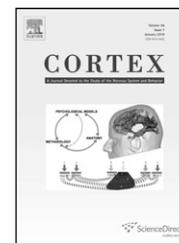




ELSEVIER

available at www.sciencedirect.comjournal homepage: www.elsevier.com/locate/cortex

Research report

Neural response to the behaviorally relevant absence of anticipated outcomes and the presentation of potentially harmful stimuli: A human fMRI study

Louis Nahum^{a,*}, Stéphane R. Simon^b, David Sander^c, François Lazeyras^b and Armin Schneider^a

^aLaboratory of Cognitive Neurorehabilitation, Division of Neurorehabilitation, Department of Clinical Neurosciences, University Hospitals and University of Geneva, Switzerland

^bDepartment of Radiology, Geneva University Hospitals and Centre d'Imagerie Biomédicale, University of Geneva, Switzerland

^cSwiss Centre for Affective Sciences and Department of Psychology, University of Geneva, Switzerland

ARTICLE INFO

Article history:

Received 4 June 2009

Reviewed 12 August 2009

Revised 15 October 2009

Accepted 23 November 2009

Action editor Jordan Grafman

Published online xxx

Keywords:

Behavioral control

Emotion

Outcome processing

Orbitofrontal cortex

fMRI

ABSTRACT

Adaptive behavior requires the ability to react to potentially harmful stimuli, characterized by high negative inherent emotional salience (iES) (e.g., spiders, snakes), and to the unexpected non-occurrence of anticipated events. When presented simultaneously, threatening stimuli and unexpected absence of anticipated outcomes induce distinct electrocortical responses in different time periods. In this study, we used fMRI to test whether processing of the absence of anticipated outcomes (prediction errors) was anatomically dissociated from the processing of iES or whether iES simply modulated activity of areas processing the non-occurrence of anticipated outcomes. Participants saw two alternating pairs of faces and indicated for each pair which one would have a declared target stimulus on its nose. Depending on the condition, the target stimulus was either a spider (high iES stimulus) or a disk (low iES stimulus). The target stimulus switched to the other face after several consecutive correct responses, with the switch being indicated by the appearance of the alternative stimulus (disk when the spider was the declared target; spider when the disk was the declared target). We found that the spider induced stronger activation in visual areas than the disk. By contrast, the absence of anticipated outcomes specifically activated the orbitofrontal cortex (OFC), irrespective of the iES of the outcome stimulus. The findings support a generic role of the OFC in outcome monitoring.

© 2009 Elsevier Srl. All rights reserved.

1. Introduction

Human behavior can be driven by emotional or cognitive aspects of an event, such as potential threats or unexpected

outcomes. The ability to react adaptively to both is essential for survival. Neuronal response to absence of anticipated outcomes mainly involves the orbitofrontal cortex (OFC) (Rolls, 2000; Schneider et al., 2005). Indeed, the OFC is crucial for

* Corresponding author. Service de Neurorééducation, Hôpitaux Universitaires de Genève, 26, av. de Beau-Séjour, CH-1211 Geneva 14, Switzerland.

E-mail address: louis.nahum@hcuge.ch (L. Nahum).

0010-9452/\$ – see front matter © 2009 Elsevier Srl. All rights reserved.

doi:10.1016/j.cortex.2009.11.007

the adaptation of behavior when expected rewards are absent in animals (Rosenkilde et al., 1981; Thorpe et al., 1983; Tremblay and Schultz, 1999, 2000) and humans (Fellows and Farah, 2003; Kringelbach and Rolls, 2004; O'Doherty et al., 2003). OFC activity in response to the absence of expected outcomes seems to be independent of the reward value of the outcome, as shown in human PET (Schnider et al., 2005) and EEG studies (Schnider et al., 2007).

These results indicate that the human OFC has a generic role in the monitoring of outcomes, irrespective of the reward value of outcomes. In agreement with this interpretation, clinical data showed that patients having a lesion of the OFC may confuse memories that relate to ongoing reality and memories that do not; they act on the basis of currently inappropriate anticipations (Schnider, 2003, 2008). Accordingly, we recently found that patients confusing reality after OFC lesions, as manifested in disorientation about time, place, and their current role, had deficient extinction capacity, that is, they failed to abandon previously valid anticipations (Nahum et al., 2009a).

Apart from the ability to adapt thought and behavior when previously correct anticipations no longer apply, an organism also needs the ability to react to potential threats, probably even faster. Neuroimaging studies suggested that the amygdala (Öhman et al., 2007; Sander et al., 2003; Sergerie et al., 2008) and the insular cortex (Mataix-Cols et al., 2008; Mathews et al., 2004) are involved in the processing of threatening stimuli. As compared to neutral stimuli, threat-related stimuli also increase activity in sensory systems: visual and auditory areas (Armony and Dolan, 2002; Pourtois et al., 2004; Phan et al., 2002; Sander et al., 2005; Vuilleumier and Schwartz, 2001; Vuilleumier et al., 2004).

The interaction between these two components of adaptive behavior – reaction to the non-occurrence of an anticipated event versus reaction to a potentially threatening stimulus – is largely unknown. Using event-related potentials (ERP) we found that the presentation of a stimulus with high inherent emotional salience (iES) – a spider – as the outcome elicited a strong, early electrocortical reaction compared to a low iES stimulus at 100–200 msec, irrespective of the need to adapt behavior (Nahum et al., 2009b). Conversely, when the anticipated outcome was absent, indicating that a switch of response was required in the next trial, there was an electrocortical correlate after 200–300 msec, irrespective of the iES of the outcome. Confirmation of the anticipated outcome induced an electrocortical correlate after 400 msec, which was independent of the iES of the outcome, too. Source estimation revealed that the distinction between the high and low iES of the outcome was mediated by an anterior medial temporal region and by visual areas. Processing of the behaviorally relevant absence of the declared target outcome was mediated by orbitofrontal areas and the presence of the target outcome by the posterior medial temporal cortex. The findings indicated that the occurrence of a potentially harmful stimulus elicits an earlier electrocortical response than the behaviorally relevant non-occurrence of an anticipated outcome and that the processing of such events is temporally and spatially distinct.

In the present study, we used event-related fMRI and a similar experimental paradigm as in the evoked potential

study (Nahum et al., 2009b) to test with high spatial resolution whether the iES of an outcome anatomically interacts with the processing of its learned behavioral relevance. Based on our previous PET study using neutral stimuli (Schnider et al., 2005), we expected that the OFC would monitor outcomes depending on the need to subsequently adapt behavior. We hypothesized that a potentially threatening stimulus would induce activity in the amygdala, the insula, and visual areas (Phelps and LeDoux, 2005) but would not influence the way in which absence of anticipated outcomes is processed.

2. Materials and methods

2.1. Participants

Sixteen right-handed students (8 males, 8 females; 24.7 ± 3.2 years, 19–29 years) with no history of neurological or psychiatric disease gave written, informed consent to participate in the study. None had participated in the previous ERP study (Nahum et al., 2009b). The study was approved by the ethical Committee of the University Hospital of Geneva.

Prior to the experiment, participants rated the iES of the two types of target stimuli that were used in the experiment: a spider and a disk. They answered six questions each for assessing anxiety, disgust, and distress induced by those stimuli on a Likert scale from 0 (not at all) to 9 (extreme). All participants declared not to be spider phobic, which was confirmed by low scores in a French version of the Fear of Spider Questionnaire (FSQ, Szymanski and O'Donohue, 1995) (mean score \pm standard deviation – SD, 27.37 ± 11.21). Men and women did not differ on FSQ scores [$F(1, 14) = .27, p = .61$].

2.2. Stimuli and experimental task

During scanning, participants performed two conditions of a reversal learning task previously used in a ERP study, adapted to the fMRI environment (Nahum et al., 2009b). Choice stimuli were two pairs of neutral faces from Ekman and Friesen (1975), alternately presented on a computer screen. Subjects had to indicate in each trial, which one of the two faces would have a specific target stimulus (outcome) on its nose. Depending on the task condition, the target outcome was either a black, schematic spider (high iES outcome) or a black disk (low iES outcome).

Fig. 1 illustrates the design of the task for the condition with the spider as a target outcome. Trials began with the presentation of a pair of faces (choice stimulus). Participants had to choose one of the two faces within 2 sec by pressing with the index or middle finger of the right hand the button corresponding to the side of the chosen face. After the choice, the non-chosen face disappeared and a fixation cross appeared on the nose of the chosen face for 1 sec. Then, the outcome stimulus was displayed on the nose of the chosen face for 1.5 sec. After a varying interval of 3 ± 1 sec with presentation of a blank screen, the next trial with the alternate pair of faces started. Presentation was controlled using e-prime (©Psychology Software Tools, Pittsburgh, PA).

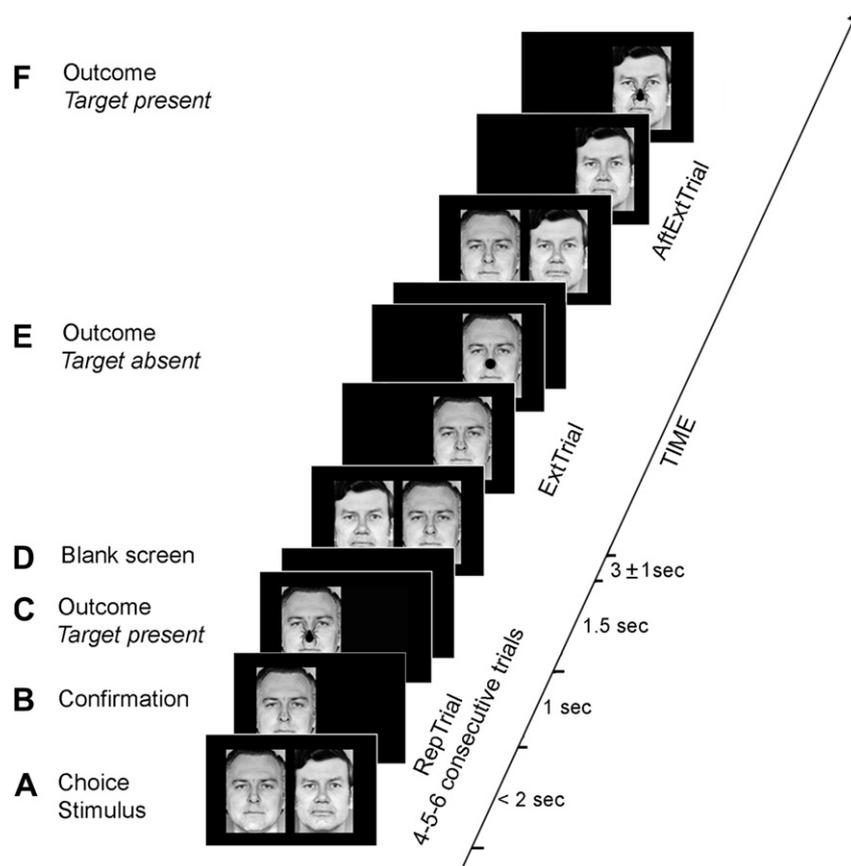


Fig. 1 – Sequential order of trials in the course of the experiment demonstrating the two different outcome types (presence or absence of the target) when the target was the spider. Every trial consisted of the same three steps: (A) presentation of the first pair of faces (choice stimulus); participants had to predict by button press which one of the two faces would have a spider on its nose. (B) After the choice, a cross appeared on the respective face (confirmation), then the outcome was presented on the nose of the chosen face: a spider (target outcome) if the choice was correct (C) or a disk (non target outcome), indicating absence of the target outcome. (D) A blank screen of 3 ± 1 sec followed the outcome. Then, the second pair of faces (for clarity, not illustrated) was presented and participants had also to predict which one of the two faces would have a spider on its nose. (E) For each pair, after four to six consecutive correct responses (RepTrial), a disk indicating absence of the target outcome (spider) was presented [“Extinction” trial (ExtTrial)]. (F) After the ExtTrial, participants had to choose the other face (AftExtTrial). Note that incorrect choice put back the counter of consecutive correct responses to zero.

Subjects were informed that the target would normally reappear on the same face, but that it would occasionally switch to the other face. In this case, absence of the target outcome would be signaled by the appearance of the alternative outcome on the nose of the chosen face (disk when the spider was the declared target outcome; spider when the disk was the declared target outcome). The target outcome switched to the other face after every fourth to sixth consecutive correct response with switches occurring independently in the two pairs of faces. Thus, subjects had to continuously update their idea of the face associated with the target outcome for both pairs.

The task constitutes a variant of a reversal learning task. However, it emphasizes the extinction component (i.e., absence of the expected outcomes) by the fact that the non-chosen face disappears from the screen after the subject’s response, which – in the case of “extinction trials” – creates the impression that the target stimulus is absent from the chosen face rather than switched to the other face (in reversal

learning tasks, the alternative, “correct” choice is presented together with the incorrectly chosen stimulus).

Trials in which the target outcome was the same as in the previous trial were called repetition trials (RepTrials). Trials in which the target was absent, as indicated by presentation of the alternative outcome, were called extinction trials (ExtTrials). The effectiveness of extinction was measured by the performance in the immediately following trial with the same pair of faces, called “after-extinction” trial (AftExtTrials).

The number of RepTrials was sufficiently variable for participants not to anticipate target switches: questioning after the experiment revealed that they were unaware that ExtTrials would always appear after four to six RepTrials.

Two scanning sessions were made: one with the spider as the target and the other with the disk as the target. At the beginning of each session, the target outcome to be looked for was announced, instructions were repeated, and 20 practice trials using the respective target were performed. For both sessions, subjects made a total of three blocks with 70 trials

each, separated by a resting phase of 1 min in which participants were told to look at a white fixation cross on a black background. This resting phase was used as a baseline condition in order to discriminate between positive and negative activations. The order of sessions was randomized; half of participants started with a session having the spider as the target and the other half with a session having the disk as target. The scanner stopped only in between the two scanning sessions. In total, participants made six blocks of 70 trials, that is, 420 trials. Total scanning time was 56 min. There were approximately 60% RepTrials, 20% ExtTrials and 20% AftExtTrials.

2.3. Analysis of behavioral data

Repeated-measures ANOVAs were made on reaction times and accuracy with the iES of the previous outcome (high or low) and the trial type of interest (RepTrial, ExtTrial, AftExtTrial) as the repeated within-participant factors. Correlation analyses were performed between reaction time and accuracy for the different questionnaire scores (FSQ, subjective ratings) using Pearson correlations.

2.4. Image acquisition and post-processing

Image data were acquired on a Siemens 3 T Trio MRI scanner (Siemens Medical Solutions, Erlangen, Germany). Structural images were acquired with high-resolution T1-weighted sequence for subsequent anatomic localization. T2*-weighted echo planar imaging (EPI) was used to obtain blood oxygenation level dependent contrast (BOLD). Two time series (including the resting periods) containing 890 volumes of 30 slices (repetition time/TR = 2 sec; echo time/TE = 30; flip angle = 90°; spatial resolution 2 × 1.7 × 4 mm; IPAT factor = 2) were acquired for each participant. Functional images were then processed using SPM5 software (www.fil.ion.ucl.ac.uk). The processing step included slice timing correction in each volume, time-series rigid body realignment (Ashburner and Friston, 2007), co-registration with the T1 anatomical reference image, normalization to the segmented T1 image (Ashburner and Friston, 2005), spatially smoothing (6 mm isotropic FWHM Gaussian kernel) and finally high-pass (180 sec cutoff) and low-pass filtering to remove signal drift from the scanner and physiological artifacts.

2.5. Statistical analysis of images

Statistical analysis was performed within the general linear model approach (Kiebel and Holmes, 2007; Worsley et al., 1992) using the SPM5 software. Events of interest corresponded only to the presentation of outcomes (Fig. 1). Each event type was modeled as a separate regressor (duration of 0 sec) convolved with the canonical SPM hemodynamic response function (HRF) and HRF temporal and dispersion derivatives functions. According to the main question of the study (neural correlates of the behavioral relevance of outcomes and its interaction with the iES of the outcome stimulus), event types of interest were RepTrials and ExtTrials, with either the spider or the disk as the target. Analysis also showed that neither the comparison between the first

trial after an extinction trial (AftExtTrials) and all following confirmatory target outcomes (RepTrials), nor the comparison between AftExtTrials and the fourth confirmatory target outcomes showed significant differences of brain activation, so that AftExtTrials were not considered further. Omissions were very rare (<1%) and were not modeled. The resting phases were entered as box-car covariate for creation of masks (see below). Finally, regressors including error trials and 6 regressors from the motion corrections were entered as a covariate of no interest. For each condition, a characterization of linear time effect (first order time modulation), which controls for linear change of BOLD response over time, was used to test for habituation effect within blocks for all types of outcomes.

Statistical parametric maps were first generated from linear contrasts between canonical HRF and HRF temporal and dispersion derivatives functions corresponding to event types of interest in each participant, and entered in a second stage for random effect analysis. Six one-sample *t*-tests were calculated for each participant: two investigating the main effect of trial type independently of iES: ExtTrial > RepTrial, RepTrial > ExtTrial; two investigating the main effect of iES of the outcome independently of trial type: Spider (all spider outcomes from all trials) > Disk (all disk outcomes from all trials), Disk > Spider; two contrasts evaluating possible interactions between iES of the outcome and trial type: [(ExtTrial with the spider > RepTrial with the spider) > (ExtTrial with the disk > RepTrial with the disk)], [(RepTrial with the spider > ExtTrial with the spider) > (RepTrial with the disk > ExtTrial with the disk)].

All analyses were first performed across the whole cerebrum. SPMs of the *t* statistic (*df* = 14) at each voxel were thresholded at $p < .05$ FWER-corrected (“family wise error rate”) for multiple comparison. The interaction contrasts were also tested with a more liberal statistical threshold ($p < .001$, uncorrected) in order to confirm potentially null effects. We only retained activation foci which corresponded to a significant non-isotropic adjusted cluster-level thresholded at $p < .05$. In addition, we used masks for each contrast of interest (except for the interaction’s contrasts) to disentangle positive and negative hemodynamic responses. Those masks consisted of the event type of interest minus the resting phase and were thresholded at $p < .05$ (uncorrected). They were first generated individually for each participant and then used as inclusive masks in the second-level analysis (random effect).

In a second step, we conducted region of interest (ROIs) analyses for contrasts showing statistically significant differences in order to test our hypotheses regarding main effects of behavioral relevance and iES with a less conservative correction for multiple comparisons than the one used in the whole brain analysis (see below). ROIs were defined a priori on the basis of our hypotheses in the amygdala, insula, hippocampus, orbitofrontal and ventro-occipital cortices, and anatomically generated using the WFU Pickatlas software toolbox (Maldjian et al., 2003, 2004) and the Anatomical Automatic Labelling (AAL) Atlas (Tzourio-Mazoyer et al., 2002), which is based on the Talairach Daemon database (Talairach and Tournoux, 1988). MNI coordinates were finally converted as recently described by Lancaster et al. (2007). The ROI for the OFC included Brodmann’s areas

(BAs) 11, 13 and medial parts of area 47/12. The ROI for the insular cortex included BAs 13 and 14. The Marseille Region of Interest Toolbox software package (MarsBaR, www.sourceforge.net/projects/marsbar) was then used to calculate, in all ROIs for each participant, the BOLD response in response to a specific event of interest. A conjunction analysis (Friston et al., 1999) of activations in the OFC associated with ExtTrial [(ExtTrial with the Spider – RepTrial with the Spider), (ExtTrial with the Disk – RepTrial with the Disk)] was undertaken to test whether significant activation in the OFC is common to both ExtTrial with the spider and ExtTrial with the disk. The *p*-values were Bonferroni corrected by the number of tested ROIs (7); only *p*-values below .007 (.05/7) were considered significant.

3. Results

3.1. Subjective ratings

Paired *t*-test demonstrated that the schematic spider was perceived more negatively than the disk: it induced more anxiety (Likert scale, mean \pm SD, spider: 2 ± 1.4 ; disk: $.4 \pm .8$), disgust (spider: 3 ± 2.6 ; disk: $.06 \pm .25$) and distress (spider: 2.1 ± 2 ; disk: $.4 \pm 1.5$) (all comparisons, $p < .001$). The spider induced more disgust than anxiety and distress [$F(2, 30) = 3.4$, $p = .046$]. Men and women did not differ on scores of anxiety, disgust and distress induced by the spider ($p > .05$).

3.2. Behavioral data

As expected, participants easily learned the behavioral relevance of outcomes; they rapidly adapted behavior after extinction trials: they made only $3.5 \pm 2.4\%$ errors (mean \pm SD) while repeatedly seeing the face associated with the target outcome (RepTrials) and $5.4 \pm 5.7\%$ errors after extinction trials (AftExtTrials). Repeated-measures ANOVAs with iES of the previous outcome stimulus (spider, high; disk, low) and trial type (AftExtTrials, RepTrials, ExtTrials) as within-participant factors on the mean error rate and reaction times revealed a significant effect of trial type on reaction times (see Table 1): participants responded faster after an extinction trial (AftExtTrials, 877 ± 128 msec, mean \pm SD) than after a confirmatory trial [RepTrials, 914 ± 139 msec; $F(1, 15) = 10.18$, $p = .006$]. By contrast, there was no significant interaction between iES and trial type and main effect of iES ($p > .05$).

Table 1 – Behavioral results.

Trial type	iES of the previous outcome stimulus	Mean proportion of errors (SD)	Mean reaction times (in msec; SD)
RepTrial	High	3.6 (2)	911 (121)
	Low	3.3 (2)	917 (159)
ExtTrial	High	–	912 (132)
	Low	–	916 (157)
AftExtTrial	High	4.7 (5.4)	873 (98)
	Low	6.1 (6.1)	882 (155)

Reaction times in ExtTrials (914 ± 126 msec) and RepTrials (914 ± 123 msec) were similar, confirming that subjects had no differential expectations about the likelihood of upcoming reversal. Performance did not differ between participants starting with the condition having the spider as target and those starting with the circle as target [$F(2, 14) = .04$, $p = .83$]. FSQ (Szymanski and O'Donohue, 1995) scores and subjective measures of anxiety, disgust and distress in response to the spider did not correlate with task performance (all *p*-values $> .05$).

3.3. fMRI data

3.3.1. Behavioral relevance: presence versus absence of anticipated outcomes

To identify brain regions responding to the behaviorally relevant absence of anticipated outcomes, we compared whole brain activity elicited by outcome stimuli from ExtTrials relative to outcome stimuli from RepTrials (ExtTrial $>$ RepTrial), irrespective of the iES of the outcome. Fig. 2 shows stronger activation in frontal and parietal areas, in the right lateral OFC and in bilateral insula and adjacent posterior lateral OFC for ExtTrials compared to RepTrials. There was also activation in the cerebellum. In addition, the whole brain analysis showed significant increase in the pallidum and in occipital areas (Table 2). A ROI analysis within the OFC confirmed the stronger activation in ExtTrials compared to RepTrials in the right lateral OFC ($x, y, z, = 25, 49, 2$; $t(15) = 8.28$; $p = .0001$) and in bilateral insula and adjacent posterior lateral OFC [Left: $x, y, z, = -30, 20, -3$, $t(15) = 9.64$; Right: $x, y, z, = 44, 17, 1$, $t(15) = 7.05$, all $p < .0001$].

Fig. 2C indicates that RepTrials induced stronger activity than ExtTrials in the right superior frontal gyrus, in the right insula, and in the hippocampus bilaterally. The whole brain analysis also indicated additional differences in, the left putamen, the left middle frontal gyrus and the right post-central gyrus (Table 2). In the ROI analysis centered on the hippocampus, there was stronger activation of the left [$x, y, z, = -29, -31, -6$, $t(15) = 11.99$, $p < .0001$] and right [$x, y, z, = 29, -31, -5$, $t(15) = 6.66$, $p = .001$] hippocampus in RepTrials than ExtTrials. The conjunction analysis [(ExtTrial with the Spider – RepTrial with the Spider), (ExtTrial with the Disk – RepTrial with the Disk)] confirms that the OFC was commonly activated by both ExtTrials with the Spider and ExtTrials with the Disk. Fig. 3 shows that the right superior lateral OFC [$x, y, z, = 27, 56, 8$, $t(30) = 3.2$, $p < .001$] and left [$x, y, z, = -31, 23, 1$, $t(30) = 3.8$, $p < .001$] and right [$x, y, z, = 31, 23, 2$, $t(30) = 3.9$, $p < .001$] insula and adjacent posterior lateral posterior OFC were significantly activated, irrespective of the iES of the stimulus. No habituation effect over time (i.e., inclusion of linear time modulation in the statistical analysis of images) was found concerning behavioral relevance.

3.3.2. iES: spider versus disk

To investigate brain regions responding to the iES of the outcome stimulus, we used whole brain analysis to compare the activation in response to the spider with activation in response to the disk (Spider $>$ Disk), irrespective of the behavioral relevance of the outcome. Fig. 2 shows that the spider induced stronger activation in visual associative areas

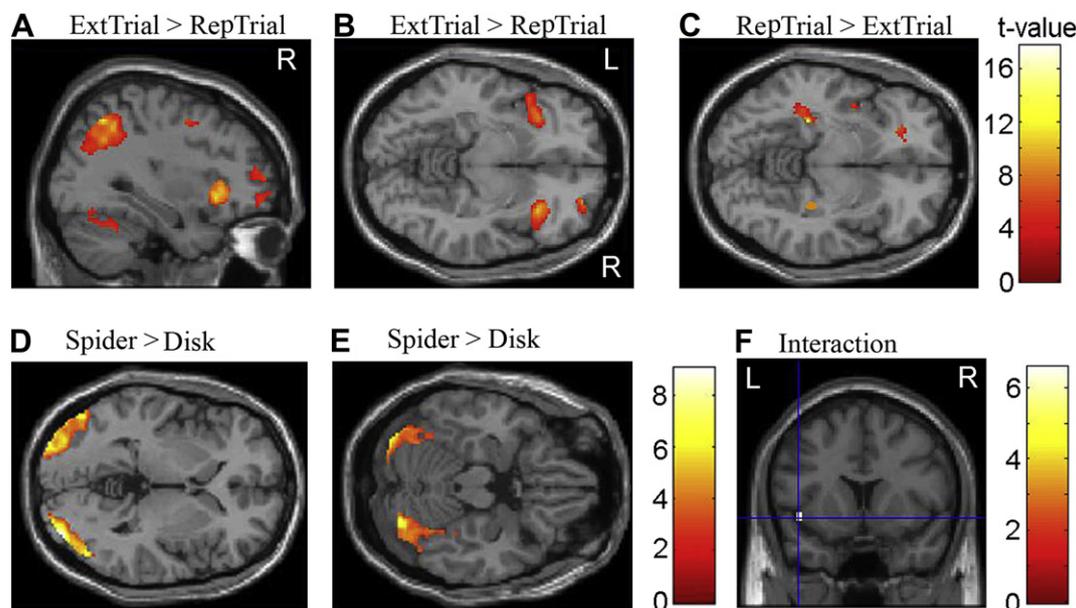


Fig. 2 – (A, B) Areas of activation revealed by the whole brain analysis for the contrast ExtTrial > RepTrial with sagittal and axial slices. (C) Areas of activation for the contrast RepTrial > ExtTrial with axial slice. (D, E) Areas of activation for the contrast Spider > Disk with axial slices. (F) Areas of activation for the interaction contrast [(RepTrial with the spider > ExtTrial with the spider) > (RepTrial with the disk > ExtTrial with the disk)] with sagittal slice. All $p < .05$ FWER-corrected with k_E at $p < .05$, except for the interaction contrast: $p < .001$, uncorrected.

(Fig. 2D) and in the primary visual cortex (Fig. 2E) than the disk. When habituation over time was taken into account, the spider still elicited stronger activation than the disk in visual areas, plus in the cerebellum (Table 2). No significant effects were detected in the whole brain for the contrast Disk > Spider.

In contradiction to our initial hypothesis, ROI analyses in the amygdala and the insula yielded no significant effect for the contrast Spider > Disk, even when a linear time modulation was used.

3.3.3. Interaction between iES and behavioral relevance

The interaction contrast [(ExtTrial with the spider > RepTrial with the spider) > (ExtTrial with the disk > RepTrial with the disk)] showed no significant effect in the whole brain analysis, even when the t statistic at each voxel was uncorrected for multiple comparisons and when a linear time modulation was used; activation in ExtTrials did not differ according to iES of the target outcome. Conversely, Fig. 2F shows that RepTrials were associated with significantly stronger activation [$t(15) = 7.31$, $p < .001$, uncorrected] in the left anterior insula than ExtTrials when the outcome was a spider than when it was a disk for the interaction contrast: [(RepTrial with the spider > ExtTrial with the spider) > (RepTrial with the disk > ExtTrial with the disk)].

4. Discussion

This study shows that two crucial components of adaptive behavior anatomically dissociate: behaviorally relevant absence of anticipated outcomes – requiring adaptation of behavior on the next trial – activated the OFC, irrespective of

the iES of the outcome (high iES: spider, low: disk). Conversely, the spider stimulus – a proxy for a potentially harmful stimulus – enhanced activation in visual areas. This result is in agreement with a previous study using a similar paradigm and evoked potentials, which indicated that these components of adaptive behavior have temporally and also anatomically distinct representations (Nahum et al., 2009b).

The non-occurrence of anticipated outcomes (ExtTrials), requiring adaptation of subsequent behavior, was associated with activation in the lateral OFC. This result confirms previous evidence that the OFC processes deviations from expected outcomes and hence plays a key role in behavioral flexibility (Cools et al., 2002; Hampshire and Owen, 2006; Schnider et al., 2005; Schoenbaum et al., 2007). Animal and human lesion studies showed that orbitofrontal damage did not impair acquisition of an association but impaired switching to the alternative stimulus in reversal learning (Bohn et al., 2003; Butter et al., 1963; Fellows and Farah, 2003).

In an earlier study using PET, we had obtained medial OFC activation (pole and area 13) while subjects anticipated and monitored neutral outcomes (Schnider et al., 2005), whereas the present study using fMRI yielded lateral OFC activation in extinction trials. The difference between the two studies is partly explained by the imaging method: H2[15]O-PET, with its temporal resolution of 45 sec, which integrates multiple trials (confirmatory and extinction trials) and does not distinguish between anticipation and outcome monitoring, while fMRI allows to analyze single events and to focus on the moment of outcome processing. Conversely, fMRI is less sensitive to posterior medial OFC activation because of typical susceptibility artifacts (Ojemann et al., 1997). Notwithstanding this potential limitation of fMRI, the present study shows an anatomical

Table 2 – Brain regions of significant activations revealed by whole cerebrum between conditions (trial type and iES of the outcome stimulus) and interaction contrast.

Contrast	Side	x	y	z	k_E	T	p
Brain region							
<i>ExtTrial > RepTrial</i>							
Lateral superior OFC (BA 10)	R	25	49	2	117	8.28	.005
Inferior frontal gyrus pars orbitalis (BA 47)	L	-45	12	5	1111	9.48	.001
Anterior insula	L	-31	19	4	-	9.39	.001
	R	31	19	1	496	10.37	<.001
Precentral gyrus (BA 6)	L	-38	-2	30	1497	11.57	<.001
Precentral gyrus (BA 6)	R	39	6	36	47	7.64	.012
Middle frontal gyrus (BA 9)	R	36	30	26	185	7.75	.01
Superior parietal lobule (BA 7)	R	30	-62	44	1402	11.87	<.001
Superior occipital gyrus (BA 19)	R	30	-61	31	-	8.34	.005
Inferior parietal lobule (BA 39)	L	-35	-55	34	2470	13.26	<.001
Superior occipital gyrus (BA 19)	L	-5	-67	39	-	12.27	<.001
Dorsal frontal gyrus (BA 6)	L	-2	11	47	-	17.22	<.001
Pallidum	L	-12	4	4	287	10.04	<.001
Cerebellum	L	-32	-53	-28	212	7.21	.025
	R	5	-71	-20	180	7.84	.009
<i>RepTrial > ExtTrial</i>							
Hippocampus	L	-29	-31	-6	2761	11.99	<.001
	R	29	-31	-5	61	6.66	.038
Posterior insula	R	45	-9	17	163	17.73	<.001
Superior frontal gyrus (BA 9)	R	12	49	22	643	11.99	<.001
Middle frontal gyrus (BA 9)	L	-12	41	28	619	7.55	.012
Postcentral gyrus (BA 43)	R	56	-11	19	42	7.23	.018
Putamen	L	-31	-17	9	87	6.64	.039
<i>High iES > Low iES</i>							
Middle occipital gyrus (BA 19)	R	30	-86	-2	927	9.5	.001
Inferior occipital gyrus (BA 18)	L	-29	-93	-5	763	7.49	.021
Fusiform gyrus (BA 19)	L	-42	-73	-14	-	6.88	.048
<i>High iESo > Low iES with time modulator</i>							
Middle occipital gyrus (BA 19)	R	32	-86	-3	1460	9.03	.002
Inferior occipital gyrus (BA 18)	R	47	-69	-5	-	7.3	.028
Fusiform gyrus (BA 19)	L	-42	-73	-14	1883	6.98	.045
Cerebellum	L	-29	-62	-15	1883	7.84	.012
	R	25	-71	-20	1460	7.66	.016
<i>(RepTrial/High iES > ExtTrial/High iES) > (RepTrial/Low iES > ExtTrial/Low iES)</i>							
Anterior insula	L	-43	12	-1	122	7.31*	<.001*

Coordinates x, y, and z are given in Talairach space. All results at $p < .05$ FWER-corrected with k_E at $p < .05$ unless indicated otherwise. * $p < .001$ uncorrected.

dissociation between the confirmation and negation of anticipated outcomes. Whole brain analysis additionally identified stronger activity in ExtTrials than RepTrials in several prefrontal areas, bilateral anterior insular cortex, as well as in parietal and occipital lobe. These regions seem to be functionally linked to the OFC in stimulus-reinforcement switches during reversal learning (Budhani et al., 2007; den Ouden et al., 2009; O'Doherty et al., 2003; Remijne et al., 2005, 2006).

The conjunction analysis and the absence of a significant interaction between the processing of the non-occurrence of anticipated outcomes (ExtTrials) and the processing of the iES of the outcome stimulus indicate that the OFC monitors outcomes independently of their emotional arousal. These results are in line with previous studies showing that mildly painful stimuli (Schiller et al., 2008), monetary penalties (O'Doherty et al., 2003), abstract punishment (O'Doherty et al., 2001), angry faces (Kringelbach and Rolls, 2003) or purely neutral stimuli (Schnider et al., 2005) signaling absence of

predicted outcomes activate the OFC or the ventromedial prefrontal cortex. Furthermore, studies using probabilistic (Remijne et al., 2005) or all-or-nothing (Schnider et al., 2007) outcome contingencies found that the OFC processed absence of predicted outcomes.

The importance of the OFC in the monitoring of outcomes becomes absolutely evident in patients who confuse reality in thinking after an acute lesions of the posterior medial OFC: they continue to act according to currently irrelevant memories and insist on action plans (anticipations) which do not pertain to present reality, a state called behaviorally spontaneous confabulation (Schnider, 2003, 2008). These patients are also disoriented regarding current time, place and situation (Schnider et al., 1996). A recent study showed that disorientation and behaviorally spontaneous confabulation are indeed strongly associated with the inability to integrate the absence of anticipated outcomes, that is, a failure of extinction (Nahum et al., 2009a).

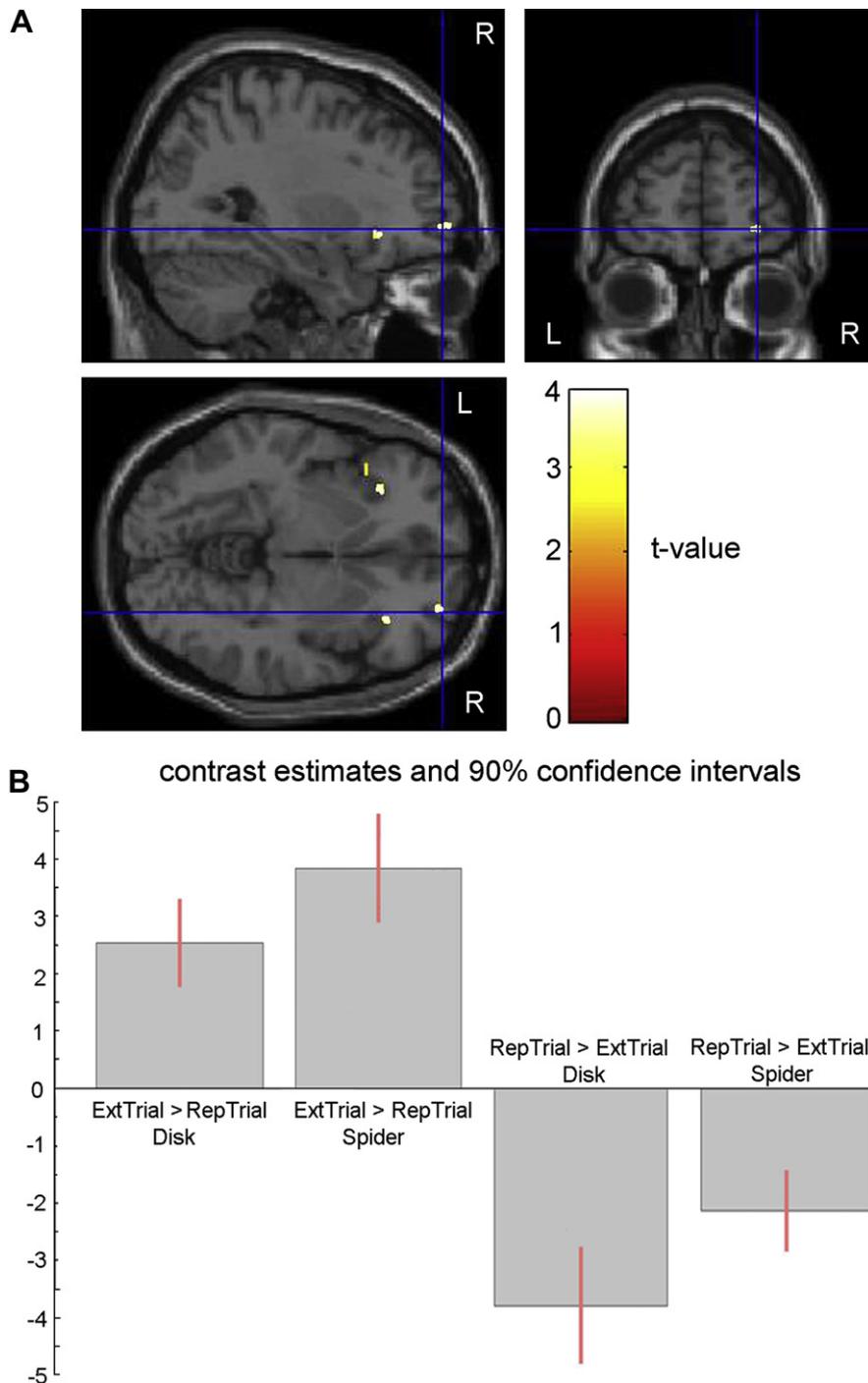


Fig. 3 – (A) Significant areas of activation in the OFC revealed by the conjunction analysis of contrasts (ExtTrial with the Disk > RepTrial with the Disk), (ExtTrial with the Spider > ExtTrial with the Spider) ($p < .001$, uncorrected). **(B)** Respective contrast estimates and 90% confidence intervals in the OFC for contrasts: ExtTrial with the Disk > RepTrial with the Disk; ExtTrial with the Spider > RepTrial with the Spider; RepTrial with the Disk > ExtTrial with the Disk; RepTrial with the Spider > ExtTrial with the Spider.

Confirmation of anticipated outcomes (RepTrials) induced bilateral hippocampal activation, which was independent of the iES of the outcome. This finding is compatible with two recent studies on alternation learning, which showed hippocampal activation when participants were requested to store the outcome of the ongoing trial to select the response in the

next trial (Budhani et al., 2007; Schneider et al., 2005) and with source estimations in our evoked potential study using the same paradigm (Nahum et al., 2009b).

In contrast to the activity of the OFC and hippocampus, the activity in the left anterior insular cortex was influenced by the iES of the outcome, as documented by a significant interaction

between RepTrials and iES of the outcome in this area: there was stronger activation in RepTrials than in ExtTrials only when the outcome was a spider, but not when it was a disk. This result is compatible with studies showing that the anterior insula is involved in the processing of varieties of disgust-related stimuli (Calder, et al., 2001; Jabbi et al., 2008; Stark et al., 2007) and in the subjective risk prediction during decision making (Preuschoff et al., 2008). These results suggest that the anterior insula is critical for interoceptive awareness (Craig, 2002, 2004). In our experiment, it is likely that the presence of anticipated spiders on the nose of faces led to more somatic feelings and disgust-related processing than the presence of anticipated disks.

While iES did not alter the distinctive processing of the absence of anticipated outcomes – there was no significant interaction between ExtTrials and the iES of the outcome – iES independently influenced brain responses: the spider induced markedly greater activation in the left fusiform gyrus and in the left inferior occipital and the right middle occipital gyri. Furthermore, the results demonstrated decreasing activation in visual areas and the cerebellum with repeated exposure to the spider, but not to the disk. These activations are unlikely to be due to differences of visual properties between the spider and the disk, as the outcomes were particularly simple visual stimuli in comparison to the faces, on which they were presented. Also, the outcome stimuli were matched with regards to size and luminosity. Our results are compatible with several studies showing that spider pictures elicit stronger activation in the cerebellum (Caseras et al., 2009) and in the visual areas than neutral pictures, in particular in the left middle occipital gyrus (Schienle et al., 2005), right primary visual cortex (Caseras et al., 2009; Dilger et al., 2003) or the bilateral fusiform gyrus (Schienle et al., 2007) in healthy subjects. From an evolutionary perspective, a spider should be rapidly detected and elicit a defensive response (Öhman and Mineka, 2001; Wendt et al., 2008) possibly driven by the amygdala modulating cortical sensory processing (Armony and Dolan, 2002; Pourtois et al., 2004; Williams et al., 2006). In the present study, however, amygdala and insula responses to spiders as compared to disks did not reach statistical significance, a result similar to some earlier fMRI and PET studies using spider stimuli in non-phobic subjects (Dilger et al., 2003; Wendt et al., 2008; Wik et al., 1993). The absence of activation in amygdala and insular cortex may reflect a relatively weak emotional response of our study participants to spiders – quite unlike subjects suffering from spider phobia, but also the fact that the behaviorally relevant target outcome (disk or spider) was explicitly specified. Nonetheless, even in this study population, spider stimuli were markedly differently processed than the neutral stimuli.

To conclude, our findings confirm that the iES of an outcome is processed separately from its learned behavioral relevance, in which the OFC plays a key role. These results give further support to the hypothesis that the OFC constitutes a generic outcome monitoring system.

Acknowledgments

We thank Dimitri van De Ville and Yann Cojan for technical advice. The study was supported by Swiss National Science

Foundation grant no. 32000-113436 to A.S and by the Center for Biomedical Imaging (CIBM) of Lausanne and Geneva.

REFERENCES

- Armony JL and Dolan RJ. Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia*, 40: 817–826, 2002.
- Ashburner J and Friston KJ. Unified segmentation. *NeuroImage*, 26: 839–851, 2005.
- Ashburner J and Friston K. Rigid body registration. In Friston KJ, Ashburner J, Kiebel SJ, Nichols TE, and Penny WD (Eds), *Statistical Parametric Mapping: the Analysis of Functional Brain Images*. Academic Press, 2007: 49–62.
- Bohn I, Gierler C, and Hauber W. Orbital prefrontal cortex and guidance of instrumental behavior in rats under reversal conditions. *Behavioral Brain Research*, 143: 49–56, 2003.
- Budhani S, Marsh AA, Pine DS, and Blair RJR. Neural correlates of response reversal: considering acquisition. *NeuroImage*, 34: 1754–1765, 2007.
- Butter CM, Mishkin M, and Rosvold HE. Conditioning and extinction of a food-rewarded response after selective ablations of frontal cortex in Rhesus monkeys. *Experimental Neurology*, 7: 65–75, 1963.
- Calder AJ, Lawrence AD, and Young AW. Neuropsychology of fear and loathing. *Nature Review Neuroscience*, 2: 352–363, 2001.
- Caseras X, Giampietro V, Lamas A, Brammer M, Vilarroya O, Carmona S, et al. The functional neuroanatomy of blood-injection-injury phobia: a comparison with spider phobics and healthy controls. *Psychological Medicine*, 13: 1–10, 2009.
- Cools R, Clark L, Owen AM, and Robbins TW. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *Journal of Neuroscience*, 22: 4563–4567, 2002.
- Craig AD. How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Review Neuroscience*, 3: 655–666, 2002.
- Craig AD. Human feelings: why are some more aware than others? *Trends in Cognitive Sciences*, 8: 239–241, 2004.
- den Ouden HEM, Friston KJ, Daw ND, McIntosh AR, and Stephan KE. A dual role for prediction error in associative learning. *Cerebral Cortex*, 19: 1175–1185, 2009.
- Dilger S, Straube T, Mentzel H-J, Fitzek C, Reichenbach JR, Hecht H, et al. Brain activation to phobia-related pictures in spider phobic humans: an event-related functional magnetic resonance imaging study. *Neuroscience Letters*, 348: 29–32, 2003.
- Ekman P and Friesen W. *Unmasking the Human Face: a Guide to Recognizing Emotions from Facial Expressions*. Englewood Cliffs, NJ: Prentice-Hall, 1975.
- Fellows LK and Farah MJ. Ventromedial frontal cortex mediates affective shifting in humans: evidence from a reversal learning paradigm. *Brain*, 126: 1830–1837, 2003.
- Friston KJ, Holmes AP, Price CJ, Büchel C, and Worsley KJ. Multisubject fMRI studies and conjunction analyses. *NeuroImage*, 10: 385–396, 1999.
- Hampshire A and Owen AM. Fractionating attentional control using event-related fMRI. *Cerebral Cortex*, 16: 1679–1689, 2006.
- Jabbi M, Bastiaansen J, and Keysers C. A common anterior insula representation of disgust observation, experience and imagination shows divergent functional connectivity pathways. *PLoS ONE*, 3: e2939, 2008.
- Kiebel SJ and Holmes AP. The general linear model. In Friston KJ, Ashburner J, Kiebel SJ, Nichols TE, and Penny WD (Eds), *Statistical Parametric Mapping: the Analysis of Functional Brain Images*. Academic Press, 2007: 101–125.

- Kringelbach ML and Rolls ET. Neural correlates of rapid reversal learning in a simple model of human social interaction. *NeuroImage*, 20: 1371–1383, 2003.
- Kringelbach ML and Rolls ET. The functional neuroanatomy of the human orbitofrontal cortex: evidence from neuroimaging and neuropsychology. *Progress in Neurobiology*, 72: 341–372, 2004.
- Lancaster JL, Tordesillas-Gutierrez D, Martinez M, Salinas F, Evans A, Zilles K, et al. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Human Brain Mapping*, 28: 1194–1205, 2007.
- Maldjian JA, Laurienti PJ, and Burdette JH. Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage*, 21: 450–455, 2004.
- Maldjian JA, Laurienti PJ, Kraft RA, and Burdette JH. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, 19: 1233–1239, 2003.
- Mataix-Cols D, An SK, Lawrence NS, Caseras X, Speckens A, Giampietro V, et al. Individual differences in disgust sensitivity modulate neural responses to aversive/disgusting stimuli. *European Journal of Neuroscience*, 27: 3050–3058, 2008.
- Mathews A, Yiend J, and Lawrence AD. Individual differences in the modulation of fear-related brain activation by attentional control. *Journal of Cognitive Neuroscience*, 16: 1683–1694, 2004.
- Nahum L, Morand S, Barcellona-Lehmann S, and Schnider A. Instinctive modulation of cognitive behavior: a human evoked potential study. *Human Brain Mapping*, 30: 2120–2131, 2009b.
- Nahum L, Ptak R, Leemann B, and Schnider A. Disorientation, confabulation, and extinction capacity: clues on how the brain creates reality. *Biological Psychiatry*, 65: 966–972, 2009a.
- O'Doherty J, Critchley H, Deichmann R, and Dolan RJ. Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. *Journal of Neuroscience*, 23: 7931–7939, 2003.
- O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, and Andrews C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4: 95–102, 2001.
- Öhman A, Carlsson K, Lundqvist D, and Ingvar M. On the unconscious subcortical origin of human fear. *Physiology and Behavior*, 92: 180–185, 2007.
- Öhman A and Mineka S. Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, 108: 483–522, 2001.
- Ojemann JG, Akbudak E, Snyder AZ, McKinstry RC, Raichle ME, and Conturo TE. Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. *NeuroImage*, 6: 156–167, 1997.
- Phan KL, Wager T, Taylor SF, and Liberzon I. Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, 16: 331–348, 2002.
- Phelps A and LeDoux JE. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, 48: 175–187, 2005.
- Pourtois G, Grandjean D, Sander D, and Vuilleumier P. Electrophysiological correlates of rapid spatial orienting toward fearful faces. *Cerebral Cortex*, 14: 619–633, 2004.
- Preusschoff K, Quartz SR, and Bossaerts P. Human insula activation reflects risk prediction errors as well as risk. *Journal of Neuroscience*, 28: 2745–2752, 2008.
- Remijnse PL, Marjan MA, Uylings BM, and Veltman DJ. Neural correlates of a reversal learning task with an affectively neutral baseline: an event-related fMRI study. *NeuroImage*, 26: 609–618, 2005.
- Remijnse PL, Nielen MMA, van Balkom AJLM, Cath DC, van Oppen P, Uylings HBM, et al. Reduced orbitofrontal-striatal activity on a reversal learning task in obsessive-compulsive disorder. *Archives of General Psychiatry*, 63: 1225–1236, 2006.
- Rolls ET. The orbitofrontal cortex and reward. *Cerebral Cortex*, 10: 284–294, 2000.
- Rosenkilde CE, Bauer RH, and Fuster JM. Single cell activity in ventral prefrontal cortex of behaving monkeys. *Brain Research*, 209: 375–394, 1981.
- Sander D, Grafman J, and Zalla T. The human amygdala: an evolved system for relevance detection. *Reviews in the Neurosciences*, 14: 303–316, 2003.
- Sander D, Grandjean D, Pourtois G, Schwartz S, Seghier ML, Scherer KR, et al. Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *NeuroImage*, 28: 848–858, 2005.
- Schienle A, Schäfer A, Hermann A, Rohrmann S, and Vaitl D. Symptom provocation and reduction in patients suffering from spider phobia. *European Archives of Psychiatry and Clinical Neuroscience*, 257: 486–493, 2007.
- Schienle A, Schäfer A, Walter B, Stark R, and Vaitl D. Brain activation of spider phobics towards disorder-relevant, generally disgust- and fear-inducing pictures. *Neuroscience Letters*, 388: 1–6, 2005.
- Schiller D, Levy I, Niv Y, LeDoux JE, and Phelps EA. From fear to safety and back: reversal of fear in the human brain. *Journal of Neuroscience*, 28: 11517–11525, 2008.
- Schnider A. Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nature Review Neuroscience*, 4: 662–671, 2003.
- Schnider A. *The Confabulating Mind: How the Brain Creates Reality*. New York: Oxford University Press, 2008.
- Schnider A, Mohr C, Morand S, and Michel CM. Early cortical response to behaviorally relevant absence of anticipated outcomes: a human event-related potential study. *NeuroImage*, 35: 1348–1355, 2007.
- Schnider A, Treyer V, and Buck A. The human orbitofrontal cortex monitors outcomes even when no reward is at stake. *Neuropsychologia*, 43: 316–323, 2005.
- Schnider A, von Däniken C, and Gutbrod K. Disorientation in amnesia: a confusion of memory traces. *Brain*, 119: 1627–1632, 1996.
- Schoenbaum G, Saddoris MP, and Stalnaker TA. Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Annual New York Academy of Sciences*, 1121: 320–335, 2007.
- Sergerie K, Chochol C, and Armony JL. The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 32: 811–830, 2008.
- Stark R, Zimmermann M, Kagerer S, Schienle A, Walter B, Weygandt M, et al. Hemodynamic brain correlates of disgust and fear ratings. *NeuroImage*, 37: 663–673, 2007.
- Szymanski J and O'Donohue W. Fear of spiders questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 26: 31–34, 1995.
- Talairach J and Tournoux P. *Co-planar Stereotaxic Atlas of the Human Brain*. New York: Thieme, 1988.
- Thorpe SJ, Rolls ET, and Maddison S. The orbitofrontal cortex: neuronal activity in the behaving monkey. *Experimental Brain Research*, 49: 93–115, 1983.
- Tremblay L and Schultz W. Relative reward preference in primate orbitofrontal cortex. *Nature*, 398: 704–708, 1999.
- Tremblay L and Schultz W. Reward-related neuronal activity during go-nogo task performance in primate orbitofrontal cortex. *Journal of Neurophysiology*, 83: 1864–1876, 2000.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, 15: 273–289, 2002.

- Vuilleumier P, Richardson MP, Armony JL, Driver J, and Dolan RJ. Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, 7: 1271–1278, 2004.
- Vuilleumier P and Schwartz S. Emotional facial expressions capture attention. *Neurology*, 56: 153–158, 2001.
- Wendt J, Lotze M, Weike A, Hosten N, and Hamm A. Brain activation and defensive response mobilization during sustained exposure to phobia-related and other affective pictures in spider phobia. *Psychophysiology*, 45: 205–215, 2008.
- Wik G, Fredrikson M, Ericson K, Eriksson L, Stone-Elander S, and Greitz T. A functional cerebral response to frightening visual stimulation. *Psychiatry Research*, 50: 15–24, 1993.
- Williams LM, Das P, Liddell BJ, Kemp AH, Rennie CJ, and Gordon E. Mode of functional connectivity in amygdala pathways dissociates level of awareness for signals of fear. *Journal of Neuroscience*, 26: 9264–9271, 2006.
- Worsley KJ, Evans AC, Marrett S, and Neelin P. A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism*, 12: 900–918, 1992.