.05 |
| Right posterior thalamus | 20, −26, 10 | <0.05 |
| Right insula | 40, 16, −12 | <0.05 |
| Left insula | 38, −6, −6 | <0.05 |

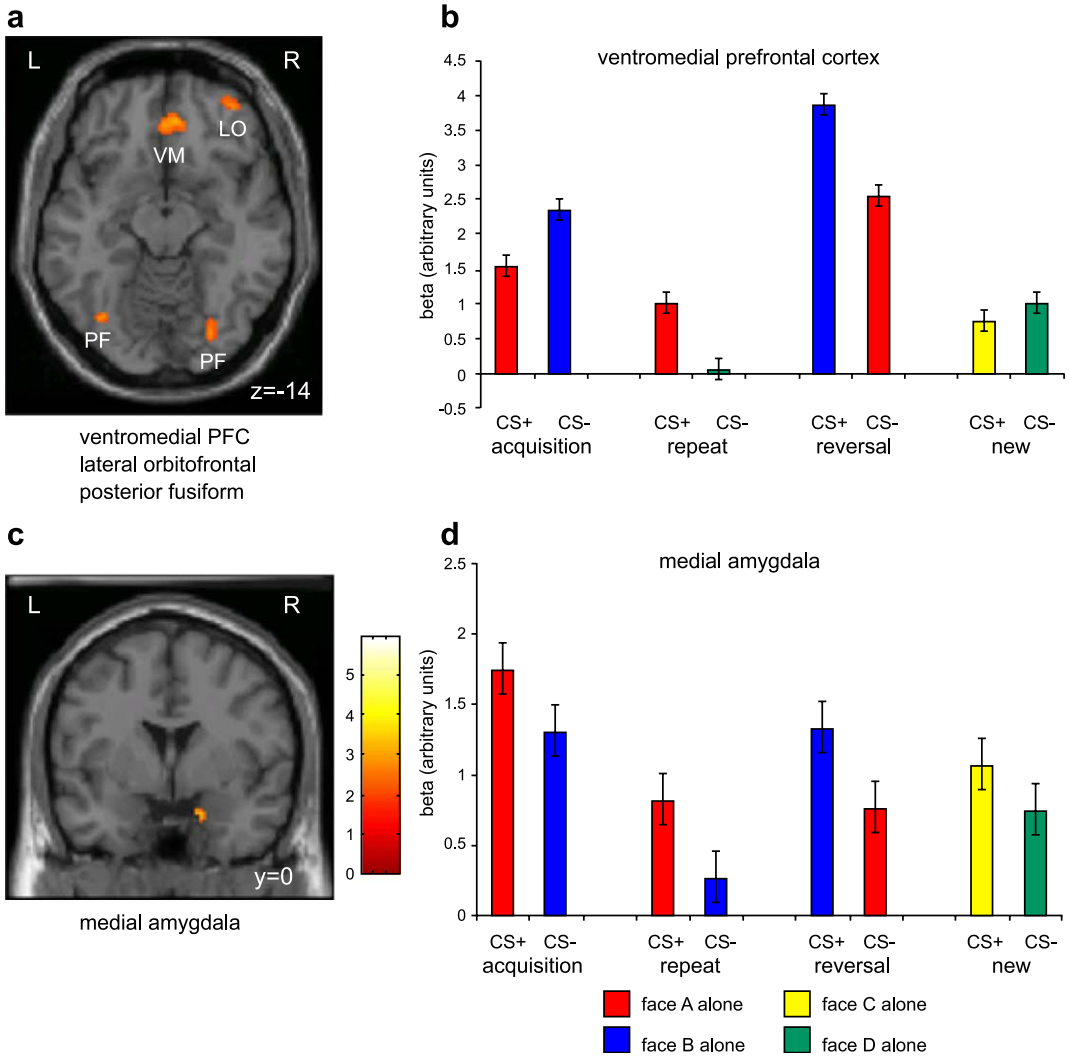


Fig. 3. Reversing CS+ responses. Statistical parametric maps (SPMs) showing (a) regions of right ventromedial prefrontal cortex (VM), right lateral orbitofrontal cortex (LO), bilateral posterior fusiform gyrus (PF) and (c) right medial amygdala with increased responses to the CS+ face relative to the CS − face in both the repeat and reversal phases. Note that the identity of the CS+ face changes in the reversal phase. Activations are displayed on (a) transverse and (c) coronal sections of a canonical MRI. Graphical representations of parameter estimates for responses to the CS+ face are shown for (b) right ventromedial prefrontal cortex and (d) right medial amygdala.

ly" and "late" reversal subgroups in learning-related neural responses reported in Table 1.

## Discussion

The present results provide evidence for rapid, experience-dependent modification of learned emotional responses in human orbitofrontal cortex and medial amygdala. During a single phase, lasting a few minutes, orbitofrontal cortex and a region of medial amygdala switched responses to reflect changed stimulus contingencies. By contrast, the data also indicate a persistent, nonreversing, "memory" for biologically significant stimulus associations in ventral amygdala. Differential responses to (old) CS+ faces in ventral amygdala persisted despite reversal of the conditioning contingency, suggesting that this structure is representing a more long-lasting association between conditioned stimulus and UCS. Other brain structures implicated in emotional processing (e.g., anterior cingulate and insula) also showed intraregional functional segregation of learning-related responses (Table 1a–d).

Dorsal amygdala showed increased early CS+ responses, which then declined with time. Right dorsal amygdala responses to the CS+ face declined rapidly within the acquisition phase (Figs. 2c,d; Table 1b). Left dorsal amygdala responses declined more slowly, showing a mean differential response to the CS+ face in the acquisition phase, but not in the later repeat phase (Figs. 2a,b; Table 1a). These time-dependent differential amygdala responses concur with previous studies of learning-related responses in amygdala (Buchel et al., 1998; LaBar et al., 1998; Morris et al., 2001a). The lateralization of amygdala responses also concurs with previous data, showing more rapid habituation in right amygdala than left (Phillips et al., 2001; Wright et al., 2001).

A ventral region of right amygdala (Fig. 4a) showed a different response pattern to that seen in the adjacent dorsal region. Although significant mean differential responses were not evident

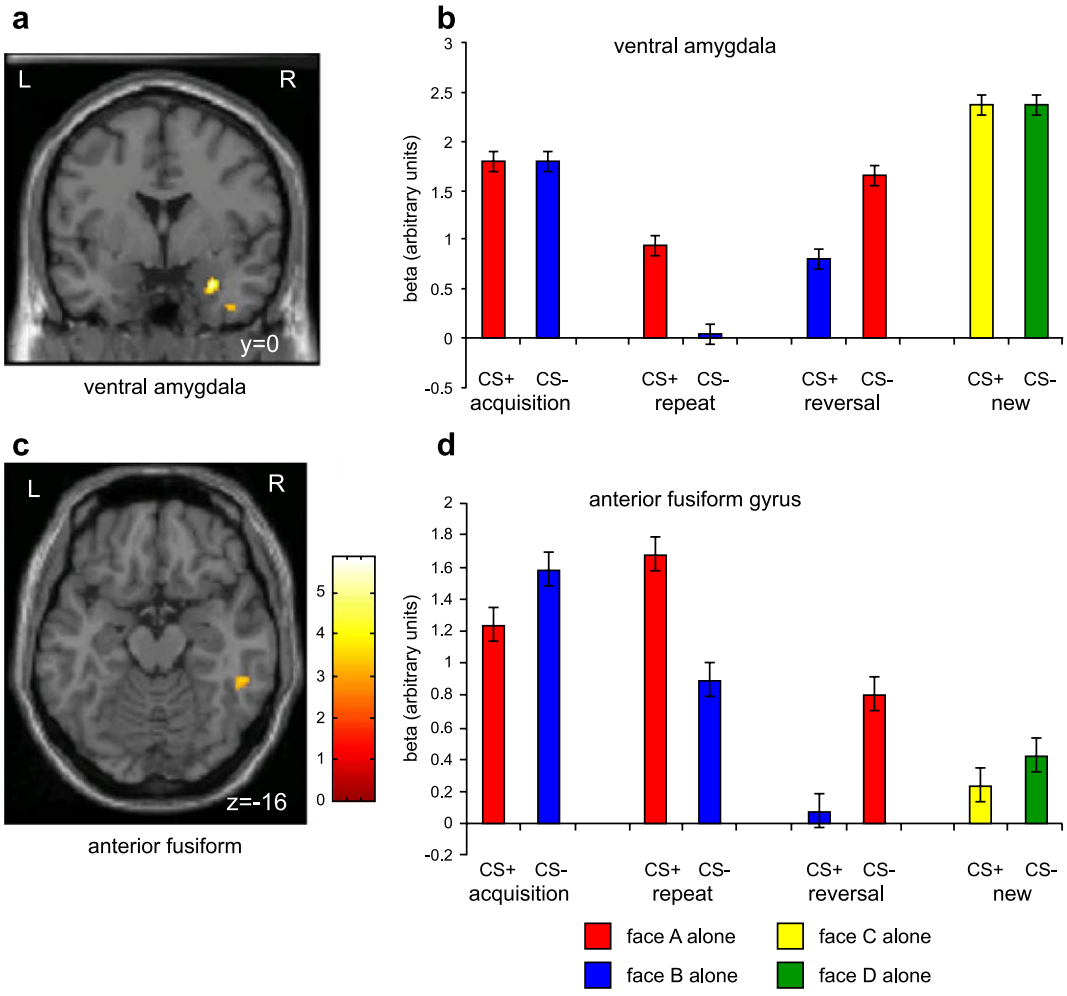


Fig. 4. Nonreversing CS+ responses. Statistical parametric maps (SPM) showing (a) region of right ventral amygdala and (b) right anterior fusiform gyrus with increased responses to face A (the original CS+) in both the repeat and reversal phases. Note that increased ventral amygdala and anterior fusiform responses to face A in the reversal phase are observed even though face A is the CS − (i.e., CS+ responses do not show rapid reversal). Activations are displayed on (a) coronal and (c) transverse sections of a canonical MRI. Graphical representations of parameter estimates for responses to face A (old CS+) are shown for (b) right ventral amygdala and (d) right anterior fusiform gyrus.

in the acquisition phase, right ventral amygdala showed increased CS+ responses in the later repeat phase. In addition to a slower acquisition of learning-related responses, right ventral amygdala also showed stable (persistent) learning-related responses, even in the face of reversal of stimulus contingencies (Fig. 4b). This learning-related functional segregation accords with a previous neuroimaging study of fear conditioning that also observed rapidly habituating responses in a dorsal amygdala subregion and non-habituating responses in an adjacent ventral amygdala subregion (Morris et al., 2001a). Electrophysiological recording of amygdala activity in rats has also revealed distinct populations of learning-related cells (Repa et al., 2001). A dorsal subregion of lateral amygdala was found to exhibit "transiently plastic" responses, characterized by rapid increases and equally rapid decreases of learning-related activity. By contrast, a ventral region of lateral amygdala showed "long-term plastic" responses, which were slower to develop, but which persisted during an extinction phase. The parallel between the functional anatomical segregation reported by Repa et al. (2001) and our present neuroimaging data is striking, although limitations in the spatial resolution of fMRI and differences in anatomical homology between human and rodent preclude identification of corresponding subnuclei.

In contrast to the persistent "memory" responses observed in ventral amygdala (Fig. 4), orbitofrontal and ventromedial prefrontal responses to conditioned stimuli reflected dynamic changes in conditioning contingency across the different experimental sessions (Fig. 3). These findings accord with animal and human data implicating orbitofrontal and ventromedial prefrontal cortex in the modification of emotional learning (Fellows and Farah, 2003; Jones and Mishkin, 1972; Rolls et al., 1994, 1996; Thorpe et al., 1983). Patients with damage to ventromedial prefrontal cortex show a characteristic clinical syndrome of irresponsible, disinhibited behaviour and fatuous affect (Eslinger and Damasio, 1985) and are impaired in changing strategies on gambling tasks despite receiving feedback indicating that the current strategy will lose money (Bechara et al., 1995; Fellows and Farah, 2003). Functional neuroimaging of healthy subjects in an instrumental reward reversal task has shown activation of orbitofrontal cortex by both feedback cues and response switching (O'Doherty et al., 2003).

Whereas previous studies implicate ventromedial prefrontal cortex in reversal of instrumental learning (e.g., O'Doherty et al., 2003; Rolls et al., 1994), the present data indicate a role for ventromedial prefrontal cortex in reversal of classical (Pavlovian) conditioning. To reverse conditioning contingencies so that, for example, stimulus A (initial CS+) becomes CS− and stimulus B (initial CS−) becomes CS+, the originally acquired conditioned responses to stimulus A must be extinguished. The process of extinction is, therefore, intrinsic to "reversal" of classical conditioning. Animal lesion studies indicate that ventromedial prefrontal cortex plays a critical role in extinction of classical (fear) conditioning (Morgan and LeDoux, 1995; Morgan et al., 1993), while electrophysiological studies in animals have shown fear extinction "memory" responses in ventromedial prefrontal neurons (Milad and Quirk, 2002). The ventromedial prefrontal and orbitofrontal responses observed in the reversal session of the present study (Figs. 3a,b) may, therefore, reflect neural processes involved in extinguishing previously acquired conditioned responses in addition to processes involved in forming new conditioned associations. Our present neuroimaging data cannot distinguish between these different components of conditioning "reversal".

A medial amygdala region (Fig. 3c) showed a further distinct pattern of learning-related activity. Responses in this region were nonhabituating but showed "orbitofrontal-like" reversal of responses in the reversal phase (Fig. 3d). This pattern of response is therefore functionally distinct from the "transient" activity in the dorsal subregion (Figs. 2a,b) and also the persistent "memory" response in the posterior ventral subregion (Figs. 4a,b). The "reversing" pattern of response observed in this medial region suggests a functional relationship with orbitofrontal cortex. Animal studies have shown that interconnections between orbitofrontal cortex and amygdala are critical for the acquisition and reversal of learned responses (Baxter et al., 2000; Schoenbaum et al., 2000). In an olfactory discrimination task, initial learning-related responses were first observed in amygdala and then later in orbitofrontal cortex (Schoenbaum et al., 2000). During reversal learning, however, correlated neural activity increased in orbitofrontal cortex, but not in amygdala (Schoenbaum et al., 2000). These results suggest that the similarity of response patterns seen in medial amygdala and orbitofrontal cortex may arise from reciprocal modulatory interactions between these structures.

Conditioning-related responses were also observed in fusiform gyrus (Figs. 3a and 4c,d). These responses accord with previous neuroimaging studies that have reported emotional and learning-related modulation of face-related activity in fusiform gyrus (Morris et al., 1998a; 2001a; Rotshtein et al., 2001). Functional segregation of activity was evident in fusiform gyrus, with an anterior region showing a "ventral amygdala-like" nonreversing pattern (Figs. 4c,d), and a posterior area showing an "orbitofrontal-like" reversing pattern (Fig. 3a). The anterior fusiform region (Figs. 4c,d), therefore, showed a persistence of learning-related responses to an individual face (i.e., face A is the CS+ initially, but CS− in the reversal phase), whereas the posterior region (Fig. 3a) responded to whichever face stimulus was currently the CS+ (i.e., irrespective of the individual identity of the face). This functional segregation mirrors previous neuroimaging data whereby posterior fusiform responses were related to general discriminability of face stimuli, while responses in an anterior fusiform region indexed the recognition of a particular individual (George et al., 2001).

Our present data accord with previous neuroimaging studies (Buchel et al., 1998; Knight et al., 1999; LaBar et al., 1998) in showing involvement of anterior cingulate in aversive conditioning. However, our findings suggest functional segregation of anterior cingulate into "reversing" and "nonreversing" patterns of learning-related activity. Intriguingly, these responses were seen in discrete, but adjacent regions of anterior cingulate ("reversing" maximal voxel $x = 12$, $y = 4$, $z = 56$, Table 1c; "nonreversing" maximal voxel $x = 18$, $y = 0$, $z = 40$, Table 1d), similar to the functional segregation observed in adjacent medial (Figs. 3c,d; Table 1c) and lateral (Figs. 4a,b; Table 1d) amygdala subregions. The extensive anatomical connectivity between anterior cingulate, amygdala, and orbitofrontal cortex helps to account for these distinct functional patterns. Increased responses in anterior cingulate have been observed during selective attention to internal emotional states (Lane et al., 1997) and during response conflict in an emotional Stroop task (George et al., 2001; Whalen et al., 1998). It has been proposed that one of the functions of anterior cingulate is to monitor and evaluate external stimuli and select appropriate responses with respect to ongoing emotional goals and priorities (Davidson et al., 2002). The complex functional segregation observed in anterior cingulate in the present study is consistent with such a role.

A "nonreversing" pattern of learning-related activity was also observed in posterior thalamus (Table 1d). Increased responses in posterior thalamus to aversively conditioned faces have been previously reported in a positron emission tomography (PET) study (Morris et al., 1997). It is notable that the posterior thalamic CS+ responses observed in the PET study were found to correlate positively with right amygdala responses, indicating the operation of learning-related functional interactions between these two structures (Morris et al., 1997). Neuroimaging studies employing backward masking of stimuli, or involving patients with occipital lesions, have also provided evidence of posterior thalamus–amygdala interactions in aversive conditioning (Morris et al., 1999, 2001b). These functional interactions are consistent with the known anatomical connections between posterior thalamus and amygdala (Jones and Burton, 1976; Linke et al., 1999).

## Conclusion

These neuroimaging data provide novel insights into the neural mechanisms underlying the modification of learned emotional responses in humans. Consistent with animal data on reinforcement learning, we found that responses in orbitofrontal cortex altered rapidly and flexibly with reversal of aversive stimulus associations. In contrast, ventral amygdala and associated brain regions showed inflexible, persistent "memory" responses that reflected previous conditioning. These observations are relevant to several phenomena described in animal learning theory. Conditioned responses are extinguished (i.e., cease to be elicited) by repeated presentations of the CS without the UCS. However, in the appropriate experimental context, extinguished conditioned responses can be renewed, reinstated, or show spontaneous recovery, indicating that a "memory" of the original conditioning remains even in the absence of "overt" conditioned behavioural responses (Bouton, 1994; LeDoux, 1996). The persistent conditioning-related responses observed in ventral amygdala and associated brain regions may represent the neurophysiological substrate for this "indelible" fear learning.

## Acknowledgment

RJD is supported by a programme grant from the Wellcome Trust.

## References

Amorapanth, P., LeDoux, J.E., Nader, K., 2000. Different lateral amygdala outputs mediate reactions and actions elicited by a fear-arousing stimulus. Nat. Neurosci. 3, 74–79.

Baxter, M.G., Parker, A., Lindner, C.C., Izquierdo, A.D., Murray, E.A., 2000. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. J. Neurosci. 20, 4311–4319.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., Damasio, A.R., 1995. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. Science 269, 1115–1118.

Blair, R.J., 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.

Bouton, M.E., 1994. Conditioning, remembering and forgetting. J. Exp. Psychol., Anim. Behav. Processes 20, 219–231.

Buchel, C., Morris, J., Dolan, R.J., Friston, K.J., 1998. Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 20, 947–957.

Cahill, L., McGaugh, J.L., 1998. Mechanisms of emotional arousal and lasting declarative memory. Trends Neurosci. 21, 294–299.

Cahill, L., Babinsky, R., Markowitsch, H.J., McGaugh, J.L., 1995. The amygdala and emotional memory. Nature 377, 295–296.

Cahill, L., Haier, R.J., Fallon, J., Alkire, M.T., Tang, C., Keator, D., Wu, J., McGaugh, J.L., 1996. Amygdala activity at encoding correlated with long-term, free recall of emotional information. Proc. Natl. Acad. Sci. U. S. A. 93, 8016–8021.

Critchley, J.L., et al., 2002. Fear conditioning in humans: the influence of awareness and autonomic arousal on functional neuroanatomy. Neuron 33, 653–663.

Davidson, R.J., Pizzagalli, D., Nitschke, J.B., Putnam, K., 2002. Depression: perspectives from affective neuroscience. Annu. Rev. Psychol. 53, 545–574.

Eslinger, P.J., Damasio, A.R., 1985. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. Neurology 35, 1731–1741.

Fellows, L.K., Farah, M.J., 2003. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. Brain 126, 1830–1837.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D., Frackowiak, R.S.J., 1995. Statistical parametric maps in functional imaging: a general linear approach. Hum. Brain Mapp. 2, 189–210.

George, N., Driver, J., Dolan, R.J., 2001. Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. NeuroImage 13, 1102–1112.

Gottfried, J.A., O'Doherty, J., Dolan, R.J., 2002. Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. J. Neurosci. 22, 10829–10837.

Jones, E.G., Burton, H., 1976. A projection from the medial pulvinar to the amygdala in primates. Brain Res. 104, 142–147.

Jones, B., Mishkin, M., 1972. Limbic lesions and the problem of stimulus-reinforcement associations. Exp. Neurol. 36, 362–377.

Josephs, O., Turner, R., Friston, K.J., 1997. Event-related fMRI. Hum. Brain Mapp. 5, 243–248.

Knight, D.C., Smith, C.N., Stein, E.A., Helmstetter, F.J., 1999. Functional MRI of human Pavlovian fear conditioning: patterns of activation as a function of learning. NeuroReport 10, 3665–3670.

LaBar, K.S., Gatenby, J.C., Gore, J.C., LeDoux, J.E., Phelps, E.A., 1998. Human amygdala activation during conditioned fear acquisition and extinction: a mixed trial study. Neuron 20, 937–945.

Lane, R.D., Reiman, E.M., Ahern, G.L., Schwartz, G.E., Davidson, R.J., 1997. Neuroanatomical correlates of happiness, sadness, and disgust. Am. J. Psychiatry 154, 926–933.

LeDoux, J.E., 1995. In search of an emotional system in the brain: leaping from fear to emotion and consciousness. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences. The MIT press, Cambridge, MA, pp. 1049–1061.

LeDoux, J.E., 1996. The Emotional Brain. Simon and Shuster, New York.

LeDoux, J.E., Iwata, J., Cicchetti, P., Reis, D.J., 1988. Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. J. Neurosci. 8, 2517–2529.

Linke, R., de Lima, A.D., Scwegler, H., Pape, H.-C., 1999. Direct synaptic connections of axons from superior colliculus with identified thalamo-amygdaloid projection neurons in the rat: possible substrates of a subcortical visual pathway to the amygdala. J. Comp. Neurol. 403, 158–170.

Marks, I., 1987. Fears, Phobias and Rituals. Oxford Univ. Press, New York.

McIntosh, A.R., Rajah, M.N., Lobaugh, N.J., 2003. Functional connectivity of the medial temporal lobe relates to learning and awareness. J. Neurosci. 23, 6520–6528.

Milad, M.R., Quirk, G.J., 2002. Neurons in medial prefrontal cortex signal memory for fear extinction. Nature 420, 70–74.

Morgan, M.A., LeDoux, J.E., 1995. Differential contribution of dorsal and ventral medial prefrontal cortex to the acquisition and extinction of conditioned fear in rats. Behav. Neurosci. 109, 681–688.

Morgan, M.A., Romanski, L.M., LeDoux, J.E., 1993. Extinction of emotional learning: contribution of medial prefrontal cortex. Neurosci. Lett. 163, 109–113.

Morris, J.S., Friston, K.J., Dolan, R.J., 1997. Neural responses to salient visual stimuli. Proc. R. Soc. Lond., B 264, 760–775.

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

Morris, J.S., Öhman, A., Dolan, R.J., 1998b. Conscious and unconscious learning in the human amygdala. Nature 393, 467–470.

Morris, J.S., Ohman, A., Dolan, R.J., 1999. A subcortical pathway to the right amygdala mediating "unseen" fear. Proc. Natl. Acad. Sci. U. S. A. 96, 1680–1685.

Morris, J.S., Buechel, C., Dolan, R.J., 2001a. Parallel neural responses in amygdala subregions and sensory cortex during implicit aversive conditioning. NeuroImage 13, 1044–1052.

Morris, J.S., DeGelder, B., Weiskrantz, L., Dolan, R.J., 2001b. Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. Brain 124, 1241–1252.

O'Doherty, J., Critchley, H., Deichmann, R., Dolan, R.J., 2003. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J. Neurosci. 23, 7931–7939.

Phillips, M.L., Medford, N., Young, A.W., Williams, L., Williams, S.C., Bullmore, E.T., Gray, J.A., Brammer, M.J., 2001. Time courses of left and right amygdalar responses to fearful facial expressions. Hum. Brain Mapp. 12, 193–202.

Repa, J.C., Muller, J., Apergis, J., Desrochers, T.M., Zhou, Y., LeDoux, J.E., 2001. Two different lateral amygdala cell populations contribute to the initiation and storage of memory. Nat. Neurosci. 4 (7), 724–731 (July).

Rolls, E.T., Hornak, J., Wade, D., McGrath, J., 1994. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J. Neurol., Neurosurg. Psychiatry 57, 1518–1524.

Rolls, E.T., Critchley, H.D., Mason, R., Wakeman, E.A., 1996. Orbitofrontal cortex neurons:role in olfactory and visual association learning. J. Neurophysiol. 75, 1970–1981.

Rotshtein, P., Malach, R., Hadar, U., Graif, M., Hendler, T., 2001. Feeling

or features: different sensitivity to emotion in high-order visual cortex and amygdala. Neuron 32, 747–757.

Sanghera, M.K., Rolls, E.T., Roper-Hall, A., 1979. Visual responses of neurons in the dorsolateral amygdala of the alert monkey. Exp. Neurol. 63, 610–626.

Schoenbaum, G., Chiba, A.A., Gallagher, M., 2000. Changes in functional connectivity in orbitofrontal cortex and basolateral amygdala during learning and reversal training. J. Neurosci. 20, 5179–5189.

Thorpe, S.J., Rolls, E.T., Maddison, S., 1983. Neuronal activity in the orbitofrontal cortex of the behaving monkey. Exp. Brain Res. 49, 93–115.

Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., Jenike, M.A., 1998. Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. J. Neurosci. 18, 411–418.

Worsley, K.J., Marrett, P., Neelin, A.C., Friston, K.J., Evans, A.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. 4, 58–73.

Wright, C.I., Fischer, H., Whalen, P.J., McInerney, S.C., Shin, L.M., Rauch, S.L., 2001. Differential prefrontal cortex and amygdala habituation to repeatedly presented emotional stimuli. NeuroReport 12, 379–383.