

Forming a negative impression of another person correlates with activation in medial prefrontal cortex and amygdala

Tetsuya Iidaka,¹ Tokiko Harada,¹ and Norihiro Sadato²

¹Department of Psychiatry, Graduate School of Medicine, Nagoya University, Nagoya, and ²Division of Cerebral Integration, National Institute for Physiological Sciences, Okazaki, Japan

Neural correlates involved in the formation of negative impression from face were investigated using event-related functional magnetic resonance imaging and a partial conditioning paradigm. Eighteen normal volunteers underwent imaging while they viewed the faces of two unfamiliar individuals: one individual's face was partially accompanied by negative emotion but the other's was not. After the volunteers learned the relationship between the faces and the emotion, they formed a more negative impression of the person's face when the emotion was presented. Subtraction analysis of the individuals' neutral faces revealed activation in the dorsal anterior cingulate cortex and superior temporal sulcus, but this activity did not correlate with the change of impression from face. On the other hand, the response in the left amygdala negatively correlated with the change of impression from face in the first run. Time modulation analysis revealed that activity in the dorsomedial prefrontal cortex associated with negative emotion was the largest in the initial part of the acquisition. These results suggest that a negative impression from face may be formed by orchestrated activity in the dorsomedial prefrontal cortex, dorsal anterior cingulate cortex and amygdala, and that the activity has a prominent role in the initial acquisition of negative emotion.

Keywords: fMRI; conditioning; superior temporal sulcus; social brain; expression

INTRODUCTION

The recent research on the mechanisms of emotion within the brain has largely been focused on measurement of the neural activity in the prefrontal cortex (PFC) and amygdala (LeDoux, 2000; Maren and Quirk, 2004). Functional neuroimaging studies have shown that the medial PFC and anterior cingulate cortex (ACC) are activated when emotion is induced in subjects under task conditions, but this activity is not specific to a particular kind of emotion (Phan *et al.*, 2002; Steele and Lawrie, 2004). The amygdala is considered to play a pivotal role in the perception of emotion and in the interaction between the central nervous system and the autonomic–hormonal system. A meta-analysis of neuroimaging studies on amygdala activation showed that the emotional characteristics of a face are associated with greater activation than other kinds of stimuli (Costafreda *et al.*, 2008). In these studies, however, the experimental stimuli presented to the subjects, such as a fearful face and a violent scene, inherently had emotional valence and the neural responses to the

change of valence from neutral to negative through the experiment were rarely investigated.

The emotional or non-verbal behavior of a person has a strong influence on the impression formation of that person even when the behavior is observed only briefly (Ambady, 1993). In this case, the face is perhaps the most important signal in social interactions, because we can infer two types of information immediately after seeing a face: identity and expressions (Posamentier and Abdi, 2003). In a functional magnetic resonance imaging (fMRI) study to investigate brain responses during the evaluation of facial trustworthiness, a significant involvement of the amygdala was observed during the task (Winston *et al.*, 2002). Emotional expressions on a person's face provide information about their emotional state (Haxby *et al.*, 2002) and can evoke similar feelings in the observer (Wild *et al.*, 2001). The ability to form an impression of others is considered to be a critical skill in human communication under social situations (Ambady, 1993). Recently, the neural processes involved in the formation of impression from face have been studied by using functional neuroimaging (Mitchell *et al.*, 2004; Schiller *et al.*, 2009). In these studies, the medial PFC, amygdala and posterior cingulate cortex were activated when a face was paired with sentences describing a personality (Mitchell *et al.*, 2004) or social profile (Schiller *et al.*, 2009). An fMRI study relevant to the present study used a classical conditioning paradigm and showed that activity in the amygdala and fusiform gyrus was found to be associated

Received 27 December 2009; Accepted 29 June 2010

Advance Access publication 6 August 2010

This study was supported by a Grant-in-Aid for Scientific Research on Innovative Areas, 'Face perception and recognition' and a Grant-in-Aid for Academic Frontier Project for Private Universities by the Japanese Ministry of Education, Science, Sports and Culture. We thank anonymous reviewers for critical comments on an earlier version of the manuscript.

Correspondence should be addressed to Tetsuya Iidaka, Department of Psychiatry, Graduate School of Medicine, Nagoya University, 65 Tsurumai, Showa, Nagoya, Aichi, 466-8550, Japan.
E-mail: iidaka@med.nagoya-u.ac.jp

with the learning of the affective value of a face (Petrovic *et al.*, 2008). A unique aspect of the present study is that a study by Petrovic *et al.* (2008) used primarily negative stimuli (i.e. electric shock) whereas we used social and non-descriptive information (negative emotional face and voice) as stimuli in the present study.

The aim of the present study was to examine the neural responses involved in the formation of a negative impression from face that was previously associated with neutral valence. To do this, we investigated the effect of a negative emotion expressed by vocal and facial stimuli on the impression of a neutral face of the same individual. The differences in the neural responses between the face with continuously neutral valence and the face that formed the negative impression were compared by using event-related fMRI and normal subjects. For this purpose, we used an aversive conditioning paradigm in which a previous neutral stimulus (conditioned stimulus or CS) elicits behavioral and autonomic responses after effective pairing with an unpleasant stimulus (unconditioned stimulus or US). The amygdala, ACC and medial PFC are involved in the classical conditioning of animals, in which the association between neutral and noxious stimuli is learned during the experiment (Fendt and Fanselow 1999; Gewirtz and Davis 2000; Maren 2001). A similar neural network has been shown to be involved in the conditioning of human subjects in functional neuroimaging experiments (Sehlmeyer *et al.*, 2009). Therefore, we predicted that the amygdala, ACC and medial PFC are associated with the formation or expression of impression from face when a conditioned face is compared with a control face.

However, it is methodologically difficult to segregate the effects of a CS and an US on brain activity, because the neural responses to these stimuli overlap substantially due to slow hemodynamic responses measured by fMRI. A partial conditioning paradigm in which half of a CS is combined with an aversive US, and the other half is not, would be useful to exclude the confounding effect. Such a paradigm has already been adopted in previous neuroimaging studies (Buchel *et al.*, 1998, 1999; Morris *et al.*, 2001; Petrovic *et al.*, 2008), in which either a neutral face (Buchel *et al.*, 1998; Petrovic *et al.*, 2008) or a neutral tone (Buchel *et al.*, 1999) was used as the CS. In an fMRI study by Pollak *et al.* (2010), the authors used auditory presentation of human scream as an US and showed a significant involvement of the amygdala in conditioning. A unique aspect of the present study is that a voice and a face expressing a negative emotion were used as the US. In a previous study, this paradigm successfully elicited conditioned responses in normal subjects, as measured by the skin conductance response (Iidaka *et al.*, 2010). We thought that the combination of a face and a voice with negative emotional valence would simulate real-life situations, in which impression from face is modulated by the observed behavior. Thus, in the present study, we measured the neural activation in response to the formation of impression from face in subjects who viewed a person showing a

negative emotional attitude by using fMRI and a partial conditioning paradigm. In particular, the degree of activation in the amygdala would be correlated with the change of impression, because the amygdala plays a modulatory role in human behavior and decision making (Seymour and Dolan, 2008).

METHODS

Subjects

Eighteen healthy volunteers (12 males and 6 females, all right handed, mean age \pm s.d. = 20.7 \pm 1.9 years) were recruited for the fMRI study. The subjects provided written informed consent for the procedure. An experimenter carefully checked their physical and mental conditions according to the standard institutional protocol. This study was approved by the ethics committee of the National Institute for Physiological Sciences.

Experimental stimuli

Selection of face pictures

Pictures of facial expressions were selected from a database of Japanese male face pictures. The faces of three individuals with neutral or negative (angry or disgusted) expressions were used as the experimental stimuli. We chose an angry or disgusted face as a stimulus rather than a fearful face in order to simulate a social situation wherein the subject is exposed to the emotional expression of another person. In the case of angry or disgusted expressions, the negative emotion is directed from the individual in the picture to the subject. In contrast, a picture of a fearful face would depict a situation wherein the individual in the picture is exposed to a threat, which did not meet our purpose. The pictures were digitized and equalized in terms of shape and luminance by using commercial software (Adobe Photoshop; Adobe Systems, Inc., San Jose, CA, USA).

In a pilot study, 12 healthy subjects (mean age = 24 years) who did not undergo fMRI were presented with these pictures. The subjects classified the expression in each picture by using six basic emotions and a neutral emotion, and rated the intensity (from 1 for neutral to 7 for extremely negative). The mean proportion of correct responses was 89% for both the neutral and the negative faces. Disgust accounted for 50% of the emotions labeled as negative, anger accounted for 39% and sadness for 11%. The results of the intensity rating showed that the mean (\pm s.d.) intensity was 0.16 (\pm 0.5) and 4.25 (\pm 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 the negative faces [$F(1,66) = 387$, $P < 0.001$], and there was no significant main effect of the stimuli [$F(2,66) = 0.21$, $P = 0.8$] or interaction of the emotion and stimuli [$F(2,66) = 1.5$, $P = 0.22$]. The mean intensity of each emotion was 4.2 (\pm 1.3) for disgust, 4.6 (\pm 0.6) for anger and 3.5 (\pm 1.3) for sadness. Thus, in the present study, the

neural substrate involved in the perception of a negative face was considered to be mainly a reflection of the combination of perceiving angry and disgusted faces. Several behavioral studies have shown that, among the six basic expressions, anger and disgust are perceived to be similar and the results of a recognition task for these two expressions are considerably intermixed (Adolphs *et al.*, 1999). Therefore, it is reasonable to assume that the perception of these two expressions share common neural substrates.

Combination with unconditioned stimulus

Two of the three faces were chosen for each subject in a counterbalanced manner to create a face–voice stimulus pair. The faces were assigned to either a CS⁺ or a CS⁻ (Figure 1A). During the experiment, each subject viewed pictures of the individuals (the CS⁺ individual and the CS⁻ individual). The CS⁺ was a conditioned stimulus that was paired with the US for half of the presentation during the acquisition (Acq) phase, whereas the CS⁻ was never paired with the US during the experiment. Every stimulus was presented for a duration of 1500 ms. The CS⁺ was further classified as CS⁺unpaired (CS⁺up) and CS⁺ paired (CS⁺p) during the Acq phase. The CS⁺up involved the presentation of only the neutral face of the CS⁺ individual for a duration of 1500 ms. The CS⁺p was a combined presentation of the face and voice, and its emotional valence changed from neutral to negative during the presentation. The CS⁺ presented during the habituation (Hab) phase was identical to the CS⁺up presented during the Acq phase.

During the Acq phase, the face was superimposed on a background picture depicting an office scene (e.g. tables, chairs and computers; Figure 1B) to investigate the effect of the conditioning on contextual information. Two pictures (B⁺ and B⁻) were chosen from the Microsoft Clip Art database (Microsoft Corporation, USA) and randomly assigned to the CS⁺ and CS⁻ faces (e.g. CS⁺ was shown on B⁺ and CS⁻ was shown on B⁻). During the Hab phase, the subjects viewed the same stimuli that were to be presented during the Acq and context (CON) phases to reduce their novelty effect. The presentation of the background pictures was fully counterbalanced across the experimental conditions and across the subjects. Therefore, there was no confounding effect on the results obtained during the Acq phase, although we have not reported the results regarding the background pictures in this article.

Unconditioned stimulus: emotional voice

A vocal 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 by using commercial audiovisual software (Symphomovie; Epson Corporation, Suwa, Japan). The time course of the audiovisual stimuli began with the presentation of the neutral face for 500 ms followed by simultaneous presentation

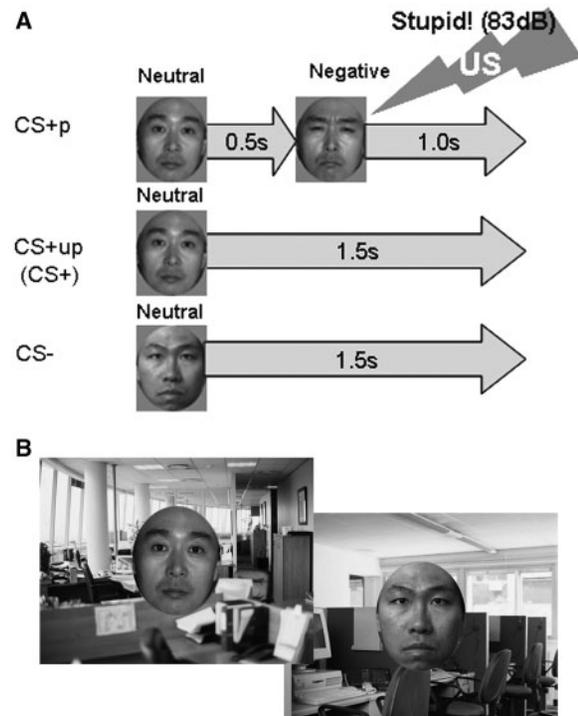


Fig. 1 (A) Schematic illustration of the experimental stimuli used in the study. Two faces from among three individuals were randomly selected for each subject. They were assigned to either the CS⁺ or the CS⁻. The US was a male voice saying ‘stupid’ with a sound pressure of 83 dB. The time course of the CS⁺p began with the presentation of the neutral face of the CS⁺ for 0.5 s followed by simultaneous presentation of the negative face of the CS⁺ and the US for 1 s. The CS⁺up presented during the acquisition (Acq) phase was identical to the CS⁺ presented during the habituation (Hab) phase. Both the CS⁺up and the CS⁺ were presentations of only the neutral face for a duration of 1.5 s. The CS⁻ represented the stimulus with the presentation of a neutral face of another individual, and it was never paired with the vocal stimulus. (B) Examples of the combinations of the facial and background pictures presented during the Hab and Acq phases.

of a negative face of the same individual and a negative voice for 1000 ms (Figure 1A; CS⁺p). The facial expression changed from neutral to negative in each aversive event, because a change of facial expression is usually associated with an emotional voice. The stimuli were converted to MPEG format and presented to the subjects by using the Presentation software (Neurobehavioral Systems Inc., Albany, CA, USA). The sound stimulus was delivered with a monaural headphone in the scanner, with a sound pressure of ~83 dB. All subjects correctly received the vocal 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 attached to the head coil of the scanner. The response was measured by using a magnetic compatible button box held in the subject’s right hand. The subjective response rate and reaction time were entered into a two-way repeated-measures ANOVA by using the conditioning (CS⁺ and CS⁻) and the phase (Hab and Acq) as factors.

Experimental procedure

The fMRI experiment consisted of the Hab (one run), Acq (four runs) and Con (one run) phases. During the Hab phase, two combined pictures and two background pictures were repeatedly and randomly presented one at a time [duration = 1.5 s, interstimulus interval (ISI) = 3.5 s] for 10 times each. The subjects pressed the left button when they saw a picture. After the Hab phase, the subjects rated their impression from faces of the CS⁺ and CS⁻ individuals by using the 7-point Likert scale (i.e. 1 for the least likeable and seven for the most likeable). During the Acq phase, the faces seen during the Hab phase (CS⁺ and CS⁻) were repeatedly presented one at a time for a duration of 1.5 s and an ISI of 13.5 s. A long ISI was chosen to segregate the responses to the US and CS for effective tracking of the hemodynamic response. Across four runs, the CS⁺ face was presented 64 times (16 times in each run). The CS⁺ was paired with the vocal stimulus (CS⁺p) for half of this presentation, whereas it was not paired for the other half (CS⁺up). The CS⁻ face was presented 64 times (16 times in each run) but never with the voice stimulus. Each face was shown against a background scene. The subjects pressed the left button whenever they saw the face, and no further judgment was required. The subjects again rated their impression from faces of the CS⁺ and CS⁻ individuals by using the 7-point Likert scale. Finally, in the Con phase, two background pictures (B⁺ and B⁻) were randomly presented 32 times each. A two-way repeated-measures ANOVA was conducted for the Likert scale scores by using the conditioning (CS⁺ and CS⁻) and the phase (Hab and Acq) as factors (statistical threshold was set at $P=0.05$).

Image acquisition and preprocessing

Functional images of the whole brain were obtained in an axial-oblique position by using a 3-T MRI scanner (Allegra; Siemens, Erlangen, Germany) equipped with single-shot echo planar imaging (EPI; TR = 2.3 s, TE = 30 ms, flip angle = 80°, 64 × 64 matrix and 36 slices, voxel size = 3 mm × 3 mm × 3 mm), and it was sensitive to blood oxygen level-dependent contrast. The number of images obtained was 93 in the Hab phase, 856 (across four runs) in the Acq phase and 284 in the Con phase. After discarding the first four images, the successive images in each run were subjected to the analysis. A high-resolution anatomical T1-weighted image was also acquired (MPRAGE; TR = 2.5 s, TE = 4.38 ms, flip angle = 8°, 256 × 256 matrix and 192 slices, voxel size = 0.75 mm × 0.75 mm × 1 mm) for each subject. The data were analyzed by using SPM5 (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 by using 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 subject to the EPI template image. The normalized images were spatially smoothed with an 8-mm Gaussian kernel.

First level analysis

Following preprocessing, statistical analysis of the data obtained from each subject was conducted by using a 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 s) were applied to the time series data. An AR(1) model was used to estimate the temporal autocorrelation. 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⁺ (with B⁺), CS⁻ (with B⁻), B⁺ and B⁻; the Acq phase, CS⁺p (with B⁺), CS⁺up (with B⁺) and CS⁻ (with B⁻); the Con phase, B⁺ and B⁻. Because the number of events with the CS⁺p and CS⁺up was half of that with the CS⁻ during the Acq phase, half of the events with the CS⁻ were randomly chosen to obtain statistical power for computation of a significant difference between the conditions. In each run, another condition involving incorrect responses was included. The US was not explicitly modeled but included in the CS⁺p condition because the onset of the vocal stimulus was delayed for only 500 ms from the onset of the CS. In addition, six movement parameters obtained during the realignment were entered as regressors. The parameter estimates for each condition and for the difference between the conditions were computed from the least mean square fit of the model to the time series data. Images of the parameter estimates representing event-related activity at each voxel were created for each condition and subject. The neural responses obtained during the Hab and Con phases were analyzed, but the data are not reported in this article.

Second level analysis

During the Acq phase, a contrast of the CS⁺up vs the CS⁻ was created in each subject. This is a critical contrast to investigate whether the CS⁺up, even when unaccompanied by the emotional stimulus, elicited differential activation compared with the CS⁻. At the second level, the results from 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 $P=0.001$, uncorrected for multiple comparisons for height and $k=20$ voxels for spatial extent.

We hypothesized that the difference in amygdala activation between the CS⁺up and the CS⁻ is correlated with the subjective impression from face. To test this, we computed the subtraction between the CS⁺p and CS⁺up to investigate the effect of negative emotion on amygdala activity. The result of the subtraction was used as a localizer of the amygdala where effect of negative emotional stimuli on neural activity was observed. The difference in parameter estimates

between the CS⁺up and the CS⁻ in the amygdala were extracted from the spherical ROI ($r=8$ mm) at the bilateral amygdalae by using the VOI function of SPM5. The value was correlated with the difference in the Likert scale scores for the CS⁺ face between the Hab and the Acq phases (score for CS⁺ after the Hab phase minus score for CS⁺ after the Acq phase, a positive value indicates that the subject formed a more negative impression of the person's face). The correlation analysis was conducted by collapsing four runs and separately in each run to investigate time dependent effect on brain-behavior relationship (Spearman rank order correlation, threshold set at $P<0.05$).

Time-modulation analysis

In the time modulation analysis of the CS⁺up vs the CS⁻, we hypothesized that the difference in activation between the CS⁺up and the CS⁻ would be modulated across the four runs during the Acq phase. This analysis was carried out to examine the change of responses during the Acq phase via neural activity to the CS and the manner in which they are modulated with learning across the runs. Although it is not possible to track neuronal responses with fMRI, it has been shown that the responses of an activated cluster change across experimental runs when emotional faces are presented to subjects (Phillips *et al.*, 2001). Therefore, it is possible to investigate how the CS-US relationship is established and how the neural responses are regulated across runs. At the first level analysis in each subject, we entered the contrast (2, 1, -1, -2) for the CS⁺up and that (-2, -1, 1, 2) for the CS⁻ (the numbers represent the contrast in each of the four

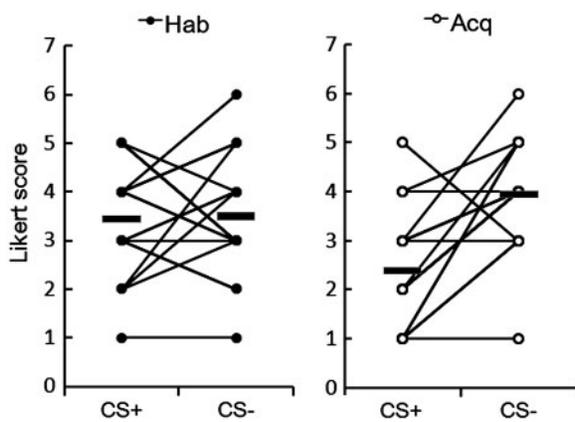


Fig. 2 The Likert scale scores (y -axis) obtained after the habituation (Hab, filled circle) and acquisition (Acq, open circle) phases for the face that was paired with a negative emotion (CS⁺) and the face that was never paired with such an emotion (CS⁻) are plotted. Horizontal bar indicates the mean of 18 subjects. Each circle represents each subject; however, two or more subjects are overlapping in the figure. Line connects the score of the same subject for CS⁺ and CS⁻. The Likert scores ranged from 1 (the least likeable) to 7 (the most likeable). A two-way ANOVA with the phases (Hab and Acq) and conditions (CS⁺ and CS⁻) as factors revealed a significant interaction effect ($P<0.01$). The results of the *post hoc* paired t -test showed a significant difference in the scores between the CS⁺ and the CS⁻ after the Acq phase ($P<0.01$).

runs). The analysis investigated the region in which the neural activity measured from the baseline condition decreases for the CS⁺up and increases for the CS⁻ across time. At the second level analysis, the statistical threshold was set at $P=0.001$, uncorrected for height and at $k=20$ voxels for spatial extent.

The changes in the Likert scale scores obtained after the Hab and Acq phases are plotted in Figure 2. A comparison of the neural responses to the CS⁺p and CS⁺up during the Acq phase is shown on the T1 template image (Figure 3, top image). The mean parameter estimate extracted from the spherical ROI at the left amygdala was correlated with the change in the Likert scale scores and plotted (Figure 3, bottom image). In Figure 4, the results of the subtraction analysis of the CS⁺up and CS⁻ are superimposed on the T1 template image. The result of the time modulation effects on the CS⁺up minus the CS⁻ is shown in Figure 5. The mean

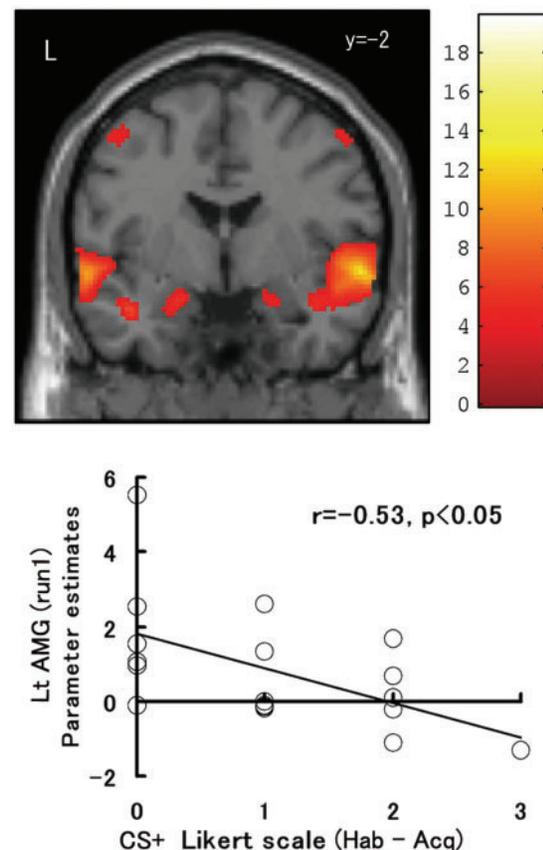


Fig. 3 Top: there was significant activation in the bilateral amygdalae in the subtraction analysis of the CS⁺p and CS⁺up. The cluster was superimposed on the T1 template image of SPM5. The threshold was set at $P=0.001$ (uncorrected) for height and $k=20$ voxels for spatial extent. Bottom: the activity in the left amygdala during the first run of the acquisition phase (run 1, CS⁺up minus CS⁻, y -axis) extracted from the region delineated by the above subtraction showed a significant negative correlation with the change of impression from face (CS⁺, Hab minus Acq, x -axis). The subjects with a high value on the x -axis had a bad impression of the face that was paired with the negative emotion and reduced amygdala activity. A regression line, correlation coefficient, and P -value are shown in the figure.

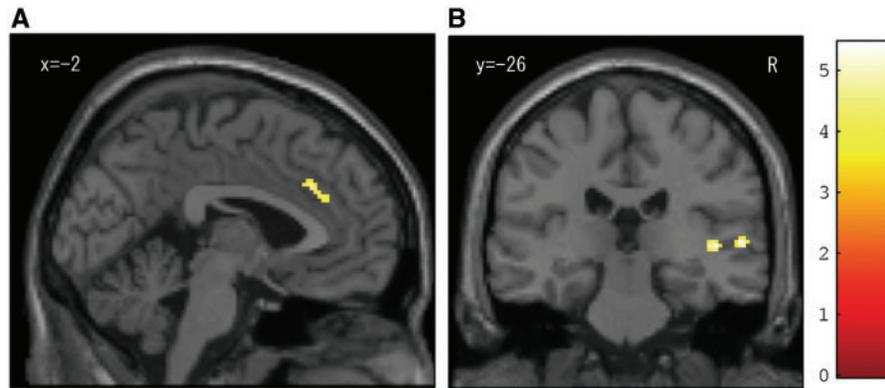


Fig. 4 A significant difference in activation between the CS⁺up and the CS⁻ was found in the dorsal anterior cingulate cortex during the acquisition phase (A). Two regions were located in the superior temporal sulcus (B). The clusters were superimposed on the T1 template image of SPM5. The threshold was set at $P=0.001$ (uncorrected) for height and $k=20$ voxels for spatial extent. Detailed information on the clusters is given in Table 1.

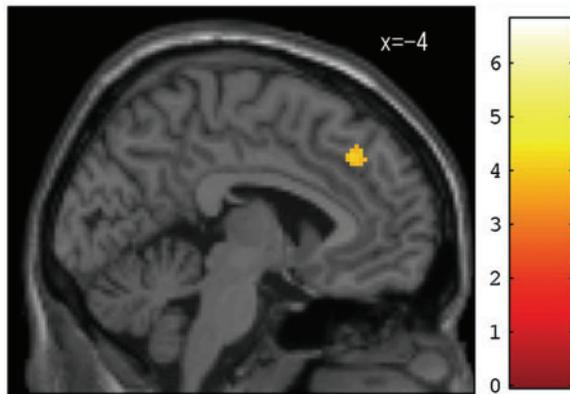


Fig. 5 The result of the time modulation analysis is superimposed on the T1 template image. The activity in the dorsomedial prefrontal cortex ($x, y, z = -4, 34, 40$) adjacent to the dorsal anterior cingulate cortex showed a time-by-condition interaction effect. Changes in the parameter estimates extracted from this cluster separately in each run and each condition are shown in Figure 6. Detailed information on this cluster is given in Table 1.

and standard error (s.e.) of the parameter estimates in the dorsomedial PFC and amygdala are plotted in Figure 6. The parameter estimates were extracted from the second level analysis conducted separately in each condition (CS⁺p, CS⁺up, and CS⁻) and each run (from run 1 to 4). In Table 1, the voxel size, T -value, coordinates, Brodmann's area and region name of the significant clusters are tabulated.

RESULTS

Behavioral data

During the Acq phase, the mean (\pm s.d.) reaction time did not differ significantly between the conditions [CS⁺p, 399 ± 119 ms; CS⁺up, 393 ± 117 ms; CS⁻, 397 ± 120 ms; $F(2,34) = 0.46$, $P = 0.62$]. The mean response rate was 96% and also did not differ significantly between the conditions [$F(2,34) = 0.58$, $P = 0.56$, see supplementary table]. In Figure 2, changes of the Likert scale score are plotted. The

two-way ANOVA for the mean Likert scale scores of the CS⁺ and CS⁻ individuals obtained after the Hab and Acq phases revealed a significant interaction effect [$F(1,17) = 17.0$, $P = 0.01$]. The *post hoc* paired t -test (two-tailed) showed that the mean scale scores for each face were equivalent after the Hab phase ($P = 0.87$) and impression from face of the CS⁺ was significantly worse than that of the CS⁻ facial impression after the Acq phase ($P < 0.01$).

fMRI data

The subtraction analysis of the CS⁺p and CS⁺up showed significant activation in the bilateral amygdalae (FDR corrected at $P = 0.05$ in the whole brain; left: $x, y, z = -20, -2, -18$, $T = 5.45$; right: $x, y, z = 22, -4, -18$, $T = 4.63$; Figure 3, top image). The correlation analysis of the signal in the left amygdala extracted from the subtraction between the CS⁺up minus CS⁻ and the change in the Likert scale scores revealed a negative relationship between the neural activity and impression from face (Figure 3, bottom image). The more negative the subject's impression (larger change from the Hab phase to Acq phase), the lower the amygdala activation. The significant correlation was observed when the four runs were collapsed ($r = -0.49$, $P < 0.05$). A separate analysis for each run showed a significant correlation only in the first run (run 1: $r = -0.53$, $P < 0.05$; run 2: $r = 0.35$, n.s.; run 3: $r = -0.02$, n.s.; run 4: $r = -0.22$, n.s.). There was no significant correlation between the scores and the activity in the right amygdala.

The results of the subtraction analysis of the CS⁺up and CS⁻ revealed significant activation in the dorsal ACC adjacent to the dorsomedial PFC during the formation of a negative impression from face (Table 1, Figure 4, left). Two clusters were located in the right STS (Figure 4, right). The change in the Likert scale scores did not correlate with the activity in these clusters (STS, $r = -0.17$ and 0.02 , n.s., dorsal ACC, $r = 0.12$, n.s.). In the time modulation analysis, the activity in the dorsomedial PFC (Figure 5 and Table 1) showed a significant time-by-condition interaction effect.

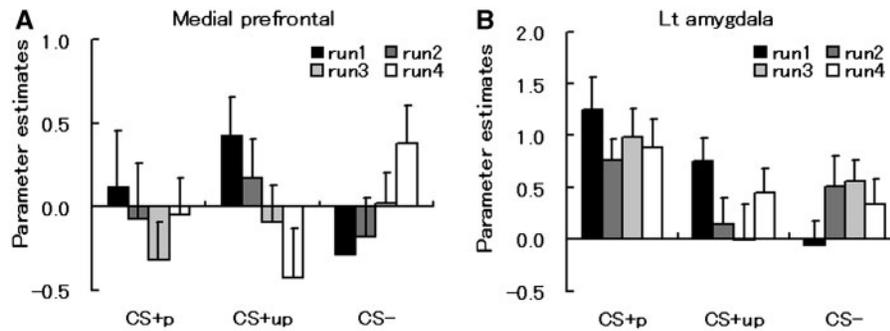


Fig. 6 (A) the mean (column) and s.e. (bar) of the parameter estimates extracted from the dorsomedial prefrontal cluster ($x, y, z = -4, 34, 40$) are plotted separately in each run and each condition. The time (from run 1 to 4)-by-condition (CS^{+up} and CS^{-}) interaction effect indicates that the activity for the CS^{+up} decreased and that for the CS^{-} increased from run 1 to 4 and (B) the mean (column) and s.e. (bar) of the parameter estimates extracted from the left amygdala ($x, y, z = -20, -2, -18$) are plotted separately in each run and each condition. The activity was higher for the CS^{+p} condition than for the CS^{+up} condition, but there was no time-by-condition interaction effect.

Table 1. Significant results of subtraction and time modulation analysis

Region name	L/R	BA	x, y, z	T	Voxels
Subtraction between CS^{+up} and CS^{-}					
Superior temporal sulcus	R	22	46, -26, 2	5.45	25
Superior temporal sulcus	R	22	58, -26, 4	5.33	27
Anterior cingulate cortex	L	32	-2, 28, 32	4.72	54
Time modulation effect on CS^{+up} minus CS^{-}					
Middle temporal gyrus	L	21	-58, -32, 0	6.79	38
Temporal pole	R	20	52, 14, -36	6.09	251
Angular gyrus	L	39	-48, -58, 26	5.61	45
Superior frontal gyrus	L	8	-12, 56, 38	5.53	34
Medial prefrontal cortex	L	8	-4, 34, 40	4.12	46

L/R, left/right; BA, Brodmann's area. Statistical threshold was set at $P = 0.001$ (uncorrected) for the height and $k = 20$ voxels for extent.

The signal in the dorsomedial PFC was the greatest in response to the CS^{+up} in the first run and the activity decreased afterward, but the response to the CS^{-} increased through the experiment (Figure 6, left). These activities in the dorsomedial PFC did not show significant correlation with the change in the Likert scale score. The signal in the left amygdala was higher in the CS^{+p} condition than in the CS^{+up} and CS^{-} conditions (Figure 6, right), but there was no time modulation effect. In particular, the difference in amygdala activity between the CS^{+p} and CS^{-} was larger in the first run than in the later runs; indicating that the amygdala has a prominent role in an initial phase of acquisition of negative information. Finally, the activity in the regions noted in the results section (peak voxels listed in Table 1 and in the bilateral amygdalae) did not show differential responses between the CS^{+} and the CS^{-} during the Hab phase at a lenient threshold of $P = 0.05$ (uncorrected for multiple comparisons).

DISCUSSION

In this study, impression from face was successfully manipulated at the behavioral level, and the change in brain activity

was monitored by using fMRI and a partial conditioning paradigm. When the subjects were frequently exposed to the negative emotional voice and expression, they formed a more negative impression of the person's face, compared with when they were not exposed such an emotion. In the fMRI experiment, the dorsomedial PFC, dorsal ACC and STS were involved in the formation or expression of the negative impression; however, the activity in these regions did not correlate with the change in the subjective rating of impression from face. On the other hand, in accordance with a prior hypothesis, the activity in the left amygdala was significantly and negatively correlated with the change in the subjective rating of impression from face.

The human face plays a key role in signaling social intentions and has a substantial influence on behaviors toward another person and interactive decision making (van't Wout and Sanfey, 2008). A brief presentation of non-verbal behavior by using video clips has been shown to have a strong influence on the formation of an impression (Ambady, 1993). In another behavioral study, a 100 ms exposure to a facial picture with a neutral expression was sufficient to make a highly confident trait judgment (e.g. likeability, trustworthiness and competence) on the face (Willis and Todorov, 2006). These judgments are based on facial features that resemble emotional expressions ordinarily signaling approach and avoidance behaviors (Todorov, 2008). In particular, angry faces that convey approach behavior by the expresser trigger avoidance responses in the observer (Adams and Ambady, 2006). However, only few studies have focused on the neural processes involving the changes of impression from face formed in subjects.

As predicted from the modulatory role of the amygdala in human behavior and decision making (Seymour and Dolan, 2008), we found that the changes of impression from face were significantly correlated with the amygdala activity. In an fMRI study by Petrovic *et al.* (2008), the authors used a classical conditioning paradigm with the combination of an electric shock and a face on normal subjects. The subjective

sympathy rating for the face worsened after the conditioning and correlated with the change in autonomic skin conductance response. The pairing of an US and a face significantly activated the bilateral amygdalae, as observed in the present study. Moreover, the degree of learning of the CS–US relationship significantly correlated with the neural activity in the amygdala and fusiform gyrus, indicating that learning of the affective value of a face is mediated by the amygdala's modulatory effect on the fusiform gyrus. In an fMRI study using facial pictures and sentences that described the person in the picture, the amygdala was significantly involved in processing impression formation and the amygdala signal negatively correlated with impression from face (Schiller *et al.*, 2009).

Although these studies (Petrovic *et al.*, 2008; Schiller *et al.*, 2009) showed a significant correlation between the signal and the impression from face, their findings are opposite in terms of the relationship between the brain responses and the subjective rating. The present study revealed that the lower the amygdala signal, the more negative subject's impression, whereas the subjects with a high amygdala signal in the Schiller *et al.* (2009) and Petrovic *et al.* (2008) studies had a low impression from face. The discrepancy between the studies could be explained by the fact that we adopted an experimental design which simulated a social situation wherein the subject was exposed to the emotional expression of another person. In a study by Schiller *et al.* (2009), the experimental task used a pair of face and sentence as social information and did not necessarily evoke emotional reaction in the subjects. In a study by Petrovic *et al.* (2008), it was not clear whether electric shock was given by a person shown on the screen, whereas in the present study, negative emotion was directly expressed by a person shown on the screen. The difference in attribution of negative emotion would be a cause of the difference in correlation between the amygdala activity and the impression from face.

Another explanation is that the subjects were explicitly instructed regarding the CS–US contingency in the present study; however, in a previous study using a conditioning paradigm, the CS–US relationship was not always known to the subjects (Sehlmeyer *et al.*, 2009). The amygdala activity has been shown to differ significantly between the condition in which the subjects were aware of the contingency and the condition in which they were not (Tabbert *et al.*, 2006). Therefore it is possible for an experimental paradigm using an explicit CS–US contingency and a mildly aversive US such as an emotional face and voice to yield different results from those of the previous studies.

The subtraction of the CS⁻ from the CS⁺up revealed the significant involvement of the dorsal ACC in the formation and expression of a negative impression from face. The time modulation analysis, to examine whether the neuronal responses to the CS changed with time and the conditions, also revealed the involvement of the dorsomedial PFC. The significant time-by-condition interaction effect observed in the

dorsomedial PFC indicates that the response to the CS⁺up was the greatest and that to the CS⁻ was the least in the first run. The differential activity between the CS⁺up and the CS⁻ was gradually reversed as the experiment proceeded from the second to the fourth run. On the other hand, no such interaction was observed in the left amygdala ($x, y, z = -20, -2, -18$) at the lenient threshold ($P = 0.05$, uncorrected). Therefore, this phenomenon might have been specific to the neural responses in the dorsomedial PFC and dorsal ACC when the subjects were exposed to the CS.

In an fMRI study using facial pictures and sentences that described personality traits such as friendly and inconsiderable, the dorsomedial PFC was activated to a significantly higher degree during the impression formation task than during the simple and non-social task for a face (Mitchell *et al.*, 2004). Activity in the anterior and ventral parts of the medial PFC was specifically involved in the later memory performance, suggesting that the medial PFC not only reflected the difference in the two tasks but also functioned significantly for successful memory encoding of the face. Activity in similar regions of the medial PFC was involved in the judgment of name–adjective pairs describing personality traits [e.g. assertive, nervous (Mitchell *et al.*, 2002)]. These regions are located in the anteriorostral medial frontal cortex (arMFC), which is related to the functions of social cognition such as self-knowledge, person perception and mentalizing (Amodio and Frith, 2006). However, in the present study, the activation was located in the posterior, not in the anterior, region of the medial PFC, which is related to conflict monitoring and error monitoring (Amodio and Frith, 2006). This difference may be due to our application of the modified version of the classical conditioning paradigm, in which the involvement of the posterior part of the ACC has often been reported (Sehlmeyer *et al.*, 2009). In fMRI experiments that used a partial conditioning paradigm and an aversive US (Buchel *et al.*, 1998, 1999), the CS⁺up events had significantly greater activation in the ACC than the CS⁻ events.

Activity of the human dorsal ACC revealed by neuroimaging experiments has been assumed to play a critical role in fear expression rather than in fear acquisition (Milad *et al.*, 2007). Several ACC regions adjacent to the activation found in the present study had a relationship with the degree of skin conductance response measured during the experiment (Milad *et al.*, 2007). In addition, a recent meta-analysis of fear studies using fMRI showed that the dorsal ACC is significantly active during both instructed and uninstructed conditioning paradigms (Mechias *et al.*, 2010). In a neuroanatomical study of monkeys, the medial PFC and ACC were found to have dense output projections from the PFC to the amygdala (Ghashghaei *et al.*, 2007). Although the differences in anatomical and physiological properties between the species must be considered, these two regions could also have a substantial effect on amygdala functioning in humans.

Two other regions associated with the subtraction analysis of the CS⁺ up and CS⁻ were found in the STS where similar activation has been observed previously (Iidaka *et al.*, 2010). The STS has been implicated in a polymodal association cortex that combines auditory and visual information (Calvert, 2001) and may be related to the cross-modal operations of the facial picture and voice in our experiment. In an fMRI study using meaningless and emotional speech (Grandjean *et al.*, 2005), the authors observed significant activation in the right STS under the angry prosody condition 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 the adjectives that were spoken (Wildgruber *et al.*, 2006). These results indicate that the right STS is involved in processing the negative emotional intonation of auditory stimuli. In the present study, the right STS was predominantly activated in the CS⁺ up condition than in the CS⁻ condition and that neither condition was paired with the voice stimuli, suggesting that reactivation of this region occurred because of previous exposure to the voice stimulus associated with negative prosody.

In conclusion, the present study revealed that the dorsomedial PFC, dorsal ACC, STS and amygdala play a role in the formation and expression of a negative impression from face. The result that the amygdala activity was significantly associated with the change of impression from face indicates the critical influence of the amygdala on decision making based on a person's emotional behavior. The dorsomedial PFC has been implicated in the conditioning paradigm where the neural response was the greatest at the beginning of the acquisition and gradually decreased. Although we investigated only a negative aspect of facial emotion, a study involving the neural mechanisms of positive impression may be conducted in the future. In addition, the method in which only two identities were shown in the conditioning paradigm may have produced the results specific to these faces; therefore, it might be better to use multiple faces in a further study. In spite of its disadvantages, the present study revealed that impression from face, particularly with negative valence, could be formed by orchestrated neural activity in the amygdala, dorsomedial PFC and dorsal ACC.

SUPPLEMENTARY MATERIAL

Supplementary material is available at SCAN online.

Conflict of Interest

None declared.

REFERENCES

- Adams, R.B., Ambady, N. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion*, 30, 179–88.
- Adolphs, R., Tranel, D., Hamann, S., et al. (1999). Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia*, 37, 1111–7.
- Ambady, N., Rosenthal, R. (1993). Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, 64, 431–41.
- Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Review Neuroscience*, 7, 268–77.
- 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, 10869–76.
- Buchel, C., Morris, J., Dolan, R.J., Friston, K.J. (1998). Brain systems mediating aversive conditioning: an event-related fMRI study. *Neuron*, 20, 947–57.
- Calvert, G.A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cereb Cortex*, 11, 1110–23.
- Costafreda, S.G., Brammer, M.J., David, A.S., Fu, C.H. (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews*, 58, 57–70.
- Fendt, M., Fanselow, M.S. (1999). The neuroanatomical and neurochemical basis of conditioned fear. *Neuroscience and Biobehavioral Reviews*, 23, 743–60.
- Gewirtz, J.C., Davis, M. (2000). Using pavlovian higher-order conditioning paradigms to investigate the neural substrates of emotional learning and memory. *Learning and Memory*, 7, 257–66.
- Ghashghaei, H.T., Hilgetag, C.C., Barbas, H. (2007). Sequence of information processing for emotions based on the anatomic dialogue between prefrontal cortex and amygdala. *Neuroimage*, 34, 905–23.
- Grandjean, D., Sander, D., Pourtois, G., et al. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, 8, 145–6.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, 51, 59–67.
- Iidaka, T., Saito, D.N., Komeda, H., et al. (2010). Transient neural activation in human amygdala involved in aversive conditioning of face and voice. *Journal of Cognitive Neuroscience*, 22, 2074–85.
- LeDoux, J.E. (2000). Emotion circuits in the brain. *Annual Review Neuroscience*, 23, 155–84.
- Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annual review of neuroscience*, 24, 897–931.
- Maren, S., Quirk, G.J. (2004). Neuronal signalling of fear memory. *Nature Review Neuroscience*, 5, 844–52.
- Mechias, M.L., Etkin, A., Kalisch, R. (2010). A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. *Neuroimage*, 49, 1760–8.
- Milad, M.R., Quirk, G.J., Pitman, R.K., Orr, S.P., Fischl, B., Rauch, S.L. (2007). A role for the human dorsal anterior cingulate cortex in fear expression. *Biological Psychiatry*, 62, 1191–4.
- Mitchell, J.P., Heatherton, T.F., Macrae, C.N. (2002). Distinct neural systems subserve person and object knowledge. *Proceedings of the National Academy of Sciences of the USA*, 99, 15238–43.
- Mitchell, J.P., Macrae, C.N., Banaji, M.R. (2004). Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Neuroscience*, 24, 4912–7.
- Morris, J.S., Buchel, C., Dolan, R.J. (2001). Parallel neural responses in amygdala subregions and sensory cortex during implicit fear conditioning. *Neuroimage*, 13, 1044–52.
- Petrovic, P., Kalisch, R., Pessiglione, M., Singer, T., Dolan, R.J. (2008). Learning affective values for faces is expressed in amygdala and fusiform gyrus. *Social Cognition and Affective Neuroscience*, 3, 109–18.
- Phan, K.L., Wager, T., Taylor, S.F., Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, 16, 331–48.
- Phillips, M.L., Medford, N., Young, A.W., et al. (2001). Time courses of left and right amygdalar responses to fearful facial expressions. *Human Brain Mapping*, 12, 193–202.

- Pollak, D.D., Rogan, M.T., Egner, T., Perez, D.L., Yanagihara, T.K., Hirsch, J. (2010). A translational bridge between mouse and human models of learned safety. *Annals of Medicine*, 42, 115–22.
- Posamentier, M.T., Abdi, H. (2003). Processing faces and facial expressions. *Neuropsychology Review*, 13, 113–43.
- Schiller, D., Freeman, J.B., Mitchell, J.P., Uleman, J.S., Phelps, E.A. (2009). A neural mechanism of first impressions. *Nature Neuroscience*, 12, 508–514.
- Sehlmeyer, C., Schoning, S., Zwitserlood, P., et al. (2009). Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS One*, 4, e5865.
- Seymour, B., Dolan, R. (2008). Emotion, decision making, and the amygdala. *Neuron*, 58, 662–71.
- Steele, J.D., Lawrie, S.M. (2004). Segregation of cognitive and emotional function in the prefrontal cortex: a stereotactic meta-analysis. *Neuroimage*, 21, 868–75.
- Tabbert, K., Stark, R., Kirsch, P., Vaitl, D. (2006). Dissociation of neural responses and skin conductance reactions during fear conditioning with and without awareness of stimulus contingencies. *Neuroimage*, 32, 761–70.
- Todorov, A. (2008). Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, 1124, 208–24.
- van't Wout, M., Sanfey, A.G. (2008). Friend or foe: the effect of implicit trustworthiness judgments in social decision-making. *Cognition*, 108, 796–803.
- Wild, B., Erb, M., Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: quality, quantity, time course and gender differences. *Psychiatry Research*, 102, 109–24.
- Wildgruber, D., Ackermann, H., Kreifelts, B., Ethofer, T. (2006). Cerebral processing of linguistic and emotional prosody: fMRI studies. *Progress in Brain Research*, 156, 249–68.
- Willis, J., Todorov, A. (2006). First impressions: making up your mind after a 100-ms exposure to a face. *Psychological Science*, 17, 592–8.
- Winston, J.S., Strange, B.A., O'Doherty, J., Dolan, R.J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience*, 5, 277–83.