Molecular Psychiatry (2015), 1–9

www.nature.com/mp

ORIGINAL ARTICLE

Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies

MA Fullana[1,2,3,12], BJ Harrison[4,12], C Soriano-Mas[5,6], B Vervliet[7], N Cardoner[3,8], A Àvila-Parcet[9] and J Radua[10,11]

Classical Pavlovian fear conditioning remains the most widely employed experimental model of fear and anxiety, and continues to inform contemporary pathophysiological accounts of clinical anxiety disorders. Despite its widespread application in human and animal studies, the neurobiological basis of fear conditioning remains only partially understood. Here we provide a comprehensive meta-analysis of human fear-conditioning studies carried out with functional magnetic resonance imaging (fMRI), yielding a pooled sample of 677 participants from 27 independent studies. As a distinguishing feature of this meta-analysis, original statistical brain maps were obtained from the authors of 13 of these studies. Our primary analyses demonstrate that human fear conditioning is associated with a consistent and robust pattern of neural activation across a hypothesized genuine network of brain regions resembling existing anatomical descriptions of the 'central autonomic–interoceptive network'. This finding is discussed with a particular emphasis on the neural substrates of conscious fear processing. Our associated meta-analysis of functional deactivations —a scarcely addressed dynamic in fMRI fear-conditioning studies—also suggests the existence of a coordinated brain response potentially underlying the 'safety signal' (that is, non-threat) processing. We attempt to provide an integrated summary on these findings with the view that they may inform ongoing studies of fear-conditioning processes both in healthy and clinical populations, as investigated with neuroimaging and other experimental approaches.

Molecular Psychiatry advance online publication, 30 June 2015; doi:10.1038/mp.2015.88

INTRODUCTION

Learning to identify and to respond to signals of threat is highly adaptive and critical to survival; however, when this process becomes dysregulated, in the form of aberrant fear responses to innocuous events, full-blown anxiety disorders can emerge.[1] Numerous experimental studies have focused on elucidating the precise mechanisms of adaptive and maladaptive fear-learning processes in animals and humans across behavioral, experiential and neural domains. To do so, the vast majority of studies have employed classical Pavlovian fear conditioning (henceforth 'fear conditioning'), a simple and powerful method to model fear learning in the laboratory. In this procedure, an initially neutral stimulus comes to elicit a fear response after being paired with an aversive stimulus (unconditioned stimulus, US). This paired association transforms the neural stimulus into a conditioned stimulus (CS) with the capacity to elicit anticipatory fear responses. Fear conditioning has become widely regarded as a valid experimental model of clinical anxiety disorders and thus it is hoped that further advances in the scientific study of fear-conditioning processes will yield translational benefits in the form of optimized pathophysiological models of these common mental health disorders.[2]

Several experimental paradigms have been developed for the study of fear conditioning. Most typically, these paradigms focus on establishing conditioned fear responses (conditioned response) to a specific foreground cue (cue-conditioning) or a general background context (context conditioning). They may include variations in the number of CSs presented (single-cue, when one CS is presented vs differential conditioning, when two or more CSs are presented) or in the temporal relationship between the CS and US (delay conditioning, when the US is presented at the end of the CS vs trace conditioning, when a gap occurs between CS offset and US onset). These paradigms have now been extensively applied, in particular with regard to investigations of the neurobiological substrates of fear/anxiety processes. One important example comes from human functional magnetic resonance imaging (fMRI) studies, which have sought to clarify the role of specific brain regions and associated networks in fear-conditioning processes, in particular delay differential cue-conditioning. In an effort to summarize the results of such studies and to reach a genuine consensus regarding the common neural substrate of human differential fear conditioning, two existing meta-analyses have been reported. Initially, Etkin and Wager[3] analyzed 10 studies published between 1998 and 2005, which included a total of 117 healthy participants. Their results indicated that fear conditioning, overall, is characterized by consistent neural activation of an extended 'fear network' including the

[1]Anxiety Unit, Institute of Neuropsychiatry and Addictions, Hospital del Mar, CIBERSAM, Barcelona, Spain; [2]IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain; [3]Department of Psychiatry, Autonomous University of Barcelona, Barcelona, Spain; [4]Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne, Melbourne, Victoria, Australia; [5]Department of Psychiatry, Bellvitge University Hospital-IDIBELL, CIBERSAM, Barcelona, Spain; [6]Department of Psychobiology and Methodology of Health Sciences, Autonomous University of Barcelona, Barcelona, Spain; [7]Center for Excellence on Generalization in Health and Psychopathology, University of KU Leuven, Leuven, Belgium; [8]Depression and Anxiety Unit, Mental Health Department, Parc Taulí Sabadell University Hospital, Barcelona, Spain; [9]Autonomous University of Barcelona, Barcelona, Spain; [10]Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK and [11]Department of Translational Neuroimaging, FIDMAG Germanes Hospitalàries-CIBERSAM, Sant Boi de Llobregat, Barcelona, Spain. Correspondence: Dr MA Fullana, Anxiety Unit, Institute of Neuropsychiatry and Addictions, Hospital del Mar, Passeig Marítim, 25-29, Barcelona 08003, Spain or Dr BJ Harrison, Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne, Office 345, Level 3, 161 Barry Street, Melbourne, Australia.
E-mail: miguelangelfullana@gmail.com or habj@unimelb.edu.au
[12]These authors equally contributed to this work.
Received 5 February 2015; revised 1 April 2015; accepted 26 May 2015

dorsal anterior cingulate cortex (dACC), bilateral anterior insular cortex (AIC), amygdala, orbitofrontal cortex (OFC), anterior thalamus, ventral putamen and pallidum, and midbrain substantia nigra/ventral tegmentum. One limitation of this meta-analysis was that it included both 'instructed' and 'uninstructed' studies. In the former situation, participants are told that a particular CS will be followed by a US (that is, they are 'contingency aware') and therefore the conditioned response during the CS–US conditioning trials relates to the expression of an already learned association. In the uninstructed studies, participants are initially unable to predict when the US will occur, that is, the CS–US conditioning trials capture the learning process itself.

In addressing this limitation, Mechias *et al.*[4] conducted separate meta-analyses of instructed ($n = 10$; 162 participants) and uninstructed ($n = 15$; 198 participants) fear-conditioning studies published between 1998 and 2008. Across both study types, some commonality was observed, particularly with regard to the consistent involvement of the rostral dACC and AIC. However, meta-analytic results for the uninstructed conditioning studies were noted as less robust or more inconsistent. Limitations of this second meta-analysis include the fact that it combined results from a trace (as opposed to delay) conditioning study,[5] which has been shown to evoke distinct neural correlates.[6] In addition, almost half of the included studies involved participants performing a concurrent cognitive task during fear conditioning. This dual-task feature is an important caveat, as previous laboratory-based studies have shown that cognitive demands may have a significant impact on fear conditioning processes.[7] Finally, both of these former meta-analyses only considered neural activations (that is, relative activity increases) occurring in response to conditioned versus non-conditioned stimuli (CS+ $>$ CS−), as opposed to considering both activations and deactivations corresponding to this experimental contrast (that is, relative activity decreases, CS+ $<$ CS−). As is now broadly recognized in the neuroimaging field, functional deactivations may be as equally informative for understanding the neural substrates of complex mental activities,[8–10] including emotion processing.[8,11,12]

To address some of the limitations of these past meta-analyses and to provide an updated characterization of the neural signatures of human fear conditioning as studied with fMRI, we have implemented anisotropic effect-size signed differential mapping (AES-SDM).[13] AES-SDM is a novel neuroimaging meta-analytic approach that is capable of combining tabulated brain activation/deactivation results (that is, regional peak statistic and coordinate information) with actual empirical voxel-wise 'brain maps' of activations and deactivations (for example, statistical parametric maps (SPMs). We concentrated on uninstructed fear-conditioning studies, because (contrary to instructed studies, which primarily capture fear expression) they focus on fear learning, which has a stronger theoretical link with the hypothesized etiology and pathophysiology of clinical anxiety disorders.[14] In doing so, we were able to compile whole-brain imaging results from 27 independent fMRI studies involving 677 healthy adult participants. As an additional and novel distinguishing feature of this analysis, original SPMs were obtained from the corresponding authors of 13 of these 27 studies. Including original SPMs, as opposed to including only peak regional effects (typically reported in tables of statistics), substantially increases the analyses statistical power[13] and may avoid reporting biases that are likely to affect certain brain regions, especially smaller subcortical regions that are poorly represented in commonly used stereotaxic atlases of the human brain.[15,16] Furthermore, inclusion of this number of SPMs allowed, for the first time, the estimation of the task-specific optimal parameters for processing the peak information of the remaining studies.

To account for the moderating influence of several methodological and sociodemographic factors that may be highly likely to influence the results of fMRI fear-conditioning studies,[17] we also performed several supplementary analyses. Of note, we purposefully investigated the influence of potential US confounding (that is, CS–US co-presentation) on both activation and deactivation patterns and also investigated 'early' versus 'late' conditioning phase effects, which have been reported to better capture the involvement of specific regions of interest, including greater involvement of the amygdala during early conditioning.[18] In conducting this meta-analysis, our primary goal was therefore to provide an updated and extended characterization of the neural signature of human fear conditioning as studied with fMRI.

MATERIALS AND METHODS

Literature search and study selection

A comprehensive literature search using PubMed, Web of Knowledge and Scopus was conducted of English-language peer-reviewed fMRI studies on cued fear conditioning in human healthy adults (age >18 years) published between January 1998 and November 2013. The search terms were: 'fMRI' or 'magnetic resonance imaging', 'fear', 'conditioning', 'Pavlovian' and their combinations. In addition, manual searches were conducted within review articles and via the reference lists of individual studies. Researchers in the area were also contacted with regard to potential unpublished data. If any studies contained participant group overlap, only the first reported study was included. If not originally reported, the corresponding authors of the identified studies were asked to provide additional details and whole-brain results where necessary and possible.

We focused on studies using a delay differential cue-conditioning paradigm given that other paradigms (for example, trace or single-cue) appear to engage non-overlapping neural responses.[6,19] We only included studies with an independent validation of successful fear-response generation (for example, via skin conductance recordings), and those that conducted direct statistical comparisons between a CS+ and a CS− during conditioning. For pharmacological challenge/intervention studies, only placebo or control groups were included.

Studies were excluded if they either used masked CSs, USs with ambiguous meaning, employed changing CS–US contingencies, presented the US before conditioning or combining context and cue conditioning in the same experiment. Dual-task studies employing attentional distraction features were also excluded, but studies where the dual task feature was intended to enhance/maintain vigilance were included (Table 1). We also excluded studies from which peak information or SPMs could not be retrieved, that did not report whole-brain statistical results and/or whereby statistical thresholds varied across the assessment of different brain regions. Studies with <10 participants were also excluded (see ref. 20).

We obtained original empirical SPMs of the primary contrasts of interest (CS+ $>$ CS−; CS+ $<$ CS−) for 13 data sets. For the remaining 14 studies, peak regional coordinate statistics were extracted and coded from the original paper or from Supplementary Data provided by corresponding authors (Supplementary Table S1). In certain studies, this contrast was based only on CSs+ not paired with the US (there was no US confounding), whereas others used 'all' CS+ trials, and thus neural responses may be confounded by US-induced activity changes.[17] In addition, in certain studies all CSs trials during conditioning were included in the analysis, whereas in others 'early' and 'late' phases were compared. When more than one contrast was available for a data set, we selected the contrast relating to the analysis of all trials. If this was not available, we focused on the 'early' contrast, given that activation in some regions appears to be more pronounced during early conditioning phases.[17] However, based on available studies, we were able to conduct an 'early' versus 'late' comparison to reconcile such findings.

The literature search, decisions on inclusion and data extraction were all performed independently by two of the authors. For each data set, several demographic and task-related variables were extracted (Table 1).

Meta-analytic approach

Functional activation differences between the CS+ and the CS− were meta-analyzed using AES-SDM software, version 4.13 (www.sdmproject.com).[13,21] This method, which has been validated and used in several structural and functional fMRI studies creates a brain map of the effect size of the difference between the two conditions (CS+ and CS−)for each study (either from SPMs or from peak information) and afterwards conducts a voxel-wise random-effects meta-analysis (weighting the studies for sample size and variance)[13,20–22] (see Supplementary Material). In the

Table 1. Characteristics of the 27 delay fear-conditioning fMRI data sets included in the meta-analysis.

Author	N	% males	Mean age (years)	Pre-conditioning?	CSs types	Number of CS+/CS − trials[a]	Average ITI (s)	Reinforcement rate	Type of CS	CS	Type of US/location	Average CS–US onset delay (ms)	Task?[b]	Independent assessment of conditioning	CS+ > CS − analysis[c]	CS+ < CS − analysis[c]	US confounding
Andreatta *et al.*[52,d]	14	50	22.8	No	1 CS+, 1 CS −	16/16	20	100%	Visual	Neutral pictures	Shock/left hand	10 000	No	Subjective ratings	Whole	Whole	No
Dunsmoor *et al.*[53]	14	50	22.6	Yes	1 CS+, 1 CS −	16/16	11	63%	Visual	Fearful faces	Shock/right wrist	4000	Yes	SCR	Whole	Whole	–
Eippert *et al.*[54,e]	32	100	26.4	No	1 CS+, 1 CS −	10/10	7.5	100%	Visual	Abstract figures	Shock/right hand	4700	Yes	SCR,HR, subjective ratings	Whole	NA	–
Haritha *et al.*[55]	25	36	20	No	1 CS+, 1 CS −, other	24/24	20	100%	Auditory	Tones	Auditory	Variable	No	SCR	Whole	NA	Yes
Harrison *et al.*[56,e]	31	29	22.4	No	1CS+, 1 CS −	16/16	6.01	50%	Visual	Neutral pictures	Auditory	1900	No	Subjective ratings	Whole, early, late	Whole, early, late	No
Hermann *et al.*[57]	74	50	24.3	No	1 CS+, 1 CS −	20/20	12	100%	Visual	Neutral pictures	Shock/left shin	7900	No	SCR	Early	Early	Yes
Holt *et al.*[46]	17	100	34.2	Yes	2 CS+, 1 CS −	16/16	15	60%	Visual	Neutral pictures	Shock/dominant hand	6000	No	SCR	Early, late	NA	Yes
Kalisch *et al.*[58,f]	16	0	26	Yes	1 CS+, 1 CS −	15/15	9	80%	Visual	Angry faces	Shock/right hand	6250	Yes	SCR	Early	NA	–
Kattoor *et al.*[59]	19	68	23.7	No	1 CS+, 1 CS −	16/16	20	75%	Visual	Neutral pictures	Rectal distension	9600	No	Subjective ratings	Early, late	Early, late	Yes
Klucken *et al.*[60,e]	20	50	23.4	No	1 CS+, 1 CS −	21/21	14	100%	Visual	Neutral pictures	Picture	8000	Yes	SCR, subjective ratings	Whole	NA	–
Knight *et al.*[61]	15	47	28.8	No	1 CS+, 1 CS −	60/60	20	100%	Auditory	Tones	Auditory	10 000	No	SCR, subjective ratings	Whole	NA	Yes
Maier *et al.*[62,d]	17	41	31	Yes	1 CS+, 1 CS −	12/12	15	50%	Visual	Neutral pictures	Shock/right wrist	5000	No	SCR, subjective ratings	Whole	Whole	No
Menon *et al.*[63,e]	12	67	36.5	Yes	1 CS+, 1 CS −	45(30)/30	8.8	33%	Visual	Neutral pictures	Shock/left hand	5000	No	SCR, subjective ratings	Whole	NA	No
Merz *et al.*[64,e]	48	50	22.3	No	1 CS+, 1 CS −	21/21	12	100%	Visual	Neutral pictures	Shock/left shin	7900	Yes	SCR	Early, late	NA	–
Merz *et al.*[65,e]	32	100	24.9	No	1 CS+, 1 CS −	16/16	10.75	63%	Visual	Neutral pictures	Shock/left shin	7900	No	SCR	Early, late	NA	Yes
Milad *et al.*[66]	14	NA	NA	Yes	2 CS+, 1 CS −	16/16	15	60%	Visual	Neutral pictures	Shock/hand	6000	No	SCR	Early, late	Early	No
Milad *et al.*[67]	17	47	25.8	Yes	2 CS+, 1 CS −	16/16	15	62%	Visual	Neutral pictures	Shock/hand	6000	No	SCR	Whole	NA	Yes
Romaniuk *et al.*[68,e]	20	70	35.1	No	1 CS+, 1 CS −	24 (12)/24	11	50%	Visual	Neutral pictures	Picture	2000	Yes	SCR, RT	Early	NA	–
Schiller *et al.*[43]	17	53	22.6	No	1 CS+, 1 CS −	18 (12)/12	12	33%	Visual	Angry face	Shock/right wrist	4000	No	SCR	Whole, early, late	NA	No
Schultz *et al.*[69]	27	56	22.3	No	1 CS+, 1 CS −	10/10	20	100%	Visual	Neutral pictures	Shock/right foot	8000	No	SCR, subjective ratings	Whole	NA	Yes
Sehlmeyer *et al.*[70,f]	32	38	23.6	Yes	1 CS+, 1 CS −	40 (30)/30	11,5	25%	Visual	Neutral faces	Auditory	2000	No	Subjective ratings	Whole	NA	No
Spoormaker *et al.*[71,e]	35	100	24.2	No	2 CS+, 1 CS −	30 (15)/15	9	50%	Visual	Neutral pictures	Shock/right hand	3100	No	SCR	Whole	NA	No
Stark *et al.*[72,e]	17	47	23.6	No	1 CS+, 1 CS −	30/30	10	100%	Visual	Neutral pictures	Shock/left shin	7900	No	SCR	Whole	NA	Yes
Tabbert *et al.*[73,e]	18	6	25.2	No	1 CS+, 1 CS −	30/30	10	100%	Visual	Neutral pictures	Shock/left shin	7900	No	SCR	Early, late	NA	Yes
van Well *et al.*[40,e]	37	35	22.3	No	1 CS+, 1 CS −	8/8	20	75%	Visual	Fear pictures	Shock/right shin	9500	No	Startle pot., subjective ratings	Whole	Whole	Yes
Visser *et al.*[74,e]	19	26	22.2	No	2 CS+, 2 CS −	26/26	21.5	50%	Visual	Neutral pictures	Shock/right shin	6500	No	Contingency awareness	Whole	Whole	No
Visser *et al.*[75,e]	38	39	23.6	No	2 CS+, 2 CS −, 2 CS neutral	26/26	19.5	50%	Visual	Neutral pictures	Shock/right shin	4500	No	Pupil dilation response	Whole	Whole	No
Total/mean	677	53%	25.37			20/20	13,91	71%				5385					

Abbreviations: CS, conditioned stimulus; CS+, CS followed by unconditioned stimulus; CS−, CS not followed by unconditioned stimulus; HR, heart rate; ITI, inter-trial interval; NA, not available; Pot, startle potentiation; RT, reaction time; SCR, skin conductance recording; SPM, statistical parametric map; US, unconditioned stimulus; US confound, possible confounding effects of US on fMRI analyses; '−' studies using cognitive tasks, which where not included in the US confound analyses. [a]In parentheses, unreinforced CS trials. [b]Studies using a task during fear conditioning. [c]Contrast available for fear conditioning: whole (whole conditioning), early (early conditioning trials) and late (late conditioning trials). [d]Studies where the CS+ versus CS− contrast was calculated during a 'test phase' immediately after conditioning. [e]Data sets for which SPMs were available. [f]Two fear-conditioning phases were reported. In the Kalisch *et al.* study, only the first conditioning phase was included, because it was followed by a fear-extinction phase. In the study by Sehlmeyer *et al.*, both conditioning phases were combined together.

present analysis, and for the first time, we were able to empirically derive the optimal parameters (40% anisotropy and kernel full width at half maximum = 20 mm) for creating the effect size maps of the studies from which peak information but not SPMs were available.[10]

To assess the robustness of findings, we conducted a jackknife sensitivity analysis.[20] We also repeated the analysis after including only those studies considering both activations and deactivations ($n = 19$), or only those studies for which SPMs were available ($n = 13$). Finally, the I^2 index and Egger's method were used to assess for heterogeneity of effect sizes and publication bias, respectively (see Supplementary Material).

We conducted an additional meta-analysis comparing studies with potential US confounding ($n = 11$) to those with no US confounding ($n = 10$) (Table 1). To reduce variability, we did not include studies with dual task features ($n = 6$). As introduced above, a meta-analysis was also conducted of studies comparing 'early' and 'late' conditioning phase effects. Although an 'early versus late' contrast was only directly available from three studies, it could be estimated from the early and late contrast results of four additional studies taking into account the correlation between the early and late phases (estimated from the initial three studies). Despite only including seven studies, SPMs of four of them were available for this meta-analysis.

The potential influence of the following variables on estimated activations and deactivations was further explored by means of meta-regression: gender (% females), mean age of participants, the presence of a pre-conditioning (habituation) phase, number of CSs during conditioning, reinforcement rate, average CS–US delay, type of US (electric shock vs other) and the use of cognitive task.

Statistical significance was assessed with AES-SDM default thresholds (voxel-level $P < 0.005$ uncorrected, minimum extent 10 contiguous voxels), as previous simulations indicate that this threshold provides an optimal balance between sensitivity and false-positive rate.[13] A more conservative threshold ($P < 0.0005$) was applied to meta-regression analyses in order to correct for the application of multiple tests.[23–25] For the sake of completeness, cluster-based corrected *P*-values are provided in all tables of results, although they should be taken with caution as cluster-based *P*-values depend on the cluster-forming threshold (we used the uncorrected *P*-value).[26] Results are reported in Montreal Neurological Institute space.

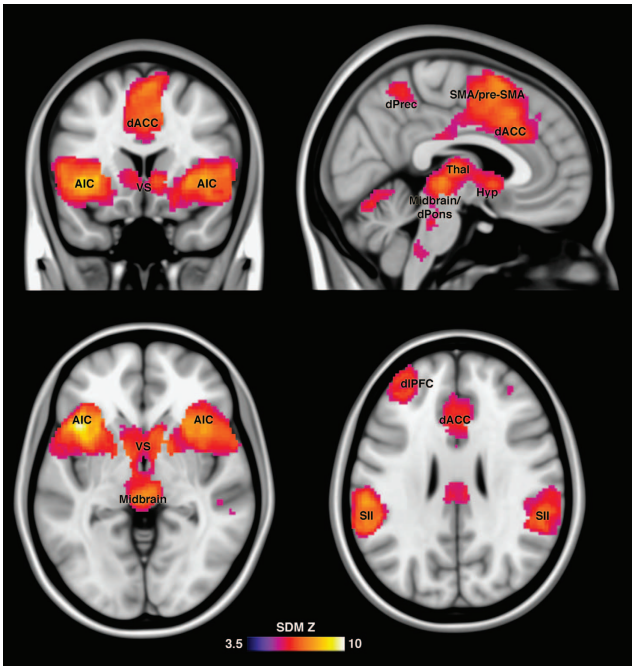

Figure 1. Significant brain functional activation to the CS+ versus CS− determined by meta-analysis. Results are displayed at $P < 0.005$ (cluster size ⩾10 voxels) on the Montreal Neurological Institute 152 T1 0.5-mm template. Abbreviations: AIC, anterior insular cortex; dACC, dorsal anterior cingulate cortex; dlPFC, dorsolateral prefrontal cortex; dPons, dorsal pons; dPrec, dorsal precuneus; HYP, hypothalamus; SII, secondary somatosensory cortex; SMA, supplementary motor area; Thal, thalamus; VS, ventral striatum.

RESULTS

Study characteristics

The final sample (Supplementary Figure S1) consisted of 27 independent data sets reporting a CS+ > CS− contrast (of which 19 also presented a CS+ < CS− contrast) including a total of 677 subjects (54% males), with a mean age of 25.37 years (range: 20–36; s.d. = 4.19; see Table 1).

Primary meta-analytic results

Seven large bilateral regional clusters were mapped as demonstrating consistently significant functional activations during differential fear conditioning (CS+ > CS−). The major regions comprising these clusters were as follows: (1) the AIC extending to the frontal operculum; (2) the ventral striatum (including the ventral rostral putamen, ventral pallidum, ventral caudate/nucleus accumbens and approximate ventral tegmental area) and major thalamic nuclei (mediodorsal, centromedial, ventrolateral nuclei), the latter referenced against thalamic nuclei probability maps from the SPM Anatomy Toolbox; (3) a large expanse of medial wall cortex including the pre-supplementary and supplementary motor areas, and the dACC (both rostral and caudal divisions) and a distinct cluster of the dorsal–anterior precuneus; (4) the second somatosensory cortex (SII)/parietal operculum; (5) the dorsolateral prefrontal cortex (more prominently left sided); (6) the lateral premotor cortex; (7) the ventral–posterior precuneus; and (8) the lateral cerebellum. We also note the relevant involvement of smaller subcortical regions including the septal–hypothalamic zone and midbrain/dorsal pons, which contains the periaqueductal grey, parabrachial nucleus, reticular formation, raphe nuclei (pontine and midbrain) and locus coeruleus. An additional cluster was also mapped to the pontomedullary junction (Figure 1 and Supplementary Table S2).

Nine large bilateral regional clusters were also mapped as demonstrating consistently significant deactivations during differential fear conditioning (CS+ < CS−). These deactivation clusters comprised the following: (1) lateral and midline primary somatosensory cortex, as well as the dorsal posterior insular cortex; (2) dorsal anterior prefrontal cortex; (3) ventromedial prefrontal cortex (vmPFC); (4) posterior cingulate cortex (PCC), including the retrosplenial cortex, hippocampus and lateral inferior and middle temporal cortex; (5) lateral OFC; 6) inferior parietal cortex (complete angular gyrus extending to the intraparietal sulcus); (7) lateral retrosplenial cortex; (8) posterior cerebellum; (9) dorsal caudate nucleus (body); and (10) dorsal–posterior precuneus (Figure 2 and Supplementary Table S2).

Supplementary Figure S2 displays the activation/deactivation results together across 24 axial slices of a whole-brain anatomical reference volume. Our corresponding robustness analyses indicated that all results were highly replicable, and that there was neither substantial heterogeneity nor evidence of potential publication bias in the main results (Supplementary Material).

Additional meta-analyses

A direct comparison of studies with potential US confounding versus those without such confounding indicated a greater involvement of the rostral dACC/dmPFC, bilateral ventral AIC, ventral striatum (ventral–mid caudate and approximate ventral tegmental area) and the right SII in the latter scenario (Figure 3) versus a greater relative involvement of the anterior calcarine sulcus in studies with potential confounding. Comparison of the two study types also indicated greater relative deactivation of the bilateral intraparietal sulcus, vmPFC and the left hippocampus in studies without potential confounding (see Supplementary Table S3). However, it must be noted that studies with potential

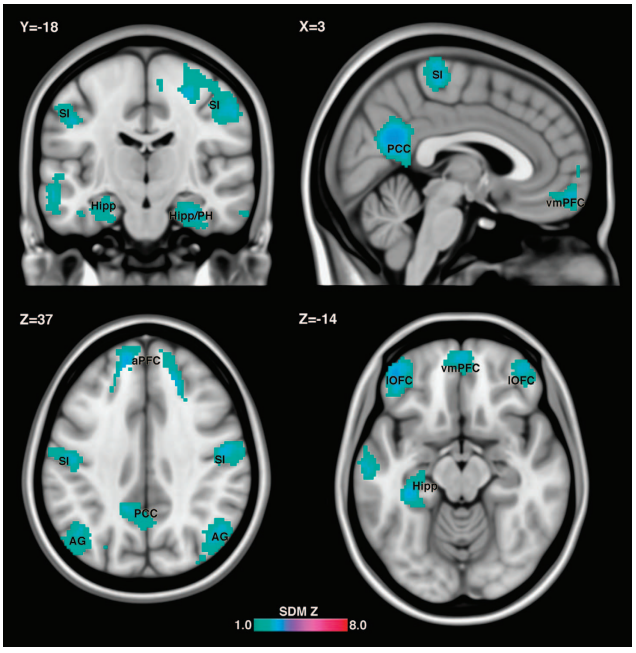

Figure 2. Significant brain functional deactivation to the CS+ versus CS − determined by meta-analysis. Results are displayed at $P < 0.005$ (cluster size $\geqslant 10$ voxels) on the Montreal Neurological Institute 152 T1 0.5-mm template. Abbreviations: AG, angular gyrus; aPFC, anterior prefrontal cortex; Hipp, hippocampus; lOFC, lateral orbitofrontal cortex; PCC, posterior cingulate cortex; PH, parahippocampal formation; SI, primary somatosensory cortex.

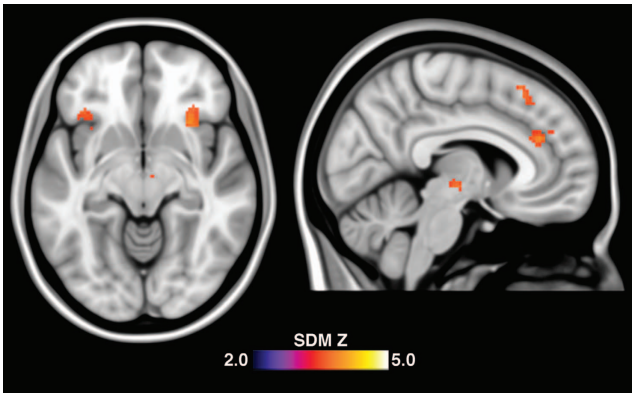

Figure 3. Influence of potential US confounding on fear-conditioned brain activation, highlighting regions that demonstrated greater relative activation when no potential confounding existed. Results are displayed at $P < 0.005$ (cluster size $\geqslant 10$ voxels) on the Montreal Neurological Institute 152 T1 0.5-mm template.

confounding also had a higher average reinforcement rate (85%) compared with those without potential confounding (50%). (See expanded discussion in Supplementary Material).

The 'early versus late' meta-analysis indicated that the early phase was associated with greater relative activation of the medial thalamus, left SII and left AIC (Supplementary Figure S3 and Supplementary Table S4). By comparison, the late phase was associated with greater activation of the subgenual ACC/vmPFC extending to the medial OFC and the right anterior hippocampus, as well as greater deactivation of the right precuneus.

Meta-regression analyses

Gender had no influence on the primary activation results. Younger age was associated with significantly greater activation of the right AIC extending to the frontal operculum, right dorsolateral prefrontal cortex, pre-supplementary motor area and the left frontal operculum, as well as greater deactivation of the left anterior hippocampus and posterior fusiform gyrus (Supplementary Table S5).

The results of the meta-regression with respect to task-specific features are summarized in Figure 4 (see also Supplementary Table S5). The use of a pre-conditioning phase significantly reduced conditioning-related activation of several regions. Of these regions, all but the ventral ACC (Figure 4a) overlapped with regions implicated in the primary meta-analysis (Figure 1). By comparison, activation of the left anterior hippocampus was greater with the use of a pre-conditioning phase (Figure 4b).

Presenting a higher number of CS trials during conditioning was associated with greater activation of the left ventral caudate nucleus extending to the nucleus accumbens (Figure 4c). A higher CS–US reinforcement rate reduced strength of activation of the rostral dACC/dmPFC (Figure 4d), right parietal operculum (SII) and a small cluster of right AIC. A longer delay between the CS and US was associated with greater activation of the subgenual ACC and vmPFC, and less activation of the right parietal operculum/SII, pre-supplementary motor area and the right premotor cortex (Figure 4e). The use of a tactile electric shock US (compared with other US types) was associated with greater activation of the left caudal dorsal ACC/ventral supplementary motor area (Figure 4f). Finally, the concurrent use of cognitive tasks reduced the activation of the bilateral mid-AIC.

DISCUSSION

This updated and extended meta-analysis has identified a highly consistent pattern of functional brain activation and deactivation associated with human differential fear conditioning. Robustness analyses confirmed the strength of the primary findings, whereas supplementary analyses were able to address the influence of specific task features on the associated patterns of brain activity.

Functional activations: an integrated perspective

The notion that human differential fear conditioning, as studied with fMRI, activates a consistent distributed set of brain regions or extended 'fear network'[17,27,28] was wholly confirmed by the current meta-analysis. Although only a small number of the same studies were included in our meta-analysis compared with the aforementioned analyses (Supplementary Table S6),[3,4] striking overlap emerged in the specific pattern of the brain activation observed. We therefore conclude that in spite of the considerable methodological diversity that exists across individual fear-conditioning studies, these studies consistently evoke a common large-scale brain activation response.[17]

With the goal of providing an integrated perspective, we hypothesize that this common pattern of activation is genuinely consistent with the engagement of a large-scale brain functional network, that is, a coordinated pattern of brain activation across anatomically distributed brain regions with well-known anatomical connectivity. In this context, we emphasize in particular the involvement of medial wall 'cingulofrontal cortex' regions, including the dACC, together with the bilateral AIC.

Considerable evidence now supports the view that these brain regions form major cortical components of a large-scale brain network with specialized functional relevance to homeostatic autonomic and behavioral (including affective) regulation.[29–32] More specifically, this brain network has been linked to interoception—the sense of the physiological condition of the

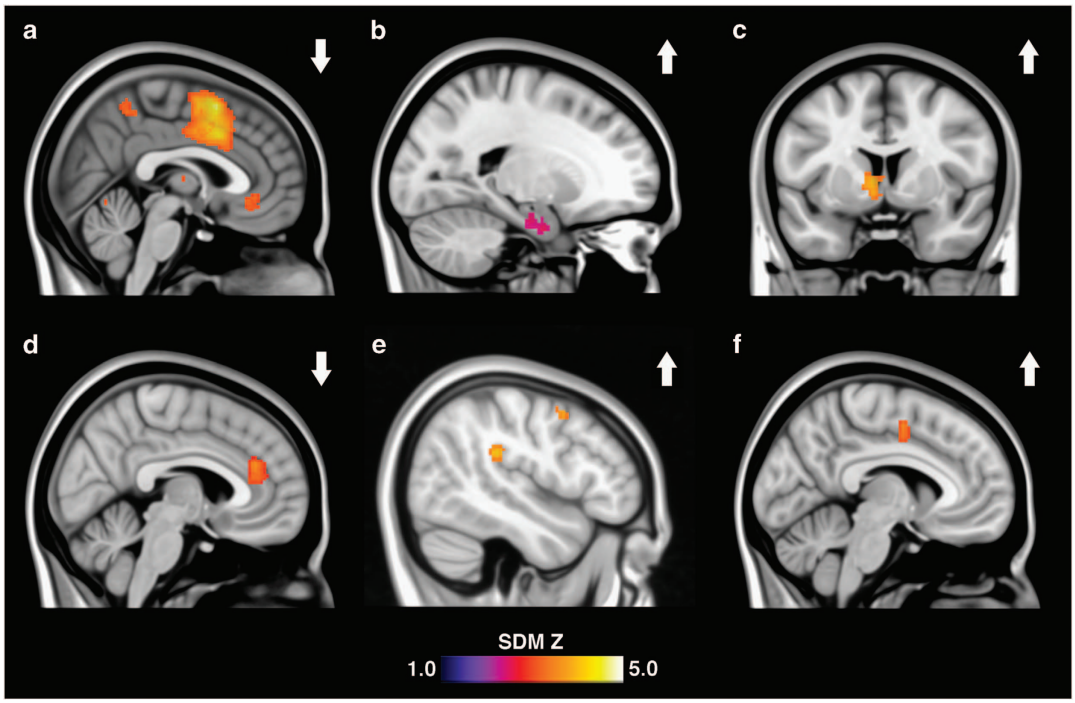

Figure 4. Summary of meta-regression analyses: influence of certain task features on fear conditioning-related brain functional activation. (**a**) Regions exhibiting reduced activation during fear conditioning following the use of a pre-conditioning task phase. (**b**) Regions exhibiting greater activation during fear conditioning following the use of a pre-conditioning task phase. (**c**) Regions exhibiting greater activation during fear conditioning in relation to a higher number of presented CS trials. (**d**) Regions exhibiting reduced activation during fear conditioning in relation to a higher CS–US reinforcement rate. (**e**) Regions exhibiting greater activation during fear conditioning in relation to a longer CS–US delay. (**f**) Regions exhibiting greater activation during fear conditioning in response to a tactile (electric shock) US compared with other US stimulus types. Results are displayed at $P < 0.005$ (cluster size $\geqslant 10$ voxels) on the Montreal Neurological Institute 152 T1 0.5-mm template. As all results correspond to changes in activation levels, a warm color display has been used and a uniform activation magnitude (SDM Z) adopted for each result. Up/down arrows indicate relative increases or decreases in activation effects.

entire body—including the representation of higher-order interoceptive feelings in terms of subjective emotional awareness.[30] Within the framework of this 'central autonomic–interoceptive network', the AIC and dACC are conceptualized as the major cortical input–output components, whereby the AIC is responsible for generating an integrated awareness of one's cognitive, affective and physical state that becomes re-represented in the dACC in order to facilitate homeostatic autonomic and behavioral responses.[30,33] Supporting this network perspective, co-activation of these brain regions has been routinely observed in human fMRI studies, either accompanying or directly (that is, temporally) correlated with changes in physiological autonomic arousal measures.[30,31,33,34] Relevantly, we also identified robust involvement of key subcortical viscerosensory (and visceromotor) processing sites, including the dorsal midbrain (periaqueductal gray and parabrachial nucleus), ventromedial thalamus and hypothalamus,[30,31,33,34] as well as the pontomedullary junction, which contains the nucleus of the solitary tract—a principal convergence site for viscerosensory afferent relay to higher areas.[31] Therefore, it seems reasonable to suggest human fMRI fear-conditioning studies primarily evoke a central autonomic–interoceptive network response that, in addition to representing autonomic responses to threat, also likely represent broader threat appraisal and response processes that cut across cognitive, motivational and psychomotor domains.

From the specific analysis of potential US confounding, we observed significantly greater activation of the rostral dACC/dmPFC, the ventral AIC, ventral striatum and SII when no such confounding existed. Hence, for these regions, it can be more confidently interpreted that their accompanying response to the CS+ was purely anticipatory in nature and unrelated to the generation of defensive autonomic responses to the US. Relevant overlap can also be noted between this analysis and the analysis of reinforcement rate, whereby higher reinforcement was associated with reduced activation of the same rostral dACC/dmPFC, right SII and ventral AIC regions (overlap shown in Supplementary Figure S4). As those studies without potential confounding had generally lower reinforcement rates, these brain responses likely represent the common influence of uncertainty or unpredictability, that is, where lower reinforcement rates increase uncertainty or unpredictability. Greater involvement of the rostral dACC/dmPFC is particularly interesting in view of recent work implicating this region specifically in the conscious appraisal of threat, with particular relevance to the subjective experience of fear and anxiety.[35] From these studies, there is evidence that a certain component of the rostral dACC/dmPFC response during fear conditioning is specifically cognitively modulated and dissociable from autonomic arousal changes.[35] By comparison, other components of the extended dACC/medial wall response are likely to represent the direct interaction between autonomic and cognitive states.[36,37] The AIC has also been proposed to mediate higher-level appraisal and anticipatory processes with relevance to the conscious experience of fear and anxiety,[38] although this proposal has yet to be conclusively demonstrated beyond the more general hypothesized link between AIC function and interoceptive awareness. In summary, fMRI fear-conditioning tasks evoke a primary neural signature that is anatomically consistent with the engagement of the central autonomic–interoceptive network. This finding, taken with other recent evidence, endorses an emerging view that the functional activity of this brain system may have direct relevance to the conscious experience of fear and anxiety, in addition to non-conscious

aspects of fear processing,[39] a view which is compatible with current appraisal theories.[35]

Functional deactivations: neural correlate of 'safety' signal processing?

Although derived from fewer studies, our analysis of functional deactivations also identified a robust and anatomically distributed pattern of activity change involving the vmPFC, lateral OFC, hippocampus and PCC. Deactivation of the vmPFC and PCC, in particular, has become recognized as a characteristic functional signature of the human 'default mode network'—a large-scale brain system that exhibits consistent functional decreases in activity in fMRI studies when conditions of goal-directed task performance are compared with conditions of low task demand, such as passive resting states.[8,9] Thus, one possible interpretation of this deactivation effect is that responding to the CS− versus CS+ corresponds with more 'resting-like' default mode network activity. Although it is difficult to completely exclude this possibility, the following factors deserve some consideration. First, processing of CS− is not a passive condition; it places specific demands on learning, attentional and cognitive evaluative processes.[40] Second, only partial involvement of the default mode network regions was observed in our meta-analysis: there was no characteristic deactivation of the extended dorsomedial PFC, which is commonly observed in fMRI studies.[41] Third, non-default mode network regions, including the lateral OFC and primary somatosensory cortex were also implicated as part of this robust deactivation pattern. Considered together, it seems more reasonable to interpret this finding as representing a specific neural correlate of processing the CS− 'safety' signal.

Previous fMRI studies of fear conditioning have argued that vmPFC deactivation is indeed likely to represent processing of the CS− as a non-threat or 'safety' signal, with particular emphasis on fear-response inhibition.[42,43] Similar ideas have also been invoked in the context of fear extinction and fear-reversal studies, where there is accumulating evidence to suggest that vmPFC activity may encode the distinction between non-threatening and threatening stimuli. Although the precise nature of vmPFC 'safety' processing may vary across fear-learning contexts, one possibility is that the distinction between CS− and CS+ signals evoke a common neural substrate for the dynamic representation of reward value, with CS− 'safety' signals having an intrinsic positive reward value.[44] Co-deactivation of the lateral OFC is also interesting in this context, having been consistently implicated as part of the extended neural circuitry of reward-associative learning, albeit with evidence existing for specialized contributions between medial and lateral vmPFC/OFC regions.[45]

Specific deactivation of the PCC/retrosplenial cortex together with the hippocampus is also interesting to consider with regards to the processing of the CS− as a 'safety' signal. Both regions are well-known components of an extended episodic memory network and it has been proposed recently that the co-engagement of these areas may specifically encode episodic memory traces of the CS/US association.[46] Although hippocampal involvement has been more traditionally linked to trace conditioning and the declarative learning of discontinuous temporal associations, our findings raise the possibility that an episodic memory representation may be established during the specific processing of the CS−, potentially related to the establishment of contingency awareness (see ref. 6). Another possibility is that greater PCC/hippocampal activation to the CS− represents a neural correlate of 'relief'—related to the US omission (see ref. 47). Thus, although the specific meaning of functional deactivations in the context of fMRI fear-conditioning studies remains speculative, the associated brain regions and the robustness of their response observed to the CS− versus CS+, as demonstrated via meta-analysis, should compel further investigations.

'Fear' versus 'threat conditioning'

It has been argued recently that the actual concept of 'fear conditioning' be abandoned, because it blurs the distinction between conscious and non-conscious fear processes.[48] In this sense, the term 'fear' should be invoked only to define its conscious experience, whereas 'threat conditioning' may be a preferable term when seeking to define the implicit (non-conscious) processes that control defense/survival responses elicited by threats. Although the boundary between these processes and their neuroanatomical representation remains a topic of debate, the idea of 'fear' as primarily relating to consciously felt experience appears to resonate with the overall findings of fMRI-conditioning studies, which consistently implicate a neural signature with direct hypothesized relevance to the conscious experience of emotions.[30,31,39] This suggestion is not to imply that such studies only engage conscious fear processes, which would be an oversimplification, but that the prominent engagement of cortical regions, in particular the AIC and rostral dACC appears consistent with such notions.

Given the strong empirical link between amygdala circuitry and threat-conditioning processes,[49] it is relevant that our meta-analysis did not characterize robust involvement of the amygdala region as was previously highlighted in various independent studies[50,51]. Other fMRI studies have reported transient amygdala responses during early conditioning;[18] however, this was neither apparent from our meta-analysis of early versus late acquisition phases. Although the absence of amygdala involvement may be partly explained by technical constraints of fMRI and other reasons,[4,17] it seems reasonable to also conclude that human fMRI fear-conditioning experiments generally do not evoke consistent responses within the classical amygdala defense/threat detection circuitry.

Strengths and limitations

Strengths of the current analysis include the following: the use of novel meta-analytic methods that combine the positive features of typical coordinate approaches with those from standard meta-analytic methods; the novel recreation of effect size maps using the optimal anisotropy and full width at half maximum for the specific data set; and the inclusion of 13 original SPMs to more effectively estimate the contrasts of interest.

Our study nevertheless has some limitations inherent to most meta-analyses; the different studies included employed various statistical thresholds, and although our methods provide excellent control for false positives it is more difficult to avoid false negatives.[13] In addition, the results of our meta-regression analyses were hampered by the low variability in some of the variables studied. Nevertheless, these analyses indicate that despite the robustness of the main findings, some demographic and experimental design features do have a tangible influence on the resulting neural signatures of fear conditioning (see expanded discussion in Supplementary Material). Despite providing an optimal balance between sensitivity and false positive rate,[13] the default AES-SDM statistical thresholds were based on uncorrected *P*-values, which may be seen as a limitation. Lastly, the influence of some important factors related to fear conditioning, such as contingency awareness, could not be assessed because such data was rarely presented among the individual studies.

CONCLUSION

Classical Pavlovian fear conditioning will continue to inform our current understanding of fear and anxiety. The results of this meta-analysis suggest that when applied in the human neuroimaging context, these experiments may be especially useful for expanding knowledge of how fear and anxiety are experienced in subjective terms with a particular relevance to central intero-

ceptive representations of brain–body interactions. It is notable that these relationships remain largely unexplored in the neuroscientific study of patients with clinical anxiety disorders, thus potentially representing a novel research avenue.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank the following authors for providing additional data or information not included in some original articles: M Andreatta, C Büchel, M Carter, H Critchley, J Dunsmoor, F Eippert, H Flor, J Hall, A Hermann, R Kalisch, J Kattoor, M Kindt, T Klucken, D Knight, TB Lonsdorf, M Menon, M Milad, C Merz, P Pauli, EA Phelps, L Romaniuk, D Schiller, D Schultz, V Spoormaker, R Stark, C Uhlmann, S van Well and R. Visser. This work was supported by 'Miguel Servet' contracts from the Carlos III Health Institute (Spain) to CS-M (CP10/00604) and JR (CP14/00041), grants from Carlos III Health Institute/FEDER to MAF/NC (PI12/00273) and CS-M (PI13/01958), and an Australian NHMRC Project Grant to BJH (APP1025619).

REFERENCES

1 Lissek S, Powers AS, McClure EB, Phelps EA, Woldehawariat G, Grillon C *et al.* Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behav Res Ther* 2005; **43**: 1391–1424.
2 Kindt M. A behavioural neuroscience perspective on the aetiology and treatment of anxiety disorders. *Behav Res Ther* 2014; **62**: 24–36.
3 Etkin A, Wager TD. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am J Psychiatry* 2007; **164**: 1476–1488.
4 Mechias ML, Etkin A, Kalisch R. A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. *Neuroimage* 2010; **49**: 1760–1768.
5 Buchel C, Dolan RJ, Armony JL, Friston KJ. Amygdala-hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *J Neurosci* 1999; **19**: 10869–10876.
6 Knight DC, Cheng DT, Smith CN, Stein EA, Helmstetter FJ. Neural substrates mediating human delay and trace fear conditioning. *J Neurosci* 2004; **24**: 218–228.
7 Carter RM, Hofstotter C, Tsuchiya N, Koch C. Working memory and fear conditioning. *Proc Natl Acad Sci USA* 2003; **100**: 1399–1404.
8 Harrison BJ, Pujol J, Contreras-Rodriguez O, Soriano-Mas C, Lopez-Sola M, Deus J *et al.* Task-induced deactivation from rest extends beyond the default mode brain network. *PLoS One* 2011; **6**: e22964.
9 Harrison BJ, Pujol J, Lopez-Sola M, Hernandez-Ribas R, Deus J, Ortiz H *et al.* Consistency and functional specialization in the default mode brain network. *Proc Natl Acad Sci USA* 2008; **105**: 9781–9786.
10 Radua J, Mataix-Cols D. Heterogeneity of coordinate-based meta-analyses of neuroimaging data: an example from studies in OCD – Authors' reply. *Br J Psychiatry* 2010; **197**: 77.
11 Radua J, Sarro S, Vigo T, Alonso-Lana S, Bonnin CM, Ortiz-Gil J *et al.* Common and specific brain responses to scenic emotional stimuli. *Brain Struct Funct* 2014; **219**: 1463–1472.
12 Ter Minassian A, Ricalens E, Humbert S, Duc F, Aube C, Beydon L. Dissociating anticipation from perception: acute pain activates default mode network. *Hum Brain Mapp* 2013; **34**: 2228–2243.
13 Radua J, Mataix-Cols D, Phillips ML, El-Hage W, Kronhaus DM, Cardoner N *et al.* A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. *Eur Psychiatry* 2012; **27**: 605–611.
14 Mineka S, Zinbarg R. A contemporary learning theory perspective on the etiology of anxiety disorders: it's not what you thought it was. *Am Psychol* 2006; **61**: 10–26.
15 David SP, Ware JJ, Chu IM, Loftus PD, Fusar-Poli P, Radua J *et al.* Potential reporting bias in fMRI studies of the brain. *PLoS One* 2013; **8**: e70104.
16 Fusar-Poli P, Radua J, Frascarelli M, Mechelli A, Borgwardt S, Di Fabio F *et al.* Evidence of reporting biases in voxel-based morphometry (VBM) studies of psychiatric and neurological disorders. *Hum Brain Mapp* 2014; **35**: 3052–3065.
17 Sehlmeyer C, Schoning S, Zwitserlood P, Pfleiderer B, Kircher T, Arolt V *et al.* Human fear conditioning and extinction in neuroimaging: a systematic review. *PLoS One* 2009; **4**: e5865.
18 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**: 937–945.
19 Cheng DT, Knight DC, Smith CN, Helmstetter FJ. Human amygdala activity during the expression of fear responses. *Behav Neurosci* 2006; **120**: 1187–1195.
20 Radua J, Mataix-Cols D. Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. *Br J Psychiatry* 2009; **195**: 393–402.
21 Radua J, Rubia K, Canales-Rodriguez EJ, Pomarol-Clotet E, Fusar-Poli P, Mataix-Cols D. Anisotropic kernels for coordinate-based meta-analyses of neuroimaging studies. *Front Psychiatry* 2014; **5**: 13.
22 Radua J, van den Heuvel OA, Surguladze S, Mataix-Cols D. Meta-analytical comparison of voxel-based morphometry studies in obsessive-compulsive disorder vs other anxiety disorders. *Arch Gen Psychiatry* 2010; **67**: 701–711.
23 Hart H, Radua J, Mataix-Cols D, Rubia K. Meta-analysis of fMRI studies of timing in attention-deficit hyperactivity disorder (ADHD). *Neurosci Biobehav Rev* 2012; **36**: 2248–2256.
24 Lim L, Radua J, Rubia K. Gray matter abnormalities in childhood maltreatment: a voxel-wise meta-analysis. *Am J Psychiatry* 2014; **171**: 854–863.
25 Radua J, Borgwardt S, Crescini A, Mataix-Cols D, Meyer-Lindenberg A, McGuire PK *et al.* Multimodal meta-analysis of structural and functional brain changes in first episode psychosis and the effects of antipsychotic medication. *Neurosci Biobehav Rev* 2012; **36**: 2325–2333.
26 Friston KJ, Worsley KJ, Frackowiak RS, Mazziotta JC, Evans AC. Assessing the significance of focal activations using their spatial extent. *Hum Brain Mapp* 1994; **1**: 210–220.
27 Buchel C, Dolan RJ. Classical fear conditioning in functional neuroimaging. *Curr Opin Neurobiol* 2000; **10**: 219–223.
28 Kim JJ, Jung MW. Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neurosci Biobehav Rev* 2006; **30**: 188–202.
29 Cameron OG. Visceral brain-body information transfer. *Neuroimage* 2009; **47**: 787–794.
30 Craig AD. How do you feel--now? The anterior insula and human awareness. *Nat Rev Neurosci* 2009; **10**: 59–70.
31 Critchley HD, Harrison NA. Visceral influences on brain and behavior. *Neuron* 2013; **77**: 624–638.
32 Saper CB. The central autonomic nervous system: conscious visceral perception and autonomic pattern generation. *Annu Rev Neurosci* 2002; **25**: 433–469.
33 Medford N, Critchley HD. Conjoint activity of anterior insular and anterior cingulate cortex: awareness and response. *Brain Struct Funct* 2010; **214**: 535–549.
34 Beissner F, Meissner K, Bar KJ, Napadow V. The autonomic brain: an activation likelihood estimation meta-analysis for central processing of autonomic function. *J Neurosci* 2013; **33**: 10503–10511.
35 Kalisch R, Gerlicher AM. Making a mountain out of a molehill: on the role of the rostral dorsal anterior cingulate and dorsomedial prefrontal cortex in conscious threat appraisal, catastrophizing, and worrying. *Neurosci Biobehav Rev* 2014; **42**: 1–8.
36 Critchley HD, Mathias CJ, Dolan RJ. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* 2001; **29**: 537–545.
37 Wager TD, van Ast VA, Hughes BL, Davidson ML, Lindquist MA, Ochsner KN. Brain mediators of cardiovascular responses to social threat, part II: Prefrontal-subcortical pathways and relationship with anxiety. *Neuroimage* 2009; **47**: 836–851.
38 Paulus MP, Stein MB. An insular view of anxiety. *Biol Psychiatry* 2006; **60**: 383–387.
39 Adolphs R. The biology of fear. *Curr Biol* 2013; **23**: R79–R93.
40 van Well S, Visser RM, Scholte HS, Kindt M. Neural substrates of individual differences in human fear learning: evidence from concurrent fMRI, fear-potentiated startle, and US-expectancy data. *Cogn Affect Behav Neurosci* 2012; **12**: 499–512.
41 Christoff K, Gordon AM, Smallwood J, Smith R, Schooler JW. Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proc Natl Acad Sci USA* 2009; **106**: 8719–8724.
42 Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL. Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biol Psychiatry* 2007; **62**: 446–454.
43 Schiller D, Levy I, Niv Y, LeDoux JE, Phelps EA. From fear to safety and back: reversal of fear in the human brain. *J Neurosci* 2008; **28**: 11517–11525.
44 Schiller D, Delgado MR. Overlapping neural systems mediating extinction, reversal and regulation of fear. *Trends Cogn Sci* 2010; **14**: 268–276.
45 Grabenhorst F, Rolls ET. Value, pleasure and choice in the ventral prefrontal cortex. *Trends Cogn Sci* 2011; **15**: 56–67.
46 Holt DJ, Coombs G, Zeidan MA, Goff DC, Milad MR. Failure of neural responses to safety cues in schizophrenia. *Arch Gen Psychiatry* 2012; **69**: 893–903.
47 Leknes S, Lee M, Berna C, Andersson J, Tracey I. Relief as a reward: hedonic and neural responses to safety from pain. *PLoS One* 2011; **6**: e17870.
48 LeDoux JE. Coming to terms with fear. *Proc Natl Acad Sci USA* 2014; **111**: 2871–2878.
49 LeDoux JE. Emotional memory: in search of systems and synapses. *Ann N Y Acad Sci* 1993; **702**: 149–157.

50 Öhman A. Human fear conditioning and the amygdala. In: Whalen PJ, Phelps EA (eds) *The Human Amygdala*. Guilford Press: New York, 2009.
51 Phelps EA, Delgado MR, Nearing KI, LeDoux JE. Extinction learning in humans: role of the amygdala and vmPFC. *Neuron* 2004; **43**: 897–905.
52 Andreatta M, Fendt M, Muhlberger A, Wieser MJ, Imobersteg S, Yarali A *et al.* Onset and offset of aversive events establish distinct memories requiring fear and reward networks. *Learn Mem* 2012; **19**: 518–526.
53 Dunsmoor JE, Prince SE, Murty VP, Kragel PA, LaBar KS. Neurobehavioral mechanisms of human fear generalization. *Neuroimage* 2011; **55**: 1878–1888.
54 Eippert F, Gamer M, Buchel C. Neurobiological mechanisms underlying the blocking effect in aversive learning. *J Neurosci* 2012; **32**: 13164–13176.
55 Haritha AT, Wood KH, Ver Hoef LW, Knight DC. Human trace fear conditioning: right-lateralized cortical activity supports trace-interval processes. *Cogn Affect Behav Neurosci* 2013; **13**: 225–237.
56 Harrison BJ, Fullana MA, Soriano-Mas C, Via E, Pujol J, Martinez-Zalacain I *et al.* A neural mediator of human anxiety sensitivity (in press).
57 Hermann A, Kupper Y, Schmitz A, Walter B, Vaitl D, Hennig J *et al.* Functional gene polymorphisms in the serotonin system and traumatic life events modulate the neural basis of fear acquisition and extinction. *PLoS One* 2012; **7**: e44352.
58 Kalisch R, Holt B, Petrovic P, De Martino B, Kloppel S, Buchel C *et al.* The NMDA agonist D-cycloserine facilitates fear memory consolidation in humans. *Cereb Cortex* 2009; **19**: 187–196.
59 Kattoor J, Gizewski ER, Kotsis V, Benson S, Gramsch C, Theysohn N *et al.* Fear conditioning in an abdominal pain model: neural responses during associative learning and extinction in healthy subjects. *PLoS One* 2013; **8**: e51149.
60 Klucken T, Schweckendiek J, Koppe G, Merz CJ, Kagerer S, Walter B *et al.* Neural correlates of disgust- and fear-conditioned responses. *Neuroscience* 2012; **201**: 209–218.
61 Knight DC, Waters NS, Bandettini PA. Neural substrates of explicit and implicit fear memory. *Neuroimage* 2009; **45**: 208–214.
62 Maier S, Szalkowski A, Kamphausen S, Perlov E, Feige B, Blechert J *et al.* Clarifying the role of the rostral dmPFC/dACC in fear/anxiety: learning, appraisal or expression? *PLoS One* 2012; **7**: e50120.
63 Menon M, Jensen J, Vitcu I, Graff-Guerrero A, Crawley A, Smith MA *et al.* Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biol Psychiatry* 2007; **62**: 765–772.
64 Merz CJ, Wolf OT, Schweckendiek J, Klucken T, Vaitl D, Stark R. Stress differentially affects fear conditioning in men and women. *Psychoneuroendocrinology* 2013; **38**: 2529–2541.
65 Merz CJ, Hermann A, Stark R, Wolf OT. Cortisol modifies extinction learning of recently acquired fear in men. *Soc Cogn Affect Neurosci* 2014; **9**: 1426–1434.
66 Milad MR, Quirk GJ, Pitman RK, Orr SP, Fischl B, Rauch SL. A role for the human dorsal anterior cingulate cortex in fear expression. *Biol Psychiatry* 2007; **62**: 1191–1194.
67 Milad MR, Furtak SC, Greenberg JL, Keshaviah A, Im JJ, Falkenstein MJ *et al.* Deficits in conditioned fear extinction in obsessive-compulsive disorder and neurobiological changes in the fear circuit. *JAMA Psychiatry* 2013; **70**: 608–618.
68 Romaniuk L, Honey GD, King JR, Whalley HC, McIntosh AM, Levita L *et al.* Midbrain activation during Pavlovian conditioning and delusional symptoms in schizophrenia. *Arch Gen Psychiatry* 2010; **67**: 1246–1254.
69 Schultz DH, Balderston NL, Helmstetter FJ. Resting-state connectivity of the amygdala is altered following Pavlovian fear conditioning. *Front Hum Neurosci* 2012; **6**: 242.
70 Sehlmeyer C, Dannlowski U, Schoning S, Kugel H, Pyka M, Pfleiderer B *et al.* Neural correlates of trait anxiety in fear extinction. *Psychol Med* 2011; **41**: 789–798.
71 Spoormaker VI, Schroter MS, Andrade KC, Dresler M, Kiem SA, Goya-Maldonado R *et al.* Effects of rapid eye movement sleep deprivation on fear extinction recall and prediction error signaling. *Hum Brain Mapp* 2012; **33**: 2362–2376.
72 Stark R, Wolf OT, Tabbert K, Kagerer S, Zimmermann M, Kirsch P *et al.* Influence of the stress hormone cortisol on fear conditioning in humans: evidence for sex differences in the response of the prefrontal cortex. *Neuroimage* 2006; **32**: 1290–1298.
73 Tabbert K, Stark R, Kirsch P, Vaitl D. Hemodynamic responses of the amygdala, the orbitofrontal cortex and the visual cortex during a fear conditioning paradigm. *Int J Psychophysiol* 2005; **57**: 15–23.
74 Visser RM, Scholte HS, Kindt M. Associative learning increases trial-by-trial similarity of BOLD-MRI patterns. *J Neurosci* 2011; **31**: 12021–12028.
75 Visser RM, Scholte HS, Beemsterboer T, Kindt M. Neural pattern similarity predicts long-term fear memory. *Nat Neurosci* 2013; **16**: 388–390.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)