PERGAMON

Neuropsychologia 41 (2003) 184–194

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Spatio-temporal dynamics of brain mechanisms in aversive classical conditioning: high-density event-related potential and brain electrical tomography analyses

Diego A. Pizzagalli[a,b,*], Lawrence L. Greischar[a,b], Richard J. Davidson[a,b]

[a] *Department of Psychology, University of Wisconsin-Madison, 1202 W. Johnson Street, Madison, WI 53706, USA*
[b] *W.M. Keck Laboratory for Functional Brain Imaging and Behavior, University of Wisconsin-Madison, 1500 Highland Avenue, Madison, WI 53705, USA*

Abstract

Social cognition, including complex social judgments and attitudes, is shaped by individual learning experiences, where affect often plays a critical role. Aversive classical conditioning—a form of associative learning involving a relationship between a neutral event (conditioned stimulus, CS) and an aversive event (unconditioned stimulus, US)—represents a well-controlled paradigm to study how the acquisition of socially relevant knowledge influences behavior and the brain. Unraveling the temporal unfolding of brain mechanisms involved appears critical for an initial understanding about how social cognition operates. Here, 128-channel ERPs were recorded in 50 subjects during the acquisition phase of a differential aversive classical conditioning paradigm. The CS+ (two fearful faces) were paired 50% of the time with an aversive noise (CS↑ + /Paired), whereas in the remaining 50% they were not (CS↑ + /Unpaired); the CS− (two different fearful faces) were never paired with the noise. Scalp ERP analyses revealed differences between CS↑ + /Unpaired and CS− as early as ∼120 ms post-stimulus. Tomographic source localization analyses revealed early activation modulated by the CS+ in the ventral visual pathway (e.g. fusiform gyrus, ∼120 ms), right middle frontal gyrus (∼176 ms), and precuneus (∼240 ms). At ∼120 ms, the CS− elicited increased activation in the left insula and left middle frontal gyrus. These findings not only confirm a critical role of prefrontal, insular, and precuneus regions in aversive conditioning, but they also suggest that biologically and socially salient information modulates activation at early stages of the information processing flow, and thus furnish initial insight about how affect and social judgments operate.

Keywords: Fear conditioning; Affect; Faces; Laterality; Source localization; LORETA

1. Introduction

Social cognition refers to the ability to evaluate and predict others' behavior, affect, intentions and beliefs, i.e. the ability to interact in complex social environments. Complex, socially-appropriate decision making, especially in situations where incomplete information is available, often involves building representation of contingencies based on prior experiences. Thus, social cognition and evaluation heavily relies on individual experiences, ultimately on learning and memory [12].

In recent years, the study of brain mechanisms governing social cognition has attracted considerable attention [7,43]. As classical conditioning is considered as "a prototype for the study of brain functions involved in unconscious acquisition of knowledge" ([21], p. 235)—a mechanism that appears to play a key role in social cognition (e.g. [1])—this paradigm represents a powerful and well-controlled tool to investigate brain mechanisms involved in important aspects of social cognition.

Aversive classical conditioning or fear conditioning [41] is a form of associative learning involving a relationship between an initially neutral event (the conditioned stimulus, CS) and an event with biological significance (the unconditioned stimulus, US). By being temporally coupled with a salient stimulus (US), the CS becomes behaviorally significant and it comes to elicit a range of species-specific defensive responses, even when the US is no longer presented [11,30]. In animals, the brain circuitry involved in the acquisition and maintenance of conditioned fear has been extensively investigated (for reviews, [11,30]). These studies have shown that CS–US pairing leads to activation of the lateral nucleus of the amygdala,

* Corresponding author. Present address: Department of Psychology, Harvard University, 1220 William James Hall, 33 Kirkland Street, Cambridge, MA 02138, USA. Tel.: +1-617-496-8896; fax: +1-617-495-3728.
E-mail address: dap@wjh.harvard.edu (D.A. Pizzagalli).

0028-3932/02/$ – see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0028-3932(02)00148-3

which in turn modulates motor and endocrine-autonomic output.

In recent years, several functional neuroimaging studies have investigated the neural substrates of classical conditioning in humans, confirming a critical role of the amygdala, orbitofrontal cortex (OFC), medial prefrontal cortex (PFC), and insula in classical aversive conditioning (for reviews, [5,44]). The amygdala, in particular, may be critical for formation and modulation of associative links with aversive character, which may result in learning-dependent plasticity in cortical regions encoding the sensory contingencies involved (e.g. [25,36]). Further, in light of their respective roles in the regulation, evaluation, and autonomic control of affect, medial PFC, OFC, and insula play an important role in anticipation of and reactivity to aversive cues (for review, [10]). Intriguingly, amygdalar activation has been found to be associated with unconscious evaluation of racial groups [43]; individual variations in amygdalar activation in response to faces of members of a racial minority were interpreted as being "a reflection of culturally acquired knowledge about social groups filtered through individual experience" ([43], p. 733).

Whereas functional neuroimaging studies have greatly advanced our understanding of the brain circuitry governing fear conditioning, their coarse temporal resolution (in the order of seconds) does not allow the investigation of the temporal dynamic of underlying brain mechanisms. This information appears important for a comprehensive understanding about how the brain processes socially relevant information. For instance, demonstration that biologically and socially relevant information is processed very quickly in the human brain, e.g. within the first 150 ms after stimulus presentation, would strongly support the notion that social cognition and evaluation operate in an automatic and preattentive way (e.g. [17]). To this end, scalp recordings of brain electrical activity (event-related potentials, ERPs) can be used to gather non-invasive access to human brain activity with a time resolution of milliseconds.

Although several ERP studies have investigated cortical correlates of classical aversive conditioning, results often diverge. Whereas some studies have demonstrated CS+ modulation during acquisition of fear conditioning on early ERP components (P100, N100, P200) [3,4,23,33,60], others have found modulations of later ERP components, particularly the P300 [53,54,61] and the CNV complex [15,50,60]. Recently, Baas et al. [3] reported accentuated exogenous ERPs starting at 60 ms over occipital regions as well as increased P300 and late frontal negativity for a threat cue (a grating predicting potential delivery of a shock) compared to a safe cue. Similarly, Montoya et al. [33] reported increased ERPs to auditory pseudo-words conditioned to electric shocks at 100 ms. The early threat-modulated activity replicated recent findings suggesting that socially salient information, such as likability/attractiveness ([46]: 80–116 ms; [47]: ∼112 ms) and facial expression ([19]: ∼110 ms) modulated initial stages of the information processing flow. As a corpus, these results are consistent with the assumption that adaptive behavior relies on rapid monitoring of potentially salient cues in the environment, a mechanism that may be implemented through enhanced sensory processing. Since the amygdala has been implicated in threat detection [5,30], amygdalar projections to pathways in the visual system have been invoked for explaining such early affect-modulated activity. Besides early exogenous ERP components, fear conditioning may modulate a late frontal negativity wave which has been observed under various experimental manipulations that involved anticipation of various arousing stimuli, such as aversive noise [50], shock [3], and high-interest pictorial stimuli [51].

In the present study, high-density ERP recordings and a tomographic source localization technique based on probabilistic brain atlases were used to investigate with increased spatial sensitivity the temporal unfolding of brain mechanisms involved in aversive conditioning. Based on the functional neuroimaging and ERP findings reviewed above, we hypothesized that the CS+ would affect activity in the medial PFC, OFC, insula, and modality-specific sensory regions, with early exogenous potentials (as indices of boosted sensory processing) and late frontal negativity (as index of anticipatory processes) as the primary ERP components involved.

2. Methods

2.1. Participants

Fifty subjects (24 female, 26 male; mean age: 19.02, S.D.: 1.07) participated in the experiment. All subjects were undergraduate students at the University of Wisconsin-Madison, were right-handed [8], and free of any medical or neurological problems. For their participation, they received US$ 40.

2.2. Procedure

After written informed consent was obtained to an experimental protocol approved by the University of Wisconsin Human Subjects Committee, subjects were comfortably seated in a room equipped with an intercom system. A chin-rest was used to keep the distance between the subject and the screen constant (50 cm). Subjects were instructed to passively observe facial stimuli serially presented on the monitor and that sometimes they will hear an annoying but not painful noise while observing the face images. They were told that the noise was used to act as a mild stressor.

For the ERP task, 224 trials were presented divided in 4 blocks each with 56 trials and separated by a 30-s break. Each block started with the presentation of a fixation cross for 3 s that was used to help to focus on the screen location where the faces were presented. Within each block, four conditions were equally presented (14 trials/condition):

"CS↑ + /Paired", "CS↑ + /Unpaired", "CS↑ − /Fear", and "CS↑ − /Happy". All stimuli were selected from the Ekman and Friesen's series [13]. For the CS+, two photographs of facial expression of fear were selected (one man, one woman). For the CS↑ − /Fear, two different photographs of facial expression of fear were used (one man, one woman). For the CS↑ − /Happy, two photographs of facial expression of happiness were selected (one man, one woman). The stimuli were size-, contrast- and brightness-adjusted portrait photographs (8 cm × 12 cm). To study aversive fear conditioning without the interference of the US, a partial reinforcement schedule was used. Thus, half the time, two of the fearful faces were paired (CS↑ + /Paired) with an aversive noise delivered through commercial earbuds; on the remaining trials they were unpaired (CS↑ + /Unpaired). The aversive noise—an overlay of several commercially available sound effects (smoke alarm, ringing bell, foghorn, and guitar electric feedback) created using Cool Edit Pro (Syntrillium Software Corporation, Phoenix, AZ) software [52]—was played with a nearly instantaneous rise time at 95 dB for 1 s immediately after offset of the CS+. The second set of fearful faces (CS↑ − /Fear) as well as the happy faces (CS↑ − /Happy) were never paired with the noise and served as control conditions. Facial stimuli were presented for 450 ms in a pseudo-randomized sequence with a variable ITI (range: 1.5–1.8 s). The pseudo-random sequence had the constraint that no more than three trials of the same condition, ITI, or model posing the facial expression could occur in sequence. For ERP computations, 56 trials/condition were available. The acquisition phase lasted approximately 10 min.

After the ERP recordings, subjects were asked to rate the US with respect to valence (from 1, "very unpleasant" to 9, "very pleasant") and arousal (from 1, "very moderate" to 9, "very intense") using two analog scales presented via the PC.

2.3. Apparatus

Presentation of facial stimuli and aversive noise was controlled by in-house software using a 100-MHz Pentium PC. The 128-channel EEG was recorded using the Geodesic Sensor Net system (Electrical Geodesic Inc., Oregon; [58]), where EEG electrodes are arrayed in a regular distribution across the head surface and the inter-sensor distance is approximately 3 cm. The EEG was digitized at 250 Hz (bandwidth: 0.01–100 Hz, with the vertex electrode (Cz) serving as recording reference. Four additional channels were used to mark the stimulus conditions.

2.4. ERP analyses

After gain and zero calibration, data were imported and analyzed in BESA (Version 4.06; MEGIS Software GmbH, Munich, Germany). Channels with corrupted signals throughout the recording (on average, 5.48 ± 2.97) were interpolated using a spline interpolation method [42]. After off-line automatic artifact rejections (amplitude: 100 μV; gradient: 75 μV; low signals: 0.032) for excluding blink, muscle and other artifacts, ERPs were computed covering 1024 ms and time-locked to the onset of CS+ or CS− presentations, with a 100-ms pre-stimulus baseline. ERP data were then baseline-corrected, lowpass filtered at 35 Hz (12 dB/octave roll-off), and re-referenced to the average reference. For these initial analyses and for the sake of brevity, only the CS↑ + /Unpaired and CS↑ − /Fear were investigated, as this contrast is the most conservative for investigating brain mechanisms involved in aversive conditioning due to the fact that it controls for the facial expression. No significant differences existed in the number of artifact-free epochs contributing to the ERPs (CS↑ + /Unpaired: 38.38 ± 9.00 versus CS↑ − /Fear: 38.62 ± 9.56; *t*-test $P > 0.50$). Finally, three grandmean ERPs were computed across the 50 subjects, one for each condition as well as one across conditions.

Scalp ERP data were analyzed with space-oriented brain electric field analysis [31]. This approach is based on the empirical observation that the brain electric field configuration (map landscape) changes step-wise and discontinuously [31]. Epochs of quasi-stable field configurations ("microstates") are concatenated by abrupt transitions in the brain electric field configurations. Accordingly, microstates index brief periods of quasi-stable spatial configurations of the active neural generators and thus are assumed to implement specific brain functions. To identify start and end times of different field configurations, a data-driven segmentation procedure previously described [26,31,45,46] was employed using the grandmean ERP across conditions. The goal of this sequential parsing procedure is to identify ERP microstates as periods characterized by stable field configuration by simultaneously satisfying two contrasting goals: stability (assignment of a maximal number of maps to a given microstate), and discrimination (detection of a maximal number of different microstates). For the segmentation procedure, a Global Map Dissimilarity index was computed at each time frame [31]. This index is a reference-free, single-value indicator for the difference of landscape configuration between two successive maps; it can vary between zero, in case of identical map topographies, and two in case of two maps with identical topography but reversed polarity.

For each microstate, global field power (GFP) peaks were subsequently determined for the grandmean ERPs. GFP is computed as the average standard deviation within the potential field, and is a reference-free, single-value measure of the amount of field strength at each time frame [31]. GFP peaks are assumed to index time points associated with maximal activity of the neuronal populations; thus, only time of GFP peaks were used for statistical analyses. To increase spatial sensitivity of scalp ERP analyses, paired *t*-tests contrasting the CS↑ + /Unpaired and CS↑ − /Fear conditions were performed at each sensor (d.f. = 49). Results were accepted as significant at $P < 0.01$ ($t_{49} > 2.68$).

2.5. Low resolution brain electromagnetic tomography (LORETA) analyses

At the GFP peaks, the cortical three-dimensional distribution of current density for the CS↑ + /Unpaired and CS− stimuli was computed using LORETA [39,40]. Without postulating a specific number of generating sources, this algorithm solves the inverse problem by assuming related orientations and strengths of neighboring neuronal sources, an assumption mathematically implemented by finding the "smoothest" of all possible activity distributions.

As in recent studies (e.g. [27,37,39,45,47]), we used the version (LORETA-KEY©) based on a three-shell spherical head model and EEG electrode coordinates derived from cross-registrations between spherical and realistic head geometry [57]. The head model has been registered to a standardized stereotactic space [55] available as digitized MRI from the Brain Imaging Centre (Montreal Neurologic Institute, MNI305). Further, the source solution space was limited to cortical gray matter and hippocampi according to digitized Probability Atlases provided by the Montreal Neurologic Institute (number of voxels: 2394; voxel dimension: 7 mm^3). Voxels were defined as gray matter if their probability of being gray matter was greater than (a) 33%, (b) the probability of being white matter, and (c) the probability of being cerebrospinal fluid. At each voxel, current density was computed as the linear, weighted sum of the scalp electric potentials, normalized to a total current density of 1, and then log-transformed to normalize the distribution. LORETA units are scaled to amperes per square meter (Å/m^2). Subsequently, whole-brain analyses using voxelwise paired *t*-tests examined differences between CS↑ + /Unpaired and CS−. Results were accepted as significant at $P < 0.005$. In cases of significant results involving regions of theoretical importance that were lateralized to one hemisphere, putative asymmetric results were formally tested by running a two-way ANOVA with *Condition* (CS↑ + /Unpaired versus CS−) and *Hemisphere* (right, left) as repeated measures [10]. Homologous clusters were defined by reversing the sign of the X coordinates. The structure-probability maps atlas [29] was used to label regions and Brodmann areas showing significant differences between the conditions.

3. Results

3.1. Post-recording affective ratings

The affective ratings collected after the ERP recordings confirmed that the US was experienced as aversive (mean valence rating: 1.80, S.D.: 1.20; mean arousal rating: 7.22, S.D.: 2.19). Eighty-six percent of the subjects rated the US with a valence value of 1 or 2 (from 1, "very unpleasant" to 9, "very pleasant") and an arousal value from 7 to 9 (from 1, "very moderate" to 9, "very intense").

3.2. Scalp ERP data

The segmentation procedure identified 84–152, 156–188, 192–532, 536–608, and 612–756 ms as stable spatial configurations. The first microstate (84–152 ms) corresponded to the P100 component, with positivity over medial occipital regions (Fig. 1A). The second microstate (156–188 ms) included the N100 component, with negativity over bilateral ventral occipito-temporal regions (Fig. 1B). The third microstate (192–532 ms) was characterized by a sustained and strong negativity over anterior-ventral frontal regions (slightly lateralized to the left hemisphere) and positivity over medial occipital regions (slightly lateralized to the right hemisphere) (Fig. 1C–F). The fourth microstate (536–608 ms) showed a similar topography, which was however attenuated compared to the third microstate (Fig. 1G). Finally, in the fifth microstate (608–756 ms), the negativity moved over ventral occipito-temporal regions, whereas a positivity emerged over centro-parietal regions (Fig. 1H).

The *t*-tests contrasting the scalp potential distribution between the two conditions at each sensor and at GFP peaks revealed several significant results ($P < 0.01$), which generally confirmed our hypotheses. At ∼120 ms, CS↑ + /Unpaired elicited a stronger negativity over a right inferior frontal sensor (sensor E8; Fig. 2A), and stronger positivity over centro-parietal regions (sensors E32, E55, E81, E88; Fig. 2B). The difference in negativity over right frontal regions increased systematically over time, as it was also observed at 240 ms (E1, E2, E8), 316 ms (E1, E21, E115, E125), 412 ms (E1, E2, E8, E9, E121, E125, E126, E127), and 464 ms (E1, E2, E8, E9, E121, E125, E126, E127). Interestingly, at GFP peaks close to the potential delivery of the US, the negativity associated with the CS+ extended from inferior frontal to inferior temporal regions (Fig. 1E). At 464 ms, the CS↑ + /Unpaired elicited a significantly stronger late positive complex over centro-parietal regions than the CS (Fig. 1F), which vanished over time.

3.3. LORETA ERP data

Fig. 3 and Table 1 summarize significant results of the voxelwise *t*-tests contrasting current density between the CS↑ + /Unpaired and CS− at GFP peaks within the epochs identified by the data-driven segmentation procedure. In the following, significant results ($P < 0.005$, uncorrected) are summarized only if the cluster involved at least three voxels (>1.03 cm^3).

As early as 120 ms after onset of the facial stimulus (Fig. 3A), the CS↑ + /Unpaired elicited stronger current density in the ventral visual pathway (inferior temporal gyrus, BA 20 and fusiform gyrus, BA 37), particularly in the right hemisphere (12 voxels; see cluster "1" in Fig. 3A; *Condition* × *Hemisphere*: $F(1, 49) = 1.84$, ns). In the left hemisphere, the activation modulated by the CS+ within the ventral visual pathway was slightly more posterior and dorsal, including the middle occipital gyrus (BAs 37, 19;

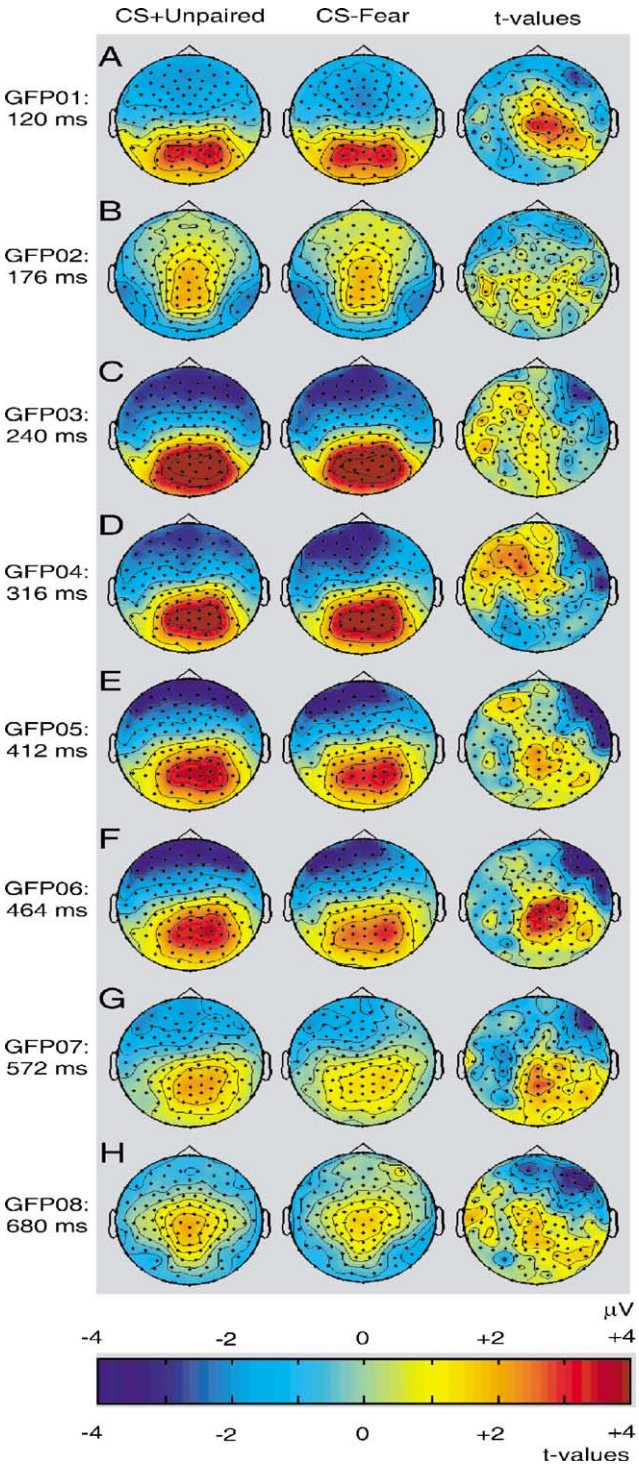

Fig. 1. Mean potential distribution maps averaged across subjects ($n = 50$) at given GFP peaks (A–H) for the CS↑ + /Unpaired (left column) and CS− (middle column) conditions (vs. average reference). In the right column, *t*-maps contrasting the two conditions are displayed (warm colors: CS+Unpaired > CS−; cold colors: CS+Unpaired < CS−; see calibration scale). Maps were created using linear interpolation with software kindly provided by Dr. Scott Makeig (http://www.sccn.ucsd.edu/~scott/ica.html).

three voxels). At the same latency, the CS− was associated with stronger current density than the CS↑ + /Unpaired in the left insula (BA 13; 11 voxels, see cluster "2"; *Condition* × *Hemisphere*: $F(1, 49) = 4.45$, $P < 0.05$) as well as in the left middle frontal gyrus (BA 9; three voxels, see cluster "3" in Fig. 3A; *Condition* × *Hemisphere*: $F(1, 49) = 1.53$, ns).

In the N100 range (GFP peak: 176 ms; Fig. 3B), the CS↑ + /Unpaired was associated with relatively stronger current density in three large clusters: a right hemispheric frontal cluster (26 voxels) that encompassed the middle and superior frontal gyri (BA 6) and precentral gyrus (BA 4; see cluster "4" in Fig. 3B; *Condition* × *Hemisphere*: $F(1, 49) = 7.03$, $P < 0.01$; *Hemisphere*: $F(1, 49) = 12.16$, $P < 0.001$, right > left), a right hemispheric cluster (14 voxels; see cluster "5" in Fig. 3B; *Condition* × *Hemisphere*: $F(1, 49) = 2.31$, ns) including the fusiform gyrus (BAs 36, 20, 37) and inferior temporal gyrus (BAs 20, 37), and finally a left hemispheric cluster (15 voxels) involving the inferior and superior parietal lobule (BAs 7, 39; see cluster "6" in Fig. 3B). In addition, the CS↑ + /Unpaired was associated with stronger current density in a small left hemispheric cluster (three voxels) that included the superior temporal gyrus (BA 29), insula (BA 13), transverse temporal gyrus (BA 41). Conversely, the CS− was associated with stronger activation in the right middle and superior frontal gyri (BAs 10, 11; 13 voxels, see cluster "7" in Fig. 3B).

At 240 ms post-stimulus, the CS↑ + /Unpaired elicited stronger activation than the CS− in a medial posterior cluster that included the cuneus and precuneus (BAs 31, 18, 19; 17 voxels; data not shown). The CS−, on the other hand, was associated with relatively stronger activation in the left middle frontal gyrus (BA 11; three voxels; *Condition* × *Hemisphere*: $F(1, 49) = 2.79$, ns; *Hemisphere*: $F(1, 49) = 4.41$, $P < 0.05$, right > left).

At 316 ms, the CS↑ + /Unpaired elicited stronger activation in the left middle and inferior temporal (BA 37) and middle occipital gyri (BA 19), whereas the CS− was associated with increased current density in the right superior frontal gyrus (BA 9), in clusters that were however very limited in size (three and two voxels, respectively; data not shown).

Shortly before the potential delivery of the loud aversive noise (US) (Fig. 3C), the CS↑ + /Unpaired elicited increased activation in right auditory regions (10 voxels; see cluster "8" in Fig. 3C; *Condition* × *Hemisphere*: $F(1, 49) = 0.92$, ns) centered on the middle temporal gyrus (BA 21) and extending ventrally to the inferior temporal gyrus (BA 20), whereas the CS− was associated with increased activation in orbitofrontal and medial frontal regions that extended dorsally to the superior frontal gyrus (BAs 9, 10, 11; 37 voxels; see cluster "9" in Fig. 3C). Other clusters surviving the statistical threshold were too small (two voxels), and are thus not reported.

Immediately after CS+ offset and US omission (464 ms), the CS↑ + /Unpaired was associated with relatively

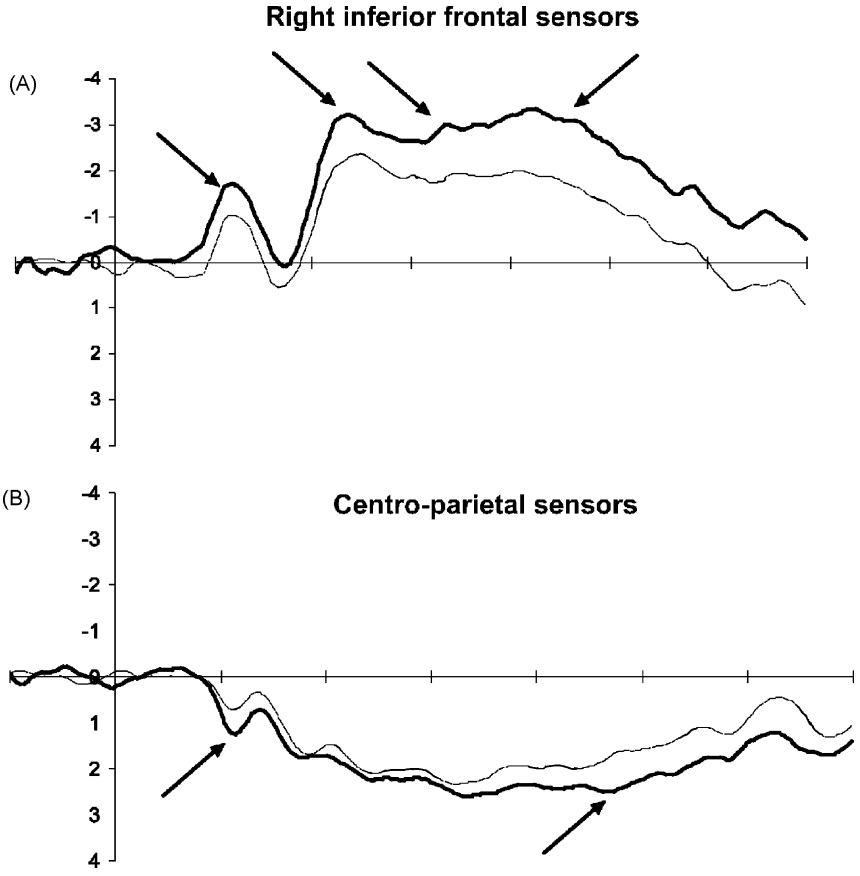

Fig. 2. Grandmean ERP waveforms ($n = 50$) elicited by CS↑ + /Unpaired (thick lines) and CS− (thin lines) at (A) a right inferior frontal sensor (sensor E8), and (B) centro-parietal sensors (E32, E55, E81, E88). Arrows denote GFP peaks where the CS↑ + /Unpaired and CS− were significantly different ($P < 0.01$). Horizontal lines: time in ms (from −100 to 700 ms from stimulus onset, tickmarks: 100 ms). Vertical lines: amplitude in μV. ERP waveforms versus average reference.

Table 1
Summary of significant results ($P < 0.005$, uncorrected) emerging from whole-brain LORETA (low resolution brain electromagnetic tomography) contrasting the CS↑ + /Unpaired and CS− at given global field peaks (GFP)

GFP peak	Contrast	*X*	*Y*	*Z*	Region	Side	BA	*t*-value	*P*-value
GFP 1: 120 ms	CS↑ + /Unpaired > CS↑ − /Fear	60	−53	−13	Inferior temporal gyrus	Right	20	4.71	0.0000
	CS↑ + /Unpaired < CS↑ − /Fear	−38	−11	1	Insula	Left	13	−3.30	0.0018
GFP 2: 176 ms	CS↑ + /Unpaired > CS↑ − /Fear	25	−4	50	Middle frontal gyrus	Right	6	4.15	0.0001
	CS↑ + /Unpaired < CS↑ − /Fear	25	59	−13	Superior frontal gyrus	Right	11	−3.86	0.0003
GFP 3: 240 ms	CS↑ + /Unpaired > CS↑ − /Fear	18	−74	29	Precuneus	Right	31	3.77	0.0004
	CS↑ + /Unpaired < CS↑ − /Fear	−45	45	−13	Middle frontal gyrus	Left	11	−3.12	0.0030
GFP 4: 316 ms	CS↑ + /Unpaired > CS↑ − /Fear	−52	−60	1	Middle temporal gyrus	Left	37	3.47	0.0011
	CS↑ + /Unpaired < CS↑ − /Fear	11	59	29	Superior frontal gyrus	Right	9	−3.33	0.0017
GFP 5: 412 ms	CS↑ + /Unpaired > CS↑ − /Fear	67	−11	−6	Middle temporal gyrus	Right	21	3.41	0.0013
	CS↑ + /Unpaired < CS↑ − /Fear	−10	66	−13	Superior frontal gyrus	Left	11	−4.05	0.0002
GFP 6: 464 ms	CS↑ + /Unpaired > CS↑ − /Fear	18	−88	−13	Lingual gyrus	Right	18	3.28	0.0019
	CS↑ + /Unpaired < CS↑ − /Fear	4	59	−6	Middle frontal gyrus	Right	10	−3.07	0.0035
GFP 7: 572 ms	CS↑ + /Unpaired > CS↑ − /Fear	18	−11	71	Superior frontal gyrus	Right	6	3.89	0.0003
	CS↑ + /Unpaired < CS↑ − /Fear	–	–	–	–	–	–	–	–
GFP 8: 680 ms	CS↑ + /Unpaired > CS↑ − /Fear	−31	−67	57	Superior parietal lobule	Left	7	3.06	0.0036
	CS↑ + /Unpaired < CS↑ − /Fear	−52	−4	−20	Middle temporal gyrus	Left	21	−3.11	0.0031

For each GFP peak, the coordinates ([55], origin at anterior commissure), anatomical regions, hemisphere, and Brodmann areas (BA) are listed for the most positive (stronger current density for CS↑ + /Unpaired than CS−) and most negative (stronger current density for CS− than the CS↑ + /Unpaired) *t*-values. X = left (−) to right (+); Y = posterior(−) to anterior (+); Z = inferior (−) to superior (+).

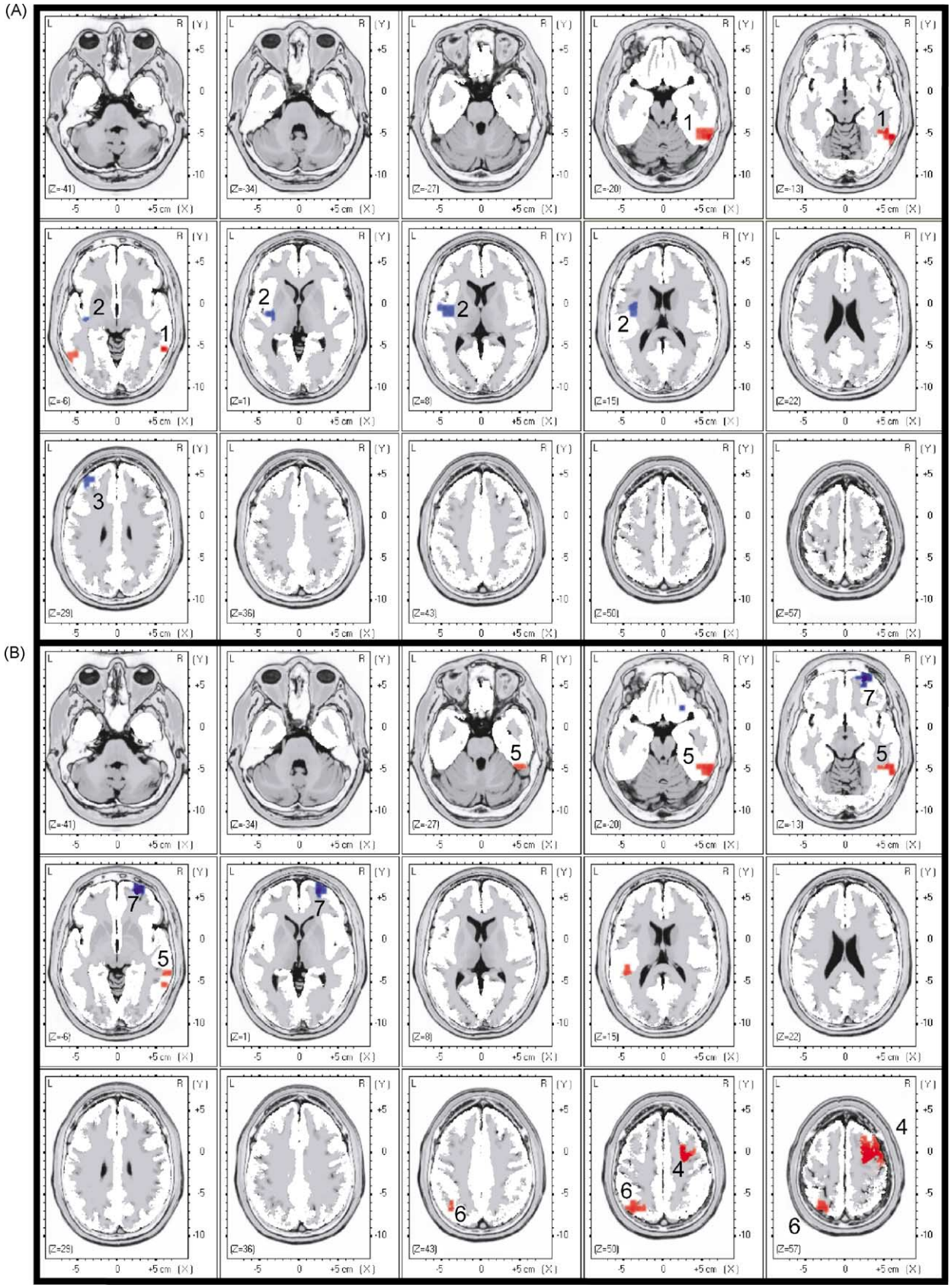

Fig. 3. Results of voxel-by-voxel *t*-tests contrasting the CS↑ + /Unpaired and CS− ($n = 50$) at (A) 120 ms, (B) 176 ms, and (C) 412 ms after onset of the facial stimulus (maps thresholded at $P < 0.005$). Fifteen axial brain slices are shown in steps of 7 mm from the most inferior level ($Z = -41$) to the most superior level ($Z = 57$). Red: stronger current density for CS↑ + /Unpaired than CS−. Blue: stronger current density for CS− than CS↑ + /Unpaired. Clusters surviving the statistical threshold and exceeding three voxels are numbered (see main text). Coordinates in mm [55], origin at anterior commissure; X = left (−) to right (+); Y = posterior (−) to anterior (+); Z = inferior (−) to superior (+).

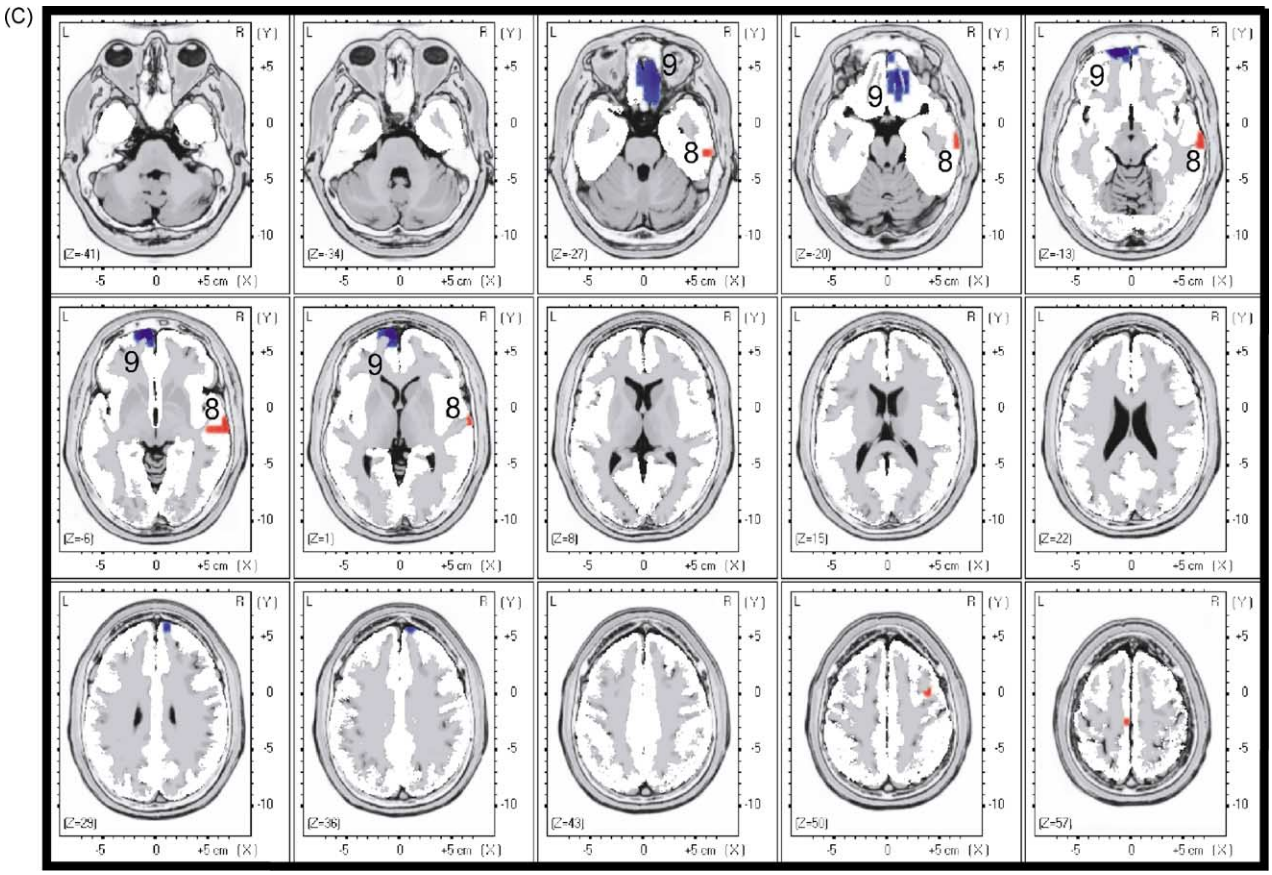

Fig. 3. (*Continued*).

increased activation in the right medial and superior frontal gyri (BA 6; eight voxels), left middle frontal gyrus (BA 6; eight voxels), and right lingual gyrus (BA 18; four voxels; data not shown). Other clusters are not reported because they were too small.

More than 100 ms after omission of the US (572 ms), the CS↑ + /Unpaired was associated with increased activation in a right frontal cluster (12 voxels) extending dorsally from the pre- and post-central (BAs 3, 4) to the superior frontal gyrus (BA 6) as well as in the cuneus/precunues (BA 19; 11 voxels). The CS+ modulated activation also in the right inferior and middle temporal gyrus (BAs 20, 21; three voxels) and left postcentral gyrus, which extended to the left precuneus (BA 7; five voxels). No significant CS− ↑CS+ differences were observed (data not shown). Finally, at 680-ms post-stimulus, clusters showing significant results were small and are not reported.

4. Discussion

In recent years, positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies have converged in their results by showing a crucial role of the amygdala, PFC, OFC, insula, and anterior cingulate cortex in classical aversive conditioning ([6,14,22,25,28,36,48]; for review [5]). Consistent with animal work [30], the amygdala, by acting as site of CS–US convergence, seems to be particularly involved in the acquisition of fear conditioning. The OFC (decoding of punishment or reward value or affective evaluation), ventromedial PFC (fear extinction and emotional regulation), and insula (autonomic control of affect) subserve important functions associated with aversive conditioning. In the present study, the use of high-density ERP recordings combined with a tomographic source localization technique added important information about the temporal unfolding of brain mechanisms involved in aversive conditioning within several of the regions previously implicated in PET/fMRI studies using similar paradigms.

4.1. Spatio-temporal aspects of brain mechanisms in aversive classical conditioning

Using a differential aversive classical conditioning paradigm with a 50% partial reinforcement strategy, we found that the presentation of CS↑ + /Unpaired stimuli (two fearful faces signaling the potential delivery of an aversive noise) was associated with higher activation than the CS− stimuli (two control fearful faces) in ventral visual pathways (120 and 176 ms), right middle and superior frontal

gyri (176, 464, and 572 ms), left parietal lobule (176 ms), cuneus/precuneus (240 and 572 ms), left middle/inferior temporal and middle occipital gyri (316 ms), right auditory cortex (412 ms), left middle frontal gyrus (464 ms), and right lingual gyrus (464 ms). Conversely, presentation of the safe stimuli (CS−) elicited relatively higher activation than the CS+ in the left insula (120 ms), left middle frontal gyrus (120 and 240 ms), right middle and superior frontal gyri (176 ms), and orbitofrontal and medial frontal PFC (412 ms).

Early modulation within ventral visual pathways (inferior temporal and fusiform gyri) is intriguing in light of the fact that these regions have been involved in early perceptual processing of faces (e.g. structural facial encoding; [56]; for review [20]). The present findings add to an emerging evidence that biologically and socially salient information modulates activation in visual cortices at early stages of the information processing flow, possibly indexing a rapid and coarse affective categorization [46,47]. Based on these and other findings, Adolphs [1] recently proposed that presentation of an affectively laden stimulus triggers an initial feed-forward information processing along occipital and temporal neocortices which leads to a coarse categorization of affect within the first 100 ms. Since the human amygdala has been shown (a) to respond to facial stimuli as early as 120 ms [18], (b) to be sensitive to threat-related cues [5], and (c) to project to ventral visual pathways [2], the early activation modulated by the CS+ in ventral visual pathways may well be linked to increased amygdalar activity in response to the CS+. Due to the difficulty of probing subcortical activity with scalp ERP data, this hypothesis awaits, however, direct empirical confirmation.

The early modulation of activation by the CS+ in right PFC regions (176 ms) is consistent with recent findings suggesting that territories of the PFC are involved in categorization of socially relevant information besides implementing more cognitive functions, such as, e.g. working memory. Intracranial ERP recordings in epileptic patients have indeed shown early categorization of faces versus objects in the right inferior frontal gyrus at ∼150 ms [32] as well as discrimination of fearful and happy facial expressions in the right ventromedial prefrontal cortex at ∼120 ms [24]. Interestingly, whereas the CS+ generally elicited stronger activation in the right PFC, the CS− was generally associated with relatively higher activation in the left PFC (120 and 240 ms; but see at 176 ms). Together with prior PET findings of an involvement of right PFC regions in aversive conditioning [14,16,22,35], these findings add to a substantial literature suggesting that the left and right PFC may be differentially involved in affect, with the left PFC being relatively more involved in positive, approach-related affect, and the right PFC in negative, withdrawal-related affect [9].

Stronger OFC and medial PFC activation at 412 ms (i.e. shortly before the US would have been presented in the CS+ condition) for the stimuli signalizing safety (CS−) is intriguing in light of animal [34] and human imaging [22,35,38] data implicating OFC and medial PFC regions in inhibition of behavioral responses and extinction of conditioned responses. Whereas future studies should directly test this hypothesis, increased medial PFC and OFC activation may be associated with re-organization of behavioral responses and re-allocation of resources in light of the safety associated with the CS−.

The modulation of auditory cortex activation by the CS+ before the potential delivery of the aversive noise (412 ms) is in agreement with previous animal [49,59] and human [25,36] work that has shown learning-dependent plasticity as a function of aversive conditioning in both the visual and auditory modality. In the Morris study, regression analyses indicated that conditioned activity within the auditory cortex (bilateral transverse temporal gyrus) covaried with activity in the amygdalae, basal forebrain, and right OFC. In general, these findings of modality-specific tuning in sensory cortices towards conditioned stimuli are consistent with the conjecture that, after initial amygdalar involvement in the formation of associative links with aversive character, learned associations are manifested in cortical regions [5]. Whereas modulation of the auditory cortex in the present study is consistent with prior results of learning-related plasticity, greater CS+ activity in auditory regions around 412 ms may alternatively reflect transient modulation of these regions due to attentional processes.

Finally, with respect to ERP components, the present findings confirm that cues associated with a potential aversive outcome modulate both early exogenous ERP components (P100, N100) and late frontal negativity (e.g. [3,4,15,23,33,50]). The earlier exogenous components may index boosted activation in early visual relays in the information processing flow, whereas the late frontal negativity that developed towards the offset of the threat cue predicting the potential delivery of the US may index anticipation of the aversive outcome (see also [3]).

4.2. Limitations

Although, the results of this study provide new insights into the spatio-temporal dynamics of brain mechanisms underlying aversive classical conditioning, some limitations exist. First, a large number of brain regions were involved, and the interpretations of the functional role of specific structures in conditioning, though in general agreement with extant literature, remain tentative. Second, although post-experiment affective ratings confirmed the aversive nature of the US, and cortical changes produced by conditioning were found, no peripheral measures (e.g. skin conductance) were obtained to confirm conditioning.

4.3. Conclusions and relevance to social cognition

Several manifestations of social cognition, including complex social evaluation, risk-based decision making, and interpersonal attitudes are rooted in (conscious or unconscious)

acquisition of social knowledge, which is often affective in nature. In the present study, we show that socially relevant cues, such as facial stimuli potentially associated with an aversive outcome, are quickly and most likely preattentively differentiated from others predicting safety through a cascade of brain processes in regions that have been implicated in social cognition, such as the fusiform gyrus and the PFC, among others [1]. Whether these mechanisms underline more complex forms of social cognition should be addressed in future studies.

Acknowledgements

The authors wish to thank Drs. A.C. Evans and P. Neelin for providing the Talairach MRI and probability atlases, Dr. V.L. Towle for useful information about electrode coordinates in realistic head geometry, as well as Andrew Hendrick, Megan Zuelsdorff, and Andrew Fox for assistance. This research was supported by grants from the Swiss National Research Foundation (81ZH-52864; DAP), "Holderbank"-Stiftung zur Förderung der wissenschaftlichen Fortbildung (DAP), and NIMH (MH40747, P50-MH52354, MH43454, K05-MH00875; RJD).

References

[1] Adolphs R. Neural systems for recognizing emotion. Current Opinion in Neurobiology 2002;12:169–77.
[2] Amaral DG, Price JL, Pitkanen A, Carmichael ST. Anatomical organization of the primate amygdaloid complex. In: Aggleton JP, editor. The amygdala: neurobiological aspects of emotion, memory and mental dysfunction. New York: Wiley, 1992. pp. 1–66.
[3] Baas JMP, Kenemans JL, Bocker KBE, Verbaten MN. Threat-induced cortical processing and startle potentiation. NeuroReport 2002;13:133–7.
[4] Begleiter H, Platz A. Cortical evoked potentials to semantic stimuli. Psychophysiology 1969;6:91–100.
[5] Buchel C, Dolan RJ. Classical fear conditioning in functional neuroimaging. Current Opinion in Neurobiology 2000;10:219–23.
[6] Buchel C, Morris J, Dolan RJ, Friston KJ. Brain systems mediating aversive conditioning: an event-related fMRI study. Neuron 1998;20:947–57.
[7] Cacioppo JT, Berntson GG, Adolphs R, Carter CS, Davidson RJ, McClintock MK, et al. Foundations in social neurosciences. Cambridge, MA: MIT Press, 2002.
[8] Chapman JL, Chapman JP. The measurement of handedness. Brain and Cognition 1987;6:175–83.
[9] Davidson RJ. Affective style, psychopathology and resilience: brain mechanisms and plasticity. American Psychologist 2000;55:1196–214.
[10] Davidson RJ, Irwin W. The functional neuroanatomy of emotion and affective style. Trends in Cognitive Sciences 1999;3:11–21.
[11] Davis M, Whalen PJ. The amygdala: vigilance and emotion. Molecular Psychiatry 2001;6:13–34.
[12] Eagly AH, Chaiken S. The psychology of attitudes. Fort Worth, TX: Harcourt, 1993.
[13] Ekman P, Friesen WV. Pictures of facial affect. Palo Alto, California: Consulting Psychologist Press, 1976.
[14] Fischer H, Andersson JL, Furmark T, Fredrikson M. Fear conditioning and brain activity: a positron emission tomography study in humans. Behavioral Neuroscience 2000;114:671–80.
[15] Flor H, Birbaumer N, Roberts LE, Feige B, Lutzenberger W, Hermann C, et al. Slow potentials, event-related potentials, gamma-band activity, and motor responses during aversive conditioning in humans. Experimental Brain Research 1996;112:298–312.
[16] Fredrikson M, Wik G, Fischer H, Andersson J. Affective and attentive neural networks in humans: a PET study of Pavlovian conditioning. NeuroReport 1995;7:97–101.
[17] Greenwald AG, Banaji MR. Implicit social cognition: attitudes, self-esteem, and stereotypes. Psychological Review 1995;102:4–27.
[18] Halgren E, Baudena P, Heit G, Clarke JM, Marinkovic K, Clarke M. Spatio-temporal stages in face and word processing. I. Depth-recorded potentials in the human occipital, temporal and parietal lobes. Journal of Physiology 1994;88:1–50.
[19] Halgren E, Raij T, Marinkovic K, Jousmaeki V, Hari R. Cognitive response profile of the human fusiform face area as determined by MEG. Cerebral Cortex 2000;10:69–81.
[20] Haxby JV, Hoffman EA, Gobbini MI. The distributed human neural system for face perception. Trends in Cognitive Sciences 2000;4:223–33.
[21] Hugdahl K. Classical conditioning and implicit learning: the right hemisphere hypothesis. In: Davidson RJ, Hugdahl K, editors. Brain asymmetry. Cambridge, MA: MIT Press, 1995. pp. 235–67.
[22] Hugdahl K, Beradi A, Thompson WL, Kosslyn SM, Macy R, Baker DP, et al. Brain mechanisms in human classical conditioning: a PET blood flow study. NeuroReport 1995;6:1723–8.
[23] Hugdahl K, Nordby H. Hemisphere differences in conditional learning: an ERP-study. Cortex 1991;27:557–70.
[24] Kawasaki H, Adolphs R, Kaufman O, Damasio H, Damasio AR, Granner M, et al. Single-neuron responses to emotional visual stimuli recorded in human ventral prefrontal cortex. Nature Neuroscience 2000;4:15–6.
[25] Knight DC, Smith CN, Stein EA, Helmstetter FJ. Functional MRI of human Pavlovian fear conditioning: patterns of activation as a function of learning. NeuroReport 1999;10:3665–70.
[26] Koenig T, Lehmann D. Microstates in language-related brain potential maps show noun-verb differences. Brain and Language 1996;53:169–82.
[27] Koles ZJ, Flor-Henry P, Lind JC. Low-resolution electrical tomography of the brain during psychometrically matched verbal and spatial cognitive tasks. Human Brain Mapping 2001;12:144–56.
[28] LaBar KS, Gatenby JC, Gore JC, LeDoux JE, Phelps EA. Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. Neuron 1998;20:937–45.
[29] Lancaster JL, Rainey LH, Summerlin JL, Freitas CS, Fox PT, Evans AC, et al. Automated labeling of the human brain—a preliminary report on the development and evaluation of a forward-transformed method. Human Brain Mapping 1997;5:238–42.
[30] LeDoux JE. Emotion circuits in the brain. Annual Review of Neuroscience 2000;23:155–84.
[31] Lehmann D, Skrandies W. Spatial analysis of evoked potentials in man: a review. Progress in Neurobiology 1984;23:227–50.
[32] Marinkovic K, Trebon P, Chauvel P, Halgren E. Localised face processing by the human prefrontal cortex: face-selective intracerebral potentials and post-lesion deficits. Cognitive Neuropsychology 2000;17:187–99.
[33] Montoya P, Larbig W, Pulvermuller F, Flor H, Birbaumer N. Cortical correlates of semantic classical conditioning. Psychophysiology 1996;33:644–9.
[34] Morgan MA, Romanski L, LeDoux JE. Extinction of emotional learning: contribution of medial prefrontal cortex. Neuroscience Letters 1993;163:109–13.
[35] Morris JS, Friston KJ, Dolan RJ. Neural responses to salient visual stimuli. Proceedings of the Royal Society of London–Series B Biological Sciences 1997;264:769–75.
[36] Morris JS, Friston KJ, Dolan RJ. Experience-dependent modulation of tonotopic neural responses in human auditory cortex. Proceedings

of the Royal Society of London–Series B Biological Sciences 1998;265:649–57.
[37] Mulert C, Gallinat J, Pascual-Marqui R, Dorn H, Frick K, Schlattmann P, et al. Reduced event-related current density in the anterior cingulate cortex in schizophrenia. NeuroImage 2001;13:589–600.
[38] Nobre AC, Coull JT, Frith CD, Mesulam MM. Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. Nature Neuroscience 1999;2:11–2.
[39] Pascual-Marqui RD, Lehmann D, Koenig T, Kochi K, Merlo MC, Hell D, et al. Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia. Psychiatry Research: Neuroimaging 1999;90:169–79.
[40] Pascual-Marqui RD, Michel CM, Lehmann D. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. International Journal of Psychophysiology 1994;18:49–65.
[41] Pavlov I. Conditioned reflexes. New York: Dover, 1927.
[42] Perrin F, Pernier J, Bertrand D, Echallier JF. Spherical splines for scalp potential and current density mapping. Electroencephalography and Clinical Neurophysiology 1989;72:184–7.
[43] Phelps EA, O'Connor KJ, Cunningham WA, Funayama ES, Gatenby JC, Gore JC, et al. Performance on indirect measures of race evaluation predicts amygdala activation. Journal of Cognitive Neuroscience 2000;12:729–38.
[44] Pizzagalli D, Shackman AJ, Davidson RJ. The functional neuroimaging of human emotion: asymmetric contributions of cortical and subcortical circuitry. In: Hugdahl K, Davidson RJ, editors. The Asymmetrical Brain. Cambridge, MA: MIT Press, 2002, pp. 511–32.
[45] Pizzagalli D, Lehmann D, Koenig T, Regard M, Pascual-Marqui RD. Face-elicited ERPs and affective attitude: brain electric microstate and tomography. Clinical Neurophysiology 2000;111:521–31.
[46] Pizzagalli D, Regard M, Lehmann D. Rapid emotional face processing in the human right and left brain hemispheres: an ERP study. NeuroReport 1999;10:2691–8.
[47] Pizzagalli DA, Lehman D, Hendrick AM, Regard M, Pascual-Marqui RD, Davidson RJ. Affective judgments of faces modulate early activity (∼160 ms) within the fusiform gyri. NeuroImage 2002;16:663–77.
[48] Ploghaus A, Tracey I, Gati JS, Clare S, Menon RS, Matthews PM, Rawlins JN. Dissociating pain from its anticipation in the human brain. Science 1999;284:1979–81.
[49] Quirk GJ, Armony JL, LeDoux JE. Fear conditioning enhances different temporal components of tone-evoked spike trains in auditory cortex and lateral amygdala. Neuron 1997;19:613–24.
[50] Regan M, Howard R. Fear conditioning, preparedness, and the contingent negative variation. Psychophysiology 1995;32:208–14.
[51] Simons RF, Ohman A, Lang PJ. Anticipation and response set: cortical, cardiac, and electrodermal correlates. Psychophysiology 1979;16:222–33.
[52] Skolnick AJ, Davidson RJ. Affective modulation of eyeblink startle with reward and threat. Psychophysiology, in press.
[53] Skrandies W, Jedynak A. Associative learning in humans: conditioning of sensory-evoked brain activity. Behavioural Brain Research 2000;107:1–8.
[54] Sugawara M, Kitajima S, Kanoh M. Enhancement and diminution of the evoked responses to conditioned stimuli during discrimination conditioning. Neuropsychologia 1977;15:243–8.
[55] Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. New York: Thieme, 1988.
[56] Tong F, Nakayama K, Moscovitch M, Weinrib O, Kanwisher N. Response properties of the human fusiform face area. Cognitive Neuropsychology 2000;17:257–79.
[57] Towle VL, Bolanos J, Suarez D, Tan K, Grzeszczuk R, Levin DN, et al. The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy. Electroencephalography and Clinical Neurophysiology 1993;86:1–6.
[58] Tucker DM. Spatial sampling of head electrical fields: the geodesic sensor net. Electroencephalography and Clinical Neurophysiology 1993;87:154–63.
[59] Weinberger NM. Physiological memory in primary auditory cortex: characteristics and mechanisms. Neurobiology of Learning and Memory 1998;70:226–51.
[60] Wong PS, Bernat E, Bunce S, Shevrin H. Brain indices of nonconscious associative learning. Consciousness and Cognition: an International Journal 1997;6:519–44.
[61] Wong PS, Shevrin H, Williams WJ. Conscious and nonconscious processes: an ERP index of an anticipatory response in a conditioning paradigm using visually masked stimuli. Psychophysiology 1994;31:87–101.