Transient Neural Activation in Human Amygdala Involved in Aversive Conditioning of Face and Voice

Tetsuya Iidaka[1], Daisuke N. Saito[2], Hidetsugu Komeda[2], Yoko Mano[2],
Noriaki Kanayama[1], Takahiro Osumi[1], Norio Ozaki[1],
and Norihiro Sadato[2]

Abstract

■ Elucidating the neural mechanisms involved in aversive conditioning helps find effective treatments for psychiatric disorders such as anxiety disorder and phobia. Previous studies using fMRI and human subjects have reported that the amygdala plays a role in this phenomenon. However, the noxious stimuli that were used as unconditioned stimuli in previous studies (e.g., electric shock) might have been ecologically invalid because we seldom encounter such stimuli in daily life. Therefore, we investigated whether a face stimulus could be conditioned by using a voice that had negative emotional valence and was collected from a real-life environment. A skin conductance response showed that healthy subjects were conditioned by using these stimuli. In an fMRI study, there was greater amygdala activation in response to the faces that had been paired with the voice than to those that had not. The right amygdala showed transient activity in the early stage of acquisition. A psychophysiological interaction analysis indicated that the subcortical pathway from the medial geniculate body to the amygdala played a role in conditioning. Modulation of the subcortical pathway by voice stimuli preceded the transient activity in the amygdala. The finding that an ecologically valid stimulus elicited the conditioning and amygdala response suggests that our brain is automatically processing unpleasant stimuli in daily life. ■

INTRODUCTION

Significant involvement of neural responses in the amygdala and related structures has been observed in the classical aversive conditioning paradigm in animals, which involves learning by association between neutral and noxious stimuli (Sotres-Bayon, Bush, & LeDoux, 2004; Quirk, Likhtik, Pelletier, & Pare, 2003; Quirk, Armony, & LeDoux, 1997). A previously neutral stimulus (conditioned stimulus or CS) elicits behavioral and autonomic responses after effective pairing with an unconditioned stimulus (US) that has unpleasant features. This phenomenon is a potential model for psychiatric disorders such as anxiety disorder (Milad, Rauch, Pitman, & Quirk, 2006), phobia, and other stress-related disorders (Ollendick & Hirshfeld-Becker, 2002) in human subjects. Investigation of the precise mechanisms of aversive conditioning helps clarify the pathogenesis of the disorders and helps us find effective treatment.

However, exposing human subjects to a US such as an electric shock or a loud tone might be ecologically invalid because we seldom encounter such intense and physical stimuli in daily life. In a study using US such as an unpleasant noise collected from the environment, the subjects were effectively conditioned to the stimuli and autonomic responses were elicited (Neumann & Waters, 2006). In the case of social anxiety disorder which exhibits elevated fear in social situations, it is hypothesized that the symptoms may be acquired through aversive conditioning in social circumstances (Ollendick & Hirshfeld-Becker, 2002). Mild emotional stress induced by simulated social situations resulted in significant changes in the plasma cortisol levels in patients with major depressive disorder (Belmaker & Agam, 2008). Thus, in addition to the degree of the stimulus intensity, we considered it important to ensure that the experimental paradigm would simulate the real-life situations in which the subjects were exposed to emotional stress.

Neural correlates involved in aversive conditioning were investigated by using fMRI and human subjects (Phelps, Delgado, Nearing, & LeDoux, 2004; Buchel, Morris, Dolan, & Friston, 1998; LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998). In these studies, in line with animal experiments, the amygdala was a key structure in the conditioning process and mediated neuronal input from several perceptual modalities to elicit autonomic responses. In particular, the amygdala response was greater in the first half than in the second half of the acquisition phase (Morris, Buchel, & Dolan, 2001; Buchel, Dolan, Armony, & Friston, 1999; Buchel et al., 1998; LaBar et al., 1998) and it showed rapid habituation. However, the temporal dynamic of the amygdala response in the conditioning paradigm has hitherto not been investigated. Although it is not possible to track

[1]Nagoya University, Nagoya, Japan, [2]National Institute for Physiological Sciences, Okazaki, Japan

neuronal responses with fMRI, it has been shown that the activity of the amygdala changed across the experimental runs when the emotional faces were presented to the subjects (Phillips et al., 2001; Wright et al., 2001). Therefore, we think that it is possible to investigate how the CS–US relationship is established and how the neural responses involved in the conditioning are regulated across experimental runs. Furthermore, it is still not clear as to how neural activity in the amygdala was modulated by combined inputs from different perceptual modalities.

In our event-related fMRI study using a 3-Tesla MRI scanner and healthy male subjects, we aim to address the following questions regarding neural mechanisms of aversive conditioning: (1) whether a US that was collected from the environment and that resembled a real-life situation causes aversive conditioning; (2) what the temporal time course of the increase and decrease in amygdala response associated with conditioning is; and (3) how neural input from the auditory modality with unpleasant emotional tone modulates amygdala activity. To answer these questions, we developed a novel paradigm that used a picture of a face and an unpleasant voice; we measured skin conductance response (SCR) in a separate group of subjects that did not participate in the fMRI experiment. The fMRI data were analyzed to examine the effect of a temporal change in amygdala activity across the experimental runs. We show that our novel paradigm successfully caused conditioning as measured by the SCR, and it enhanced the amygdala activity. The neural input from the subcortical auditory pathway might play a modulatory role, particularly in an initial stage of aversive learning.

METHODS

Subjects

Eighteen healthy male volunteers (right-handed, mean age ± SD = 21.4 ± 2.7 years) were recruited for participation in the fMRI study. Only the male subjects were used because there are significant confounding issues of sex and the menstrual cycle in fear conditioning and extinction in human subjects (Milad, Goldstein, et al., 2006). The subjects provided written informed consent according to the procedure approved by the local ethics committee. The physical and mental conditions of the subjects were carefully checked by a physician. The Structured Clinical Interview for DSM-IV was used to exclude those who had a past and/or current history of psychiatric disease. This study was approved by the ethics committee at the National Institute for Physiological Sciences.

General Concept of the Experiment

The aim of the present experiment was to simulate stressful condition in a social situation and use fMRI to measure the hemodynamic responses elicited by the CS. A partial conditioning paradigm in which half of the CS is re-

inforced by the US would be useful for excluding the confounding effect of aversive USs. Such a paradigm has already been adopted in previous neuroimaging studies (Straube, Weiss, Mentzel, & Miltner, 2007; Morris et al., 2001; Buchel et al., 1998, 1999). In these studies, the CS was either a neutral face (Buchel et al., 1998), a neutral tone (Buchel et al., 1999), or a symbol matrix (Straube et al., 2007) that was paired with the conventional US. According to these experiments, we used a picture of an individual face with a neutral expression as a CS that was later paired with a US. In addition, the neutral face was repeatedly presented before learning the CS–US relationship during the habituation phase. A unique aspect of the present study was that a negative emotional voice was used as a US. Several neuroimaging studies have used voices with emotional intonation as experimental stimuli. These studies have shown that activation in the superior temporal gyrus is specifically involved in processing the emotional prosody of the voice (Wildgruber, Ackermann, Kreifelts, & Ethofer, 2006; Grandjean et al., 2005). Thus, we think that combining an individual face and a voice with negative emotional valence would be appropriate for simulating real-life situations where a person feels emotional stress from other person's behavior. The facial expression changed from a neutral one to a negative one in each aversive event because change in facial expression is usually associated with the emotional voice in real-life situations.

Experimental Stimuli

Pictures of neutral and negative (angry or disgusted) facial expressions were selected from a Japanese male face picture database. The faces of three different individuals with neutral and negative expressions were used as the experimental stimuli. We chose an angry or disgusted face as the stimulus rather than a fearful face because we wanted to simulate a social situation wherein the subjects were exposed to emotional stress. In the case of angry and disgusted expressions, negative emotion is directed from the individual in the picture to the subjects. In contrast, a picture of a fearful face would depict the situation wherein the individual in the picture is exposed to the stress, which does not meet our purpose. These pictures were digitized and equalized in terms of shape and luminance using commercial software (Photoshop; Adobe, San Jose, CA). Twelve healthy subjects who did not participate in the fMRI or SCR experiment (mean age = 24 years) were presented with these pictures. The subjects classified the expression in each picture using six basic emotions and a neutral emotion and rated the intensity (from 0 for neutral to 7 for extremely negative). The mean proportion of correct responses was 89% for both the neutral and negative faces. Disgust accounted for 50% of the emotions labeled to the negative faces; anger, 39%; sadness, 11%. The results of the intensity rating showed that the mean (SD in parentheses) intensity was 0.16 (0.5) and 4.25 (1.13) for the neutral and negative faces, respectively. A two-way ANOVA revealed

that the mean intensity rating significantly differed between the neutral and negative faces [$F(1, 66) = 387, p < .001$], and there was no significant main effect of stimuli [$F(2, 66) = 0.21, p = .8$] or interaction of emotion and stimuli [$F(2, 66) = 1.5, p = .22$]. The mean intensity (SD) of each emotion was 4.2 (1.3) for disgust, 4.6 (0.6) for anger, and 3.5 (1.3) for sadness. A picture of a house was used for the purpose of presentation during the habituation phase.

Two of the three individual faces were chosen for each subject in a counterbalanced manner to create a face–voice stimulus pair. One of the faces was assigned to the CS+ and the other to the CS− (Figure 1). During the experiment, each subject saw pictures of two individuals (the CS+ individual and the CS− individual). The CS+ was a conditioned stimulus that was paired with the US in the half of the presentation during the acquisition phase, whereas the CS− was never paired with the US during the experiment. Every stimulus was presented for a duration of 1500 msec. CS+ was further classified into CS+ unpaired (CS+up) and CS+ paired (CS+p) during the acquisition phase. The CS+up involved the presentation of only the neutral face of the CS+ individual for a duration of 1500 msec. The CS+p was a combined presentation of the face and voice, as described in the next section, the emotional valence of which changed from neutral to negative during the presentation. The CS+ presented during the habituation and extinction phases were identical with the CS+up presented during the acquisition phase.
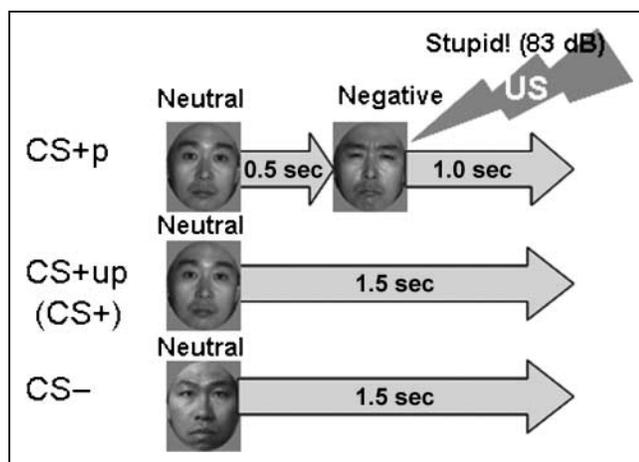

Figure 1. Schematic illustration of the experimental stimuli used in our study. Two individual faces among the three were randomly selected for each subject. They were assigned to either CS+ or CS−. The unconditioned stimulus (US) was a voice saying "stupid" with a sound pressure of 83 dB. The time course of CS+p began with the presentation of the neutral face of CS+ for 500 msec followed by simultaneous presentation of the negative face of CS+ and a US for 1000 msec. CS+up presented during the acquisition phase was identical with CS+ presented during the habituation and extinction phases. Both CS+up and CS+ were presentations of only the neutral face for a duration of 1500 msec. CS− represented the stimulus with the presentation of a neutral face of another individual, and it was never paired with the voice stimulus. The ISI was set at 13.5 sec during the Hab and Acq phases and at 3.5 sec during the Ext phase.

A voice stimulus was used as the US. The stimulus was digitally recorded from a male volunteer who said "stupid" loudly with a negative emotional valence. The voice stimulus was digitally combined with each negative face using commercial audiovisual software (Symphomovie; Epson, Suwa, Japan). Time course of the audiovisual stimuli began with the presentation of the neutral face for 500 msec followed by simultaneous presentation of a negative face of the same individual and a negative voice for 1000 msec (Figure 1, CS+p). The stimuli were converted to MPEG format and presented to the subjects using Presentation software (Neurobehavioral Systems, Albany, CA). The sound stimulus was delivered with a monaural headphone in the scanner with a sound pressure of approximately 83 dB. All subjects correctly received the voice stimulus in a noisy environment during the fMRI experiment. The visual stimuli were projected onto a transparent screen, which was hung from the bore of a magnet, at a distance of 75 cm from the subject's eyes. The subjects viewed the stimuli through a tilted mirror that was attached to the head coil of the scanner. The response was measured using a magnetic-compatible button box held in the subject's right hand.

Experimental Procedure

The fMRI experiment consisted of three phases, namely, habituation (Hab, 1 run), acquisition (Acq, 5 runs), and extinction (Ext, 2 runs), as shown in Figure 2. During the Hab phase, two faces (CS+ and CS−) and a house were repeatedly and randomly presented, one at a time [duration = 1.5 sec, interstimulus interval (ISI) = 13.5 sec]. The face stimuli were presented 10 times each, and the house stimulus was presented 20 times. The subject pressed the left button when he saw the face and the right button when he saw the house. The house stimuli were not presented during the Acq or Ext phases that followed the Hab phase. During the Acq phase, the two faces seen during the Hab phase (CS+ and CS−) were repeatedly presented one at a time with the same duration and ISI. A long ISI was chosen because there was a need to track hemodynamic responses effectively. Across five runs, the CS+ face was presented 50 times (10 times in each run). Half of the presentation of CS+ was paired with the voice stimuli (CS+p), and the other half was not (CS+up). The CS− was presented 50 times (10 times in each run) and never paired with the voice. During each run, the presentation order was randomized, except that the initial two stimuli were always the CS+p. The subject pressed the left button whenever he saw the face, and no further judgment was required. The subject was instructed that a face on the screen would say "stupid" to the subject. During the Ext phase, two faces (CS+ and CS−) were repeatedly and randomly presented one at a time (25 times each) without the voice stimuli (duration = 1.5 sec, ISI = 3.5 sec). The subject was required to judge whether the face was paired with the voice during the Acq phase and to press the corresponding

Figure 2. Schematic illustration of the experimental design used in the present study. In the habituation phase (Run1), CS+, CS−, and a house were repeatedly and randomly presented one at a time. The subject pressed the left button when he saw the face and the right button when he saw the house. In the acquisition phase (Runs 2–6), CS+ and CS− were repeatedly presented one at a time across five runs. Half of the presentation of CS+ was paired with the voice stimuli (CS+p), whereas the other half was not
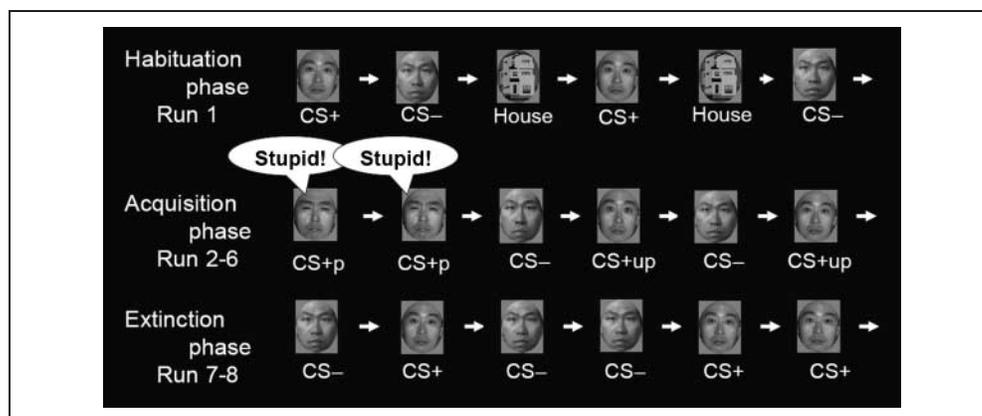
(CS+up). The CS− was never paired with the voice. The two initial stimuli were always the CS+p. The subject pressed the left button whenever he saw the face. The extinction phase (Runs 7 and 8); CS+ and CS− were repeatedly and randomly presented one at a time without the voice stimuli. The subject was required to judge whether the face was paired with the voice during the acquisition and to press the corresponding button.

button. The run was repeated twice with the same set of stimuli randomly intermixed. Thus, the number of presentations for each of the two individual faces was equal during the experiment.

Image Acquisition and Preprocessing

Functional images of the whole brain were obtained in an axial–oblique position using a 3-Tesla MRI scanner (Allegra; Siemens, Erlangen, Germany) that was equipped with a single-shot echo-planar imaging (EPI) (TR = 2.3 sec, TE = 30 msec, flip angle = 80°, 64 × 64 matrix and 36 slices, voxel size = 3 × 3 × 3 mm) and it was sensitive to BOLD contrast. The number of images obtained during each phase was as follows: the Hab phase, 266; the Acq phase, 690; and the Ext phase, 232. After discarding the first six images, the next successive images in each run were subjected to the analysis. A high-resolution anatomical T1-weighted image was also acquired (MP-RAGE, TR = 2.5 sec, TE = 4.38 msec, flip angle = 8°, 256 × 256 matrix and 192 slices, voxel size = 0.75 × 0.75 × 1 mm) for each subject. Data were analyzed using SPM2 (The Wellcome Department of Imaging Neuroscience, London, UK). First, all the volumes were realigned spatially to the first volume, and the signal in each slice was realigned temporally to that obtained in the middle slice using a sinc interpolation. The resliced volumes were then normalized to the standard MNI space by using a transformation matrix obtained from the normalization process of the mean EPI image of each individual subject to the EPI template image. The normalized images were spatially smoothed with an 8-mm Gaussian kernel.

Statistical Analysis

Following preprocessing, a statistical analysis of each individual subject was conducted using the general linear model. At the first level (a fixed effects model), each event was modeled as a hemodynamic response function and its temporal derivative. High-pass frequency filters (128 sec) were applied to the time series data. The images were scaled to a grand mean of 100 over all voxels and scans within a session. In the subsequent analysis, the following conditions were modeled separately: the Hab phase, CS+, CS−, and house; the Acq phase, CS+p, CS+up, and CS−; the Ext phase, CS+ and CS−; in each run another condition that involved incorrect responses was included. The US was not explicitly modeled but was included in the CS+p condition because the onset of voice stimulus was delayed only 500 msec from the onset of the CS. In addition, six movement parameters obtained during the realignment were entered as regressors. Parameter estimates for each condition and for the difference between these conditions were computed from the least mean square fit of the model to the time series data. Images of parameter estimates representing event-related activity at each voxel were created for each condition and subject. The neural response associated with the house stimuli was analyzed, but the data have not been reported here.

Subtraction Analysis

The following contrasts of interest were created in each phase and in each subject.

The Hab Phase

(1) CS+ versus CS−. This contrast was to confirm that the CS+ and CS− did not elicit differential activation in the region of the amygdala and hippocampus before the acquisition of face–voice pair.

The Acq Phase

(2) CS+up versus CS−. This contrast investigated whether the CS+up, even when unaccompanied by the voice stimulus, elicited differential activation as compared with the CS− during the Acq phase.
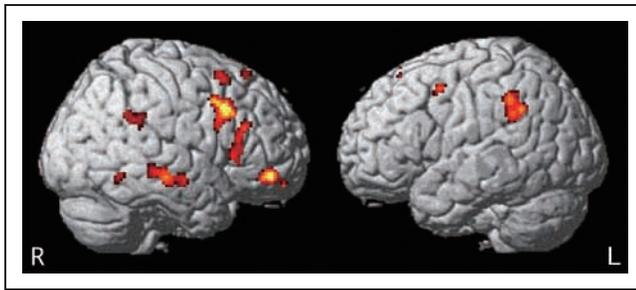
Figure 3. Brain regions where the signal was greater for the CS+up than for the CS− during the Acq phase are rendered on the surface template of SPM2. The statistical threshold was set to $p = .001$, uncorrected. Detailed information for each cluster is listed in Table 1. R = right hemisphere; L = left hemisphere.
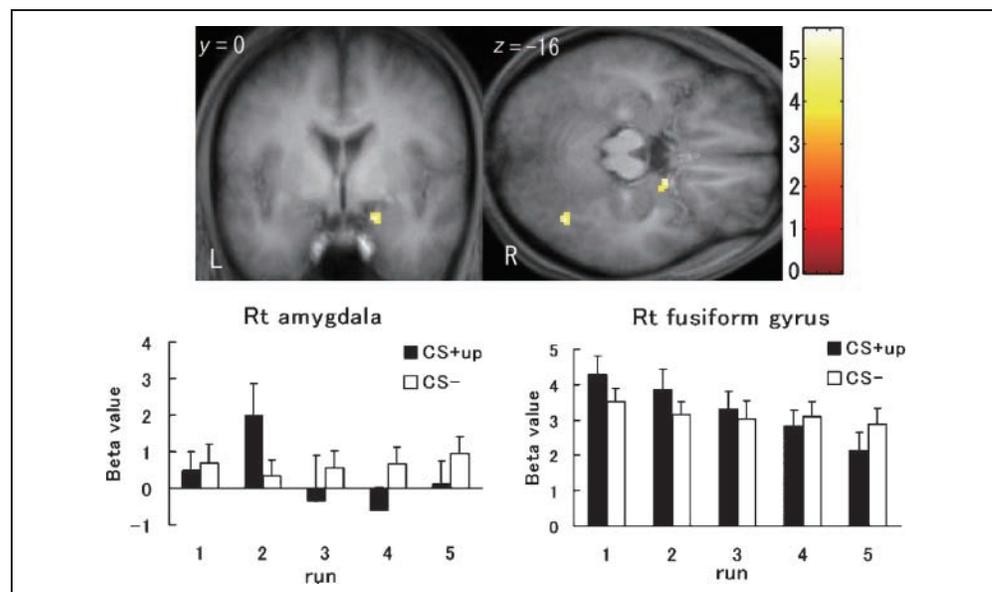
(3) Time modulation effect of CS+up versus CS−. We hypothesized that during the Acq phase the difference in activation between the CS+up and CS− would be modulated across five runs. To test this, we entered the contrast (2, 1, 0, −1, −2) for the CS+up and (−2, −1, 0, 1, 2) for the CS− (number represents contrast in each of five runs).

(4) CS+p versus CS+up. This contrast examined the effect of voice stimulus in primary auditory cortex and medial geniculate body (MGB).

The Ext Phase

(5) CS+ versus CS−. This contrast examined whether the CS+ and CS− elicited differential activation after the acquisition of the voice–face pair.

At the second level, the results for each subject were entered into a group analysis (a random effects model) by applying a one-sample t test to the contrast images.

The statistical threshold was set at $p = .001$, uncorrected for multiple comparisons for height, and at $k = 20$ voxels for spatial extent. The results for the Hab phase that compared the CS+ and CS− before the acquisition of face–voice coupling is superimposed on the mean T1 image of 18 subjects (Supplementary Figure 1). The results that compared neural responses to the CS+up and CS− during the Acq phase are rendered on the surface template (Figure 3). The time modulation effect on the difference between the CS+up and CS− was observed in the region of the amygdala and fusiform gyrus, and the results are shown on the mean T1 image (Figure 4). Beta values in each of five runs were extracted from the peak voxel in the right amygdala and fusiform gyrus to illustrate the change in activity across runs (Figure 4). The results of the Ext phase that compared the CS+ and CS− after the acquisition of face–voice coupling are shown in Figure 5. The mean parameter estimates were extracted from a spherical ROI ($r = 8$ mm) drawn at each peak voxel by using MarsBaR software (Brett, Anton, Valabregue, & Polini, 2002; Rorden & Brett, 2000) and plotted in Figure 5. In Tables 1–3, the voxel size, T value, coordinates, Brodmann's area, and region name of the significant cluster are tabulated.

Psychophysiological Interaction Analysis

Another approach that investigated the functional relationship between the regions involved in the processing of face–voice pairs during the Acq phase was a psychophysiological interaction (PPI) analysis (Friston et al., 1997). A PPI analysis aims to explain neural responses in one brain area in terms of the interaction between the influences of another brain region and a task condition. The analysis was conducted to examine whether signal coupling between the MGB and the amygdala would differ

Figure 4. The time modulation effect on the difference between the CS+up and CS− that was found in the right amygdala and fusiform gyrus is shown in the mean T1 image. The statistical threshold was set to $p = .001$, uncorrected. Plots of mean beta values extracted from the right amygdala and fusiform gyrus are illustrated. Black and white columns indicate the mean beta value in each run for the CS+up and CS−, respectively (bar, *SEM*). Detailed information for the cluster is listed in Table 2.

Figure 5. The comparison between CS+ and CS− during the Ext phase is shown in the mean T1 image. The statistical threshold is set to $p = .001$, uncorrected, and $k = 20$ voxels. Regions in the hippocampus (left), orbital gyrus (middle, arrow), and fusiform gyrus (right, arrowhead) showed increased activation in response to CS+ as compared to CS−. The mean parameter estimates (bar, *SEM*) extracted from each ROI (spherical, $r = 8$ mm) are plotted in the figure. Detailed information about the cluster is listed in Table 3. There was no region in which the signal was significantly greater for the CS− than for the CS+.
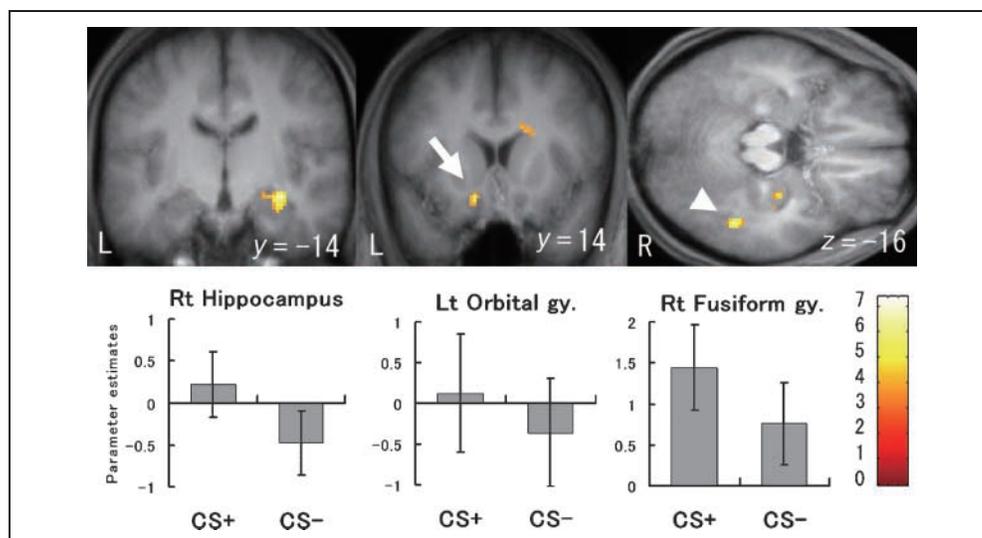
between the conditions. Our hypothesis was that when the aversive voice stimulus was paired with the face stimulus, it would have a substantial effect on amygdala activity through a subcortical pathway involving the MGB (LeDoux, Farb, & Ruggiero, 1990). Therefore, we predicted that the functional coupling between the regions would be greater in the CS+p condition than in the CS+up condition.

The PPI analysis consists of three regressors: the psychological variable, representing the task condition; the physiological variable, representing the signal response in the MGB; and the interaction term of these two. The psychological variable used was a vector coding for the task condition (1 for CS+p, −1 for CS+up) convolved with the hemodynamic response function. The individual time series for the right MGB ($x = 16$, $y = −26$, $z = −4$) was obtained by extracting the first principal component from a sphere ($r = 4$ mm) centered on the peak voxel of the group analysis, as shown in Supplementary Figure 2. These time series were mean-corrected and high-pass filtered to remove low-frequency signal drift. The physiological factor was then multiplied with the psychological factor; this constitutes the interaction term. PPI analyses were then

Table 1. Significant Difference in Activation between the CS+up and CS− during the Acq Phase

Contrast	Voxels	T	x y z (mm)	Hem	BA	Region Name
CS+up minus CS−	141	10.35	44 48 −10	R	10	lateral orbital gy.
	20	4.55	30 56 −14	R	11	anterior orbital gy.
	161	5.26	60 24 6	R	44	inferior frontal gy.
	56	5.10	34 14 56	R	6	middle frontal gy.
	402	4.79	40 16 34	R	9	middle frontal gy.
	115	5.05	6 32 60	R	6	superior frontal gy.
	211	5.93	−60 −44 32	L	40	supramarginal gy.
	75	4.67	60 −48 30	R	40	supramarginal gy.
	37	5.46	−42 4 48	L	6	precentral gy.
	241	5.61	54 −22 −10	R	21	middle temporal gy.
	24	4.73	60 −54 −10	R	37	inferior temporal gy
	38	4.65	26 26 −8	R		insula
CS− minus CS+up	30	5.15	16 −12 −28	R	28	parahippocampal gy.
	30	5.09	−50 −6 38	L	6	precentral gy.

hem = hemisphere; R/L = right/left; BA = Brodmann's area; gy. = gyrus. x, y, and z represent the MNI coordinates.

Statistical threshold was set at $p = .001$, uncorrected and $k = 20$ voxels.

Table 2. The Time Modulation Effect on the Difference between the CS+up and CS− during the Acq Phase

Voxels	T	x y z (mm)	Hem	BA	Region Name
26	5.47	18 0 −16	R		amygdala
28	5.50	40 −62 −16	R	37	fusiform gy.
66	4.70	10 −20 50	R	24	cingulate gy.
25	5.67	−64 −42 6	L	21	middle temporal gy.

Statistical threshold was set at $p = .001$, uncorrected and $k = 20$ voxels.

carried out for each subject involving the creation of a design matrix with the interaction term, the psychological factor, and the physiological factor as regressors (Neufang, Fink, Herpertz-Dahlmann, Willmes, & Konrad, 2008). The data of the five runs were included in a single design matrix for each subject.

A contrast image for the interaction term from the PPI analysis for each run was subjected to a second-level analysis. Five PPI group analyses were conducted for each of the five runs during the Acq phase. As we predicted that the functional coupling between the MGB and the amygdala would differ between the CS+p and CS+up, the small-volume correction [SVC] ($p = .05$) was conducted in a spherical ROI ($r = 4$ mm) set at the peak voxel in the left ($x = −16$, $y = −4$, $z = −22$) and right ($x = 16$, $y = −4$, $z = −22$) amygdala. The parameter estimates of the PPI were extracted from the ROI placed at peak amygdala voxels in the right hemisphere. The results of the PPI analysis and the plot of the effect of interaction term in each of the five runs are shown in Figure 6.

Skin Conductance Response

Eleven healthy male volunteers (right-handed, mean age ± $SD = 24.1 ± 2.5$ years) who did not participate in the fMRI study participated in the SCR study. SCR data were recorded in a shielded room using an MP-100 psychophysiological monitoring system (BioPac Systems, Santa Barbara, CA). For each subject, disposable Ag/AgCl electrodes (Nihon

Koden, Tokyo, Japan) were attached to the volar surface of the second phalanx of the first and the third fingers of the left hand. The experimental procedure was slightly modified to measure sufficient SCR; that is, ISI was set to 20 sec and no behavioral response was required from the subject. The subject was in a sitting position and saw the stimuli on a CRT monitor (17 in.). The voice sound was delivered by a headphone. Skin conductance was measured throughout the Acq phase and data were analyzed off-line using Acknowledge Software (BioPac Systems). SCR was defined as the change from the baseline to the peak of the response within a 0.5–4 sec time window after the onset of each stimulus. A log transformation (log [SCR + 1]) was performed on the peak SCR amplitude to normalize the data. Mean values were calculated by averaging the data across trials for each condition in each subject.

RESULTS

Skin Conductance Response and Behavioral Results

The mean SCR values (*SD*) for the CS+up and CS− were 0.01 (0.01) and 0.004 (0.003), respectively. A two-way repeated measure ANOVA on the SCR data showed that there was a significant main effect of the conditioning [$F(1, 10) = 5.91$, $p = .03$]. Although the SCR data were measured in a group of subjects that did not participate in the fMRI study, the results indicate that the CS+ was

Table 3. Significant Difference in Activation between the CS+ and CS− during the Ext Phase

Voxels	T	x y z (mm)	Hem	BA	Region Name
71	7.33	32 −14 −10	R		hippocampus–amygdala
76	6.92	48 −44 −16	R	37	fusiform gy.
23	5.39	−12 14 −12	L	11	orbital gy.
40	5.37	26 −22 8	R		thalamus
61	6.18	16 26 −4	R		caudate nucleus
135	5.79	16 26 12	R		caudate nucleus
45	4.61	18 18 26	R		caudate nucleus

Statistical threshold was set at $p = .001$, uncorrected and $k = 20$ voxels.

There was no region in which activation was larger for the CS− than for the CS+ at this threshold.
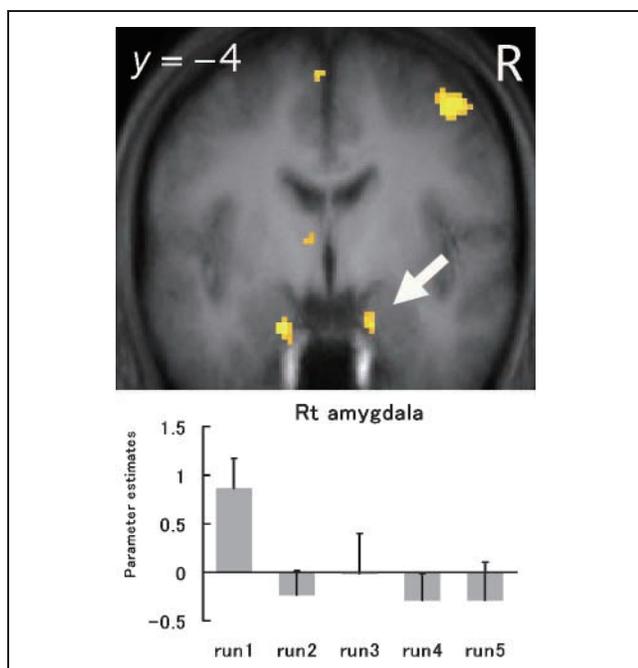
Figure 6. (Top) The results of PPI analysis in the first run of the Acq phase. Significant clusters (left and right amygdala, both survived $p = .05$, SVC) were superimposed on the mean T1 image ($y = -4$ mm). The figure was thresholded at $p = .01$, uncorrected for multiple comparisons for a presentation purpose. The correlation between the signal time course of MGB and that of the amygdala was significantly modulated by the task condition (i.e., CS+p vs. CS+up). This indicates that pairing of the voice sound significantly increased the functional coupling of the amygdala and MGB. (Bottom) Plots of the parameter estimates of the right amygdala (white arrow) for each of the five runs. An x-axis indicates the run in the Acq phase, and a y-axis indicates the degree of interaction. Each plot represents the mean (bar, *SEM*) of 18 subjects. The degree of the interaction was greater in the first run than in the later runs.

successfully conditioned by the voice stimuli to evoke autonomic responses. Neither the mean accuracy (proportion of corrected response) nor the mean reaction time measured during the fMRI experiment differed significantly between the conditions in each phase (Supplementary Table 1).

fMRI Results

The Acq Phase

Several regions in the prefrontal, temporal, and parietal cortices showed a significantly greater activation for the CS+up than for the CS−, as listed in Table 1 and shown in Figure 3. These clusters, particularly those observed in prefrontal cortex, are located mainly in the right hemisphere.

A time modulation effect on the difference between the CS+up and CS− was found in the right amygdala and fusiform gyrus and is listed in Table 2 and shown in Figure 4. A plot of mean beta values extracted from the right amygdala revealed that the predominant response to the CS+up was observed in the second run and the activity decreased afterward, whereas the response to the CS− did not change

significantly through five runs. The pattern of change in the beta value indicates that there was a time lag between the commencement of face–voice coupling and an increment of amygdala response. In the fusiform gyrus, the response to the CS+up decreased linearly across runs and that to the CS− was relatively constant. The inverse contrast of the time modulation effect on the difference between the CS+up (−2, −1, 0, 1, 2) and CS− (2, 1, 0, −1, −2) showed no significant result.

The Ext Phase

The comparison between the CS+ and CS− revealed that there was significantly greater activation in response to the CS+ than to the CS− in the right hippocampus (Figure 5). The cluster extended to the amygdala region in the anterior slices. Regions in the orbital gyrus and fusiform gyrus showed increased activation in response to the CS+ as compared with the CS− during the Ext phase (Table 3 and Figure 5). The activity in these regions did not differ between the first and second runs at the threshold of $p = .05$ (uncorrected). There was no region in which the signal was significantly greater for the CS− than for the CS+ at $p = .001$, uncorrected.

Psychophysiological Interaction

The PPI analysis revealed that during the first run of the Acq phase, there was a significant interaction effect on the signal coupling between the MGB and the amygdala (left: $x = -16$, $y = -4$, $z = -22$, $T = 3.65$, $p = .013$, SVC; right: $x = 16$, $y = -4$, $z = -22$, $T = 2.97$, $p = .04$, SVC), as shown in Figure 6. The results indicate that the signal correlation between the MGB and the amygdala was significantly greater for the CS+p condition than for the CS+up condition during the early stage of the Acq phase. Plots of the parameter estimates representing the degree of the interaction show that the interaction effect is greater in the first run than in the later runs. An ANOVA conducted on these data revealed that although the main effect of the run was at trend level [$F(4, 85) = 4.37$, $p = .08$], post hoc comparisons between Run 1 and Run 2, Run 1 and Run 4, and Run 1 and Run 5 were significant at $p = .05$ (Fisher's LSD test).

DISCUSSION

Our study investigated the neural responses in the amygdala and related structures using fMRI and healthy male volunteers while the subject was conditioned to the face picture and an aversive voice. The results of the SCR conducted outside the scanner showed that the subjects were conditioned to the stimuli and evoked autonomic responses. The fMRI results indicated that the transient increase in the early stage of the Acq phase and subsequent reduction in amygdala activity may relate to the conditioning of aversive stimuli. In addition, PPI analysis showed

that neural coupling in the subcortical pathway from the MGB to the amygdala in the initial stage of the Acq phase may play a pivotal role in conditioning. This suggests that an emotional voice, although less intense than conventional US, is sufficient to cause conditioning and modulate amygdala response. The results may have implications in understanding how stressful events are processed in our brain and for preventing stress-related disorders.

In previous fMRI studies using the classical aversive conditioning paradigm, painful electrical shock (Alvarez, Biggs, Chen, Pine, & Grillon, 2008; Straube et al., 2007; Kalisch et al., 2006; Birbaumer et al., 2005; Glascher & Buchel, 2005; Knight, Cheng, Smith, Stein, & Helmstetter, 2004; Phelps et al., 2004; Cheng, Knight, Smith, Stein, & Helmstetter, 2003; Jensen et al., 2003; Thiel, Friston, & Dolan, 2002; Knight, Smith, Stein, & Helmstetter, 1999; LaBar et al., 1998) and loud tone (Dunsmoor, Bandettini, & Knight, 2008; Knight, Nguyen, & Bandettini, 2005; Morris & Dolan, 2004; Buchel et al., 1998, 1999; Morris, Ohman, & Dolan, 1998) have been used as US to evoke autonomic responses and amygdala activation. Other types of US, such as odor stimuli (Gottfried & Dolan, 2004) and negative emotional pictures (Nitschke, Sarinopoulos, Mackiewicz, Schaefer, & Davidson, 2006), also evoked amygdala activation. We combined audiovisual stimuli that resembled real-life stimuli, and we evoked autonomic responses and amygdala activation in the healthy male volunteers. In an SCR study using an unpleasant sound as the US, the subjects were successfully conditioned to the auditory stimulus that was sampled from the environment (Neumann & Waters, 2006). These results indicate the possibility that we are automatically and unwittingly conditioned to our environmental stimuli in daily life.

Rapid habituation or reduction in the amygdala response to the CS+ stimuli that is characterized by significant Time-by-Condition interaction has been reported in previous fMRI studies (Birbaumer et al., 2005; Morris & Dolan, 2004; Morris et al., 2001; Buchel et al., 1998, 1999; LaBar et al., 1998). In line with these studies, our results showed that the right amygdala activity increased in an early stage of the Acq phase and rapidly decreased afterward. However, in contrast to previous studies that compared signal changes between the first and the second halves of the experiment, our experimental procedure spanning five runs revealed the time course of the neural responses and a dramatic increase in amygdala activity in the second run. This pattern of activity accords with the pattern of firing rate of the lateral amygdala cells in rat during fear conditioning (Quirk et al., 1997) and may indicate an inhibitory effect on the amygdala cells from prefrontal cortex in the later stage of conditioning (Leal-Campanario, Fairen, Delgado-Garcia, & Gruart, 2007; Sotres-Bayon et al., 2004; Quirk et al., 2003). In the fusiform gyrus that showed significant time modulation effect, the signal in response to the CS+up linearly decreased along the time course, suggesting an effect of habituation on neural activity associated with visual processing (Ishai, Pessoa, Bikle, & Ungerleider,

2004). The difference in the signal time course between the amygdala and fusiform gyrus may be attributable to the fact that the amygdala is involved in emotional processing of faces, and the fusiform is primarily involved in low-level processing of faces.

The other cortical regions that showed tonic and consistent responses during the Acq phase were dorsolateral prefrontal cortex (DLPFC), superior temporal sulcus (STS), and temporo-parietal junction. The results, showing that the location of clusters in DLPFC and STS was strongly right lateralized, indicate that the right hemisphere may play a role in the emotional processing (Schirmer & Kotz, 2006). The DLPFC activation associated with conditioning has been reported in several neuroimaging studies that showed that the DLPFC activity has been related with explicit awareness of the CS–US relationship (Carter, O'Doherty, Seymour, Koch, & Dolan, 2006) and with expectancy of the US occurrence (Dunsmoor et al., 2008). In another study using negative emotional pictures, right DLPFC activity increased during the anticipation phase and positively correlated with subject's negative affect (Nitschke et al., 2006). These previous findings suggest that DLPFC may mediate top–down attentional processes or working memory function (LeDoux, 2000) related with conditioning that persists during the Acq phase.

Another explanation for the prefrontal and parietal activation found in the subtraction between the CS+up and CS− is that these areas may be involved in the neural processing underlying the understanding of the mental states of other people. When the CS+up face that was previously paired with the negative emotion was presented, the subject would have conjectured the intention or emotion of the CS+ individual from his neutral face. This mental processing may have activated a mirror neuron system in the brain of our subjects. The mirror neuron system was originally found in the monkey; it was activated when the subjects were imitating the actions of others (Rizzolatti & Fabbri-Destro, 2008; Iacoboni & Dapretto, 2006). Recently, this system has been considered to exist in the human brain and was found to be activated when the subject observed and imitated the facial expression of others (Carr, Iacoboni, Dubeau, Mazziotta, & Lenzi, 2003). Furthermore, the location of the system in the inferior frontal gyrus and temporo-parietal junction (Rizzolatti & Fabbri-Destro, 2008; Iacoboni & Dapretto, 2006) is similar with the location of the region where activation was observed in the present study. A novel finding of the present study was that the neutral face that was previously paired with a negative emotion evoked significant activation even in the absence of reinforcing stimuli.

The STS has been implicated in a polymodal association cortex that combines auditory and visual information (Murase et al., 2008; Calvert, 2001) and may be related to the cross-modal operations of the picture of the face and the sound of the voice in our experiment. In an fMRI study using meaningless and emotional speech (Grandjean

et al., 2005), there was significant activation in the right STS under the angry prosody condition as compared with the neutral prosody condition. The neural responses in the middle part of the right STS correlated with the degree of emotional prosody of adjectives that were spoken (Wildgruber et al., 2006). These results indicate that the right mid-STS is involved in processing the negative emotional intonation of auditory stimuli. What is intriguing in the present study is the fact that the right STS was more predominantly activated in the CS+up condition than in the CS− condition and that neither condition was paired with the voice stimuli. The present results may suggest that reactivation of this region occurred because of previous exposure to the voice stimuli that were associated with negative prosody.

The PPI analysis involving the amygdala and MGB revealed that the degree of correlation between the regions was modulated by the pairing of voice and face stimuli. A critical finding was that there was an interaction effect between functional coupling in the right amygdala–MGB connectivity and the task condition (i.e., CS+p vs. CS+up) in the first run of the Acq phase. There was no such interaction effect between the conditions and the right amygdala–MGB connectivity in the later runs. This suggests that the subcortical pathway might be modulated predominantly at an early stage in the Acq phase. Projections from the MGB to several nuclei of the amygdala have been confirmed by an anatomical study in animals (LeDoux et al., 1990). The pathways through which auditory inputs from the MGB might reach the amygdala, bypassing auditory cortex, may play a pivotal role in the processing of stimuli with emotional significance (LeDoux et al., 1990).

The result that the interaction effect was significant only in the first run suggests that this phenomenon may be brought about by the learning process involved in the CS–US relationship because the interaction effect would be found throughout the five runs if this phenomenon was based exclusively on the perceptual process. Another point was that the increasing amygdala–MGB connectivity was not necessarily associated with the increasing signal in the amygdala during the first run. Furthermore, this phenomenon preceded enhanced response in the amygdala to the CS+up during the second run. These results imply that the degree of signal correlation was critical for modulating amygdala activity, but the degree of signal intensity was not. The neuronal ensemble in the subcortical pathway may facilitate learning of the CS–US relationship in the amygdala for the transfer of emotional information from the CS+p to the CS+up. In a previous neuroimaging study using a conditioning paradigm during extinction, there was a significant interaction between the activity of the subcortical visual pathway and the task condition (Morris, Ohman, & Dolan, 1999). In an fMRI study of emotional expression (Vuilleumier, Armony, Driver, & Dolan, 2003), activation of the subcortical visual pathway through the superior colliculus was involved in the processing of the low-frequency component of a picture of a fearful face.

In line with these studies, the present study demonstrated the relationship between the auditory subcortical route and acquisition of aversive conditioning; however, it should be noted that the presentation of a negative face under the CS+p condition was a confounding factor.

Although the results of the comparison between the CS+ and CS− during the Hab phase did not show significant difference in activation in the hippocampal region ($p = .05$, uncorrected; Supplementary Figure 1), conditioning-related activation during the Ext phase has been found in the right hippocampus–amygdala (Figure 5). This suggests that the differential response observed during the Ext phase was most likely acquired during the Acq phase. The hippocampal responses have been involved in conditioning, particularly when the contextual information was associated with unpleasant stimuli (Alvarez et al., 2008; Hasler et al., 2007). The experimental procedure, where the subjects were instructed to judge whether the face was paired with the aversive voice in the Acq phase, might have utilized contextual memory associated with the US. Other structures involved in the extinction were the orbital gyrus and the fusiform gyrus. Several fMRI studies using conditioning paradigm (Alvarez et al., 2008; Gottfried & Dolan, 2004) indicated that activation in orbito-frontal cortex was associated with the extinction. In the fusiform gyrus, there was a significant difference in activation between the conditions, indicating that the faces associated with an unpleasant voice and expression elicited greater activation than those that were not. The affective significance of a visual item might have lead to enhanced perceptual processing and hemodynamic responses in visual cortex (Padmala & Pessoa, 2008). Although no difference was observed in the activation between the first and second runs, this may be due to the relatively short extinction time in the present experiment.

Some limitations of the current study should be noted. Due to a technical difficulty, we could not record the SCR within the scanner. Therefore, it may not be appropriate to use our SCR data to analyze the data obtained from the fMRI experiment. Combined presentation of a negative expression and negative voice in the CS+p condition would have had a confounding effect on visual and auditory domains that are associated with conditioning. Furthermore, we could not disentangle the semantic and prosodic effects of a negative voice on neural responses in stress-related brain regions. The partial conditioning paradigm adopted in the current study facilitated the segregation of the neural responses into the CS+p and CS+up conditions (Straube et al., 2007; Morris et al., 2001; Buchel et al., 1998, 1999); however, brain responses to the CS+up in the later stage of the Acq phase would be similar with the responses to the CS+ in the early Ext phase. In addition, the response to the CS that was not paired with the US might involve another process such as a negative prediction error. The subjects may have been aware of the partial reinforcement schedule and may have adapted their responses to the omission of the US. The response adaptation learned

during the Acq phase might be responsible for the rapid decrease in amygdala activation observed in the present study.

In conclusion, the SCR study showed that the healthy male subjects were conditioned to the face picture by using an aversive voice as the US, and the fMRI experiment revealed that enhanced amygdala activity was involved in processing the faces that had been paired with the voice. The aversive learning of face–voice coupling was associated with transient neural activation in the right amygdala that was modulated by auditory inputs from the subcortical pathway during the initial stage of the Acq phase. Modulation of the subcortical pathway by the aversive auditory input preceded the transient enhancement of the amygdala activity. These results indicate that in real life, we are automatically conditioned to aversive environmental stimuli, and that the amygdala plays a substantial role in combining perceptual inputs from different sensory modalities and integrates the information to elicit autonomic responses. Although in the present study we emphasized the potential role of the amygdala in negative emotional processing, a recent meta-analysis of neuroimaging literatures indicated that the amygdala responded to all visual emotional stimuli, regardless of their valence (Sergerie, Chochol, & Armony, 2008). This finding supports a novel model that posits a more general role of the amygdala in the detection of biologically and socially relevant information (Sander, Grafman, & Zalla, 2003). Particularly in humans, socially relevant events such as pairing of faces and voices appear to have become the dominant functions of the amygdala.

REFERENCES

Alvarez, R. P., Biggs, A., Chen, G., Pine, D. S., & Grillon, C. (2008). Contextual fear conditioning in humans: Cortical–hippocampal and amygdala contributions. *Journal of Neuroscience, 28,* 6211–6219.

Belmaker, R. H., & Agam, G. (2008). Major depressive disorder. *New England Journal of Medicine, 358,* 55–68.

Birbaumer, N., Veit, R., Lotze, M., Erb, M., Hermann, C., Grodd, W., et al. (2005). Deficient fear conditioning in psychopathy: A functional magnetic resonance imaging study. *Archives of General Psychiatry, 62,* 799–805.

Brett, M., Anton, J., Valabregue, R., & Polini, J. P. (2002). *Region of interest analysis using an SPM toolbox.* Paper presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Buchel, C., Dolan, R. J., Armony, J. L., & Friston, K. J. (1999). Amygdala–hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *Journal of Neuroscience, 19,* 10869–10876.

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.

Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex, 11,* 1110–1123.

Carr, L., Iacoboni, M., Dubeau, M. C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences, U.S.A., 100,* 5497–5502.

Carter, R. M., O'Doherty, J. P., Seymour, B., Koch, C., & Dolan, R. J. (2006). Contingency awareness in human aversive conditioning involves the middle frontal gyrus. *Neuroimage, 29,* 1007–1012.

Cheng, D. T., Knight, D. C., Smith, C. N., Stein, E. A., & Helmstetter, F. J. (2003). Functional MRI of human amygdala activity during Pavlovian fear conditioning: Stimulus processing versus response expression. *Behavioral Neuroscience, 117,* 3–10.

Dunsmoor, J. E., Bandettini, P. A., & Knight, D. C. (2008). Neural correlates of unconditioned response diminution during Pavlovian conditioning. *Neuroimage, 40,* 811–817.

Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6,* 218–229.

Glascher, J., & Buchel, C. (2005). Formal learning theory dissociates brain regions with different temporal integration. *Neuron, 47,* 295–306.

Gottfried, J. A., & Dolan, R. J. (2004). Human orbitofrontal cortex mediates extinction learning while accessing conditioned representations of value. *Nature Neuroscience, 7,* 1144–1152.

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., et al. (2005). The voices of wrath: Brain responses to angry prosody in meaningless speech. *Nature Neuroscience, 8,* 145–146.

Hasler, G., Fromm, S., Alvarez, R. P., Luckenbaugh, D. A., Drevets, W. C., & Grillon, C. (2007). Cerebral blood flow in immediate and sustained anxiety. *Journal of Neuroscience, 27,* 6313–6319.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience, 7,* 942–951.

Ishai, A., Pessoa, L., Bikle, P. C., & Ungerleider, L. G. (2004). Repetition suppression of faces is modulated by emotion. *Proceedings of the National Academy of Sciences, U.S.A., 101,* 9827–9832.

Jensen, J., McIntosh, A. R., Crawley, A. P., Mikulis, D. J., Remington, G., & Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron, 40,* 1251–1257.

Kalisch, R., Korenfeld, E., Stephan, K. E., Weiskopf, N., Seymour, B., & Dolan, R. J. (2006). Context-dependent human extinction memory is mediated by a ventromedial prefrontal and hippocampal network. *Journal of Neuroscience, 26,* 9503–9511.

Knight, D. C., Cheng, D. T., Smith, C. N., Stein, E. A., & Helmstetter, F. J. (2004). Neural substrates mediating human delay and trace fear conditioning. *Journal of Neuroscience, 24,* 218–228.

Knight, D. C., Nguyen, H. T., & Bandettini, P. A. (2005). The role of the human amygdala in the production of conditioned fear responses. *Neuroimage, 26,* 1193–1200.

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 fMRI study. *Neuron, 20,* 937–945.

Leal-Campanario, R., Fairen, A., Delgado-Garcia, J. M., & Gruart, A. (2007). Electrical stimulation of the rostral medial prefrontal cortex in rabbits inhibits the expression of conditioned eyelid responses but not their acquisition. *Proceedings of the National Academy of Sciences, U.S.A., 104,* 11459–11464.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience, 23,* 155–184.

LeDoux, J. E., Farb, C., & Ruggiero, D. A. (1990). Topographic organization of neurons in the acoustic thalamus that project to the amygdala. *Journal of Neuroscience, 10,* 1043–1054.

Milad, M. R., Goldstein, J. M., Orr, S. P., Wedig, M. M., Klibanski, A., Pitman, R. K., et al. (2006). Fear conditioning and extinction: Influence of sex and menstrual cycle in healthy humans. *Behavioral Neuroscience, 120,* 1196–1203.

Milad, M. R., Rauch, S. L., Pitman, R. K., & Quirk, G. J. (2006). Fear extinction in rats: Implications for human brain imaging and anxiety disorders. *Biological Psychology, 73,* 61–71.

Morris, J. S., Buchel, C., & Dolan, R. J. (2001). Parallel neural responses in amygdala subregions and sensory cortex during implicit fear conditioning. *Neuroimage, 13,* 1044–1052.

Morris, J. S., & Dolan, R. J. (2004). Dissociable amygdala and orbitofrontal responses during reversal fear conditioning. *Neuroimage, 22,* 372–380.

Morris, J. S., Ohman, A., & Dolan, R. J. (1998). Conscious and unconscious emotional 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. *Proceedings of the National Academy of Sciences, U.S.A., 96,* 1680–1685.

Murase, M., Saito, D. N., Kochiyama, T., Tanabe, H. C., Tanaka, S., Harada, T., et al. (2008). Cross-modal integration during vowel identification in audiovisual speech: A functional magnetic resonance imaging study. *Neuroscience Letters, 434,* 71–76.

Neufang, S., Fink, G. R., Herpertz-Dahlmann, B., Willmes, K., & Konrad, K. (2008). Developmental changes in neural activation and psychophysiological interaction patterns of brain regions associated with interference control and time perception. *Neuroimage, 43,* 399–409.

Neumann, D. L., & Waters, A. M. (2006). The use of an unpleasant sound as an unconditional stimulus in a human aversive Pavlovian conditioning procedure. *Biological Psychology, 73,* 175–185.

Nitschke, J. B., Sarinopoulos, I., Mackiewicz, K. L., Schaefer, H. S., & Davidson, R. J. (2006). Functional neuroanatomy of aversion and its anticipation. *Neuroimage, 29,* 106–116.

Ollendick, T. H., & Hirshfeld-Becker, D. R. (2002). The developmental psychopathology of social anxiety disorder. *Biological Psychiatry, 51,* 44–58.

Padmala, S., & Pessoa, L. (2008). Affective learning enhances visual detection and responses in primary visual cortex. *Journal of Neuroscience, 28,* 6202–6210.

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: Role of the amygdala and vmPFC. *Neuron, 43,* 897–905.

Phillips, M. L., Medford, N., Young, A. W., Williams, L., Williams, S. C., Bullmore, E. T., et al. (2001). Time courses of left and right amygdalar responses to fearful facial expressions. *Human Brain Mapping, 12,* 193–202.

Quirk, G. J., Armony, J. L., & LeDoux, J. E. (1997). Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. *Neuron, 19,* 613–624.

Quirk, G. J., Likhtik, E., Pelletier, J. G., & Pare, D. (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *Journal of Neuroscience, 23,* 8800–8807.

Rizzolatti, G., & Fabbri-Destro, M. (2008). The mirror system and its role in social cognition. *Current Opinion in Neurobiology, 18,* 179–184.

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12,* 191–200.

Sander, D., Grafman, J., & Zalla, T. (2003). The human amygdala: An evolved system for relevance detection. *Reviews in the Neurosciences, 14,* 303–316.

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences, 10,* 24–30.

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 Biobehavioral Reviews, 32,* 811–830.

Sotres-Bayon, F., Bush, D. E., & LeDoux, J. E. (2004). Emotional perseveration: An update on prefrontal–amygdala interactions in fear extinction. *Learning & Memory, 11,* 525–535.

Straube, T., Weiss, T., Mentzel, H. J., & Miltner, W. H. (2007). Time course of amygdala activation during aversive conditioning depends on attention. *Neuroimage, 34,* 462–469.

Thiel, C. M., Friston, K. J., & Dolan, R. J. (2002). Cholinergic modulation of experience-dependent plasticity in human auditory cortex. *Neuron, 35,* 567–574.

Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience, 6,* 624–631.

Wildgruber, D., Ackermann, H., Kreifelts, B., & Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research, 156,* 249–268.

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.