nature communications

Article

https://doi.org/10.1038/s41467-025-63078-x

Neural correlates of human fear conditioning and sources of variability in 2199 individuals

Received: 9 December 2024

Accepted: 7 August 2025

Published online: 23 August 2025

Check for updates

A list of authors and their affiliations appears at the end of the paper

Pavlovian fear conditioning is a fundamental process in both health and disease. We investigate its neural correlates and sources of variability using harmonized functional magnetic resonance imaging data from 2199 individuals across nine countries, including 1888 healthy individuals and 311 with anxiety-related or depressive disorders. Using mega-analysis and normative modeling, we show that fear conditioning consistently engages brain regions within the "central autonomic–interoceptive" or "salience" network. Several task variables strongly modulate activity in these regions, contributing to variability in neural responses. Additionally, brain activation patterns differ between healthy individuals and those with anxiety-related or depressive disorders, with distinct profiles characterizing specific disorders such as post-traumatic stress disorder and obsessive-compulsive disorder. While the neural correlates of fear conditioning are highly generalizable at the population level, variability arises from differences in task design and clinical status, highlighting the importance of methodological diversity in capturing fear learning mechanisms.

Fear conditioning, also known as threat conditioning, is a psychological paradigm developed over a century ago to study associative learning mechanisms. It remains one of the most widely used and productive experimental models for investigating both normal and pathological fear and anxiety in humans[1]. Fear conditioning models how the association between an initially neutral stimulus (conditioned stimulus (CS)) and an innately aversive stimulus (unconditioned stimulus, US) is learned. The success of learning in fear conditioning is typically assessed by comparing responses to the fear cue (CS+, paired with the US) and the safety cue (CS-, not paired with the US) across subjective, autonomic, or neural domains. Successful conditioning is indicated by greater responses to the CS+ than to the CS-[2]. In the brain, this involves activity changes within a "central autonomic–interoceptive" or "salience" network, which in humans includes functionally and anatomically connected regions like the dorsal anterior cingulate cortex (dACC) and the anterior insular cortex (AIC)[3]. Additionally, fear conditioning has been linked to decreased activity in regions like the ventromedial prefrontal cortex (vmPFC), although such decreases have been less extensively studied[3]. Although the amygdala plays a crucial role in fear conditioning in rodents[4–6], and classical lesion studies have implicated the amygdala in human fear conditioning[7], this relationship has not been consistently identified in human fMRI studies[3,8–12].

Limitations in prior research on the neural correlates of human fear conditioning include the use of small sample sizes (typically $n < 30$) and the reliance on heterogeneous neuroimaging processing and analytical methods[3,13]. While group-level meta-analyses can partially address the sample size issue[3], individual-level *mega-analyses* offer additional advantages. These include enhanced statistical power, more precise effect size estimation, standardized preprocessing and analysis techniques, and substantially improved power to detect whether activation is modulated by individual variability -one of the primary goals of the current study[14–16].

Individual differences, such as sociodemographic factors (e.g., age) and trait variables (e.g., trait anxiety), are likely to modulate the neural correlates of fear conditioning, potentially affecting the generalizability of findings across groups, such as younger versus older adults or individuals with high versus low anxiety[13]. However, existing research on individual differences has been inconsistent and often hampered by limited sample sizes ($n < 50$[13]) or sampling biases[17].

e-mail: mafullana@clinic.cat

Moreover, task-specific variables, such as task instructions or characteristics of the US, may also influence conditioning at the behavioral or neural level[2,13,18,19]. For example, compared to other USs, a tactile electric shock may elicit greater activation in the dACC and the ventral supplementary motor area[3]. A primary challenge in this field is integrating prior data to accurately assess how individual differences and task variables affect neural outcomes. This complexity arises from variations in adjustable factors and sampling across studies and participants, highlighting the need for methods that can account for and isolate specific sources of variation—such as the normative modeling approach used here. Normative modeling allows us to integrate multiple smaller-scale studies into a common reference space—a standardised baseline against which to statistically quantify individual variations. This approach allows for meaningful comparisons across diverse studies by controlling for multiple sources of variation. As a result, the variance associated with specific variables and individuals can be isolated, quantified, and systematically analysed[20]. (For details on the underlying mathematics, see refs. 21–23; for applications, see refs. 24–29).

Fear conditioning has also been used to study the development and persistence of mental health disorders marked by pathological fear, such as anxiety-related disorders[1,30–33], which are highly prevalent and rank among the leading causes of disability worldwide[33]. However, there is ongoing debate over whether anxiety-related disorders consistently show abnormal fear conditioning at behavioral or neural levels[34,35] or if these abnormalities are specific to certain clinical groups —such as post-traumatic stress disorder (PTSD[36]) but not others, like social anxiety disorder (SAD)[37]. Inconsistencies maybe due in part to small sample sizes (ns < 100 for anxiety-related disorders as a group, ns < 25 for comparisons among clinical groups). Furthermore, most research in this area has relied on case-control designs and traditional analysis techniques, both of which have limitations that could be addressed through normative modeling. This framework enables statistical inference for individual subjects relative to an expected population pattern, providing a more detailed examination of the heterogeneity underlying group-level analyses[20].

In this study with pre-registered hypotheses and analyses (cf. Materials and Methods), we used both mega-analysis and normative modelling to analyse individual-level, harmonized fMRI data acquired during fear-conditioning from 43 samples from 21 laboratories across nine countries (total n = 2199), including both healthy participants and individuals diagnosed with anxiety-related and depressive disorders. First, we assessed the overall neural correlates of fear conditioning in healthy participants to provide a comprehensive delineation of the brain regions underlying human fear conditioning. Based on previous studies, we hypothesized that during fear conditioning, the CS + > CS- contrast would be associated with robust activations in regions such as the dACC, AIC, pre/supplementary motor areas (SMA), and dorsolateral prefrontal cortex (dlPFC), whereas the CS + < CS- contrast would be associated with deactivations in the vmPFC and hippocampus. We expected the mega-analysis to be more sensitive than previous studies in detecting subtle effects in other brain regions not previously (or not consistently) identified. Second, we assessed variation among healthy participants. Given their role in mediating subjective arousal and autonomic expression of fear[38], we hypothesised that regions including the vmPFC and the anterior-to-mid cingulate cortex would show the greatest between-subject heterogeneity. Third, we examined how individual differences (e.g., age, trait anxiety) and task variables (e.g., task instructions) influenced this variation. Finally, we explored differences in the neural correlates of fear conditioning between individuals with anxiety-related and depressive disorders and healthy controls, as well as among clinical subgroups (e.g., PTSD, SAD). We show that fear conditioning is consistently associated with brain activation in regions of the central autonomic-interoceptive network, despite methodological variations. However, specific task variables also influence the responses of these regions during conditioning. Additionally, brain activation patterns during conditioning differ between healthy individuals and those with anxiety-related or depressive disorders, with certain groups displaying distinct activation profiles.

Results

All results -including effect sizes for the linear models- are available in a free open-access repository (see Data availability statement).

Conditioning is associated with extensive brain (de)activations

In the mega-analysis (Fig. 1a), we included data from 1888 healthy individuals (42 experiment samples) and used linear mixed-effect models (LMMs) to perform a mega-analysis of whole-brain activation during fear conditioning (CS + > CS− contrast). We observed significant activation encompassing clusters within the bilateral anterior and mid insular cortices; the secondary somatosensory cortices (SII); the dlPFC; the lateral premotor cortices; and the dorsal and lateral cerebellum (Fig. 1b). Significant activation was also observed in multiple regions across the cortical midline, including the dACC extending to the pre-supplementary and SMA, ventral posterior cingulate cortex (PCC), and dorsal precuneus (dPrec).

Additionally, the CS + > CS- mega-analysis revealed the broad activation of subcortical regions, particularly the thalamus and basal ganglia. The largest of these activation patterns were observed in the dorsal striatum, specifically the caudate nucleus (CN); the globus pallidus extending to the striatum; the ventral tegmental area extending to the habenula; the mediodorsal thalamus (Thal); and the midbrain tegmentum. Activation of the midbrain was noted generally across the dorsal midbrain (- substantia nigra/red nucleus and pretectal nuclei) (Supplementary Fig. S1). To specifically assess the role of the amygdala, we conducted a Region of Interest (ROI) mega-analysis focusing on this region (see Materials and Methods), which indicated that neither the left (t = 1.93, p = 0.054, Cohen's d = 0.129, 95% CI [−0.002, 0.260]) nor the right amygdala (t = 1.57, p = 0.116, Cohen's d = 0.117, 95% CI [−0.029, 0.264]) showed significant activation during fear conditioning.

We also observed significant deactivations (CS + < CS- contrast) during fear conditioning, predominantly in regions of the default mode network (Fig. 1c). This included the PCC and precuneus; the vmPFC extending to the mPFC and subgenual cingulate cortex medially, as well as the left dorsal prefrontal cortex (dPFC); the bilateral angular gyri; and the parahippocampi and hippocampi (Hipp). Additional deactivation was observed in the lateral orbitofrontal cortex; the primary somatosensory cortex (SI); as well as the left temporal (TG) and fusiform gyri (see Supplementary Fig. S2 for detailed activation and deactivation across axial, sagittal, and coronal slices).

Heterogeneity in the neural correlates of conditioning

We estimated voxel-wise normative models of fear-conditioning related activation using the CS + > CS- contrast from 894 controls (training sample), and specifying age, biological sex, sample, and task variables as covariates (see Materials and Methods for all variables. The normative modeling sample is smaller than the mega-analysis due to the requirement for participants to have data on all covariates used in model construction). Testing these models with a held-out test sample (n = 646) showed good model fit with explained variance reaching 0.3 in regions that showed activation during fear conditioning (Fig. 1b), and skew and kurtosis within acceptable limits (Supplementary Fig. S3). For each participant in our held-out test sample, we then calculated a deviation score (z-score) within each voxel. In other words, for each participant, we quantified the distance from the predicted mean activation of each voxel, relative to the normative reference distribution for that voxel (Fig. 1d). While almost every voxel had at least 5 participants with large deviations (deviations > ±2.6), including areas, such as the bilateral insula and expanses of the

https://doi.org/10.1038/s41467-025-63078-x

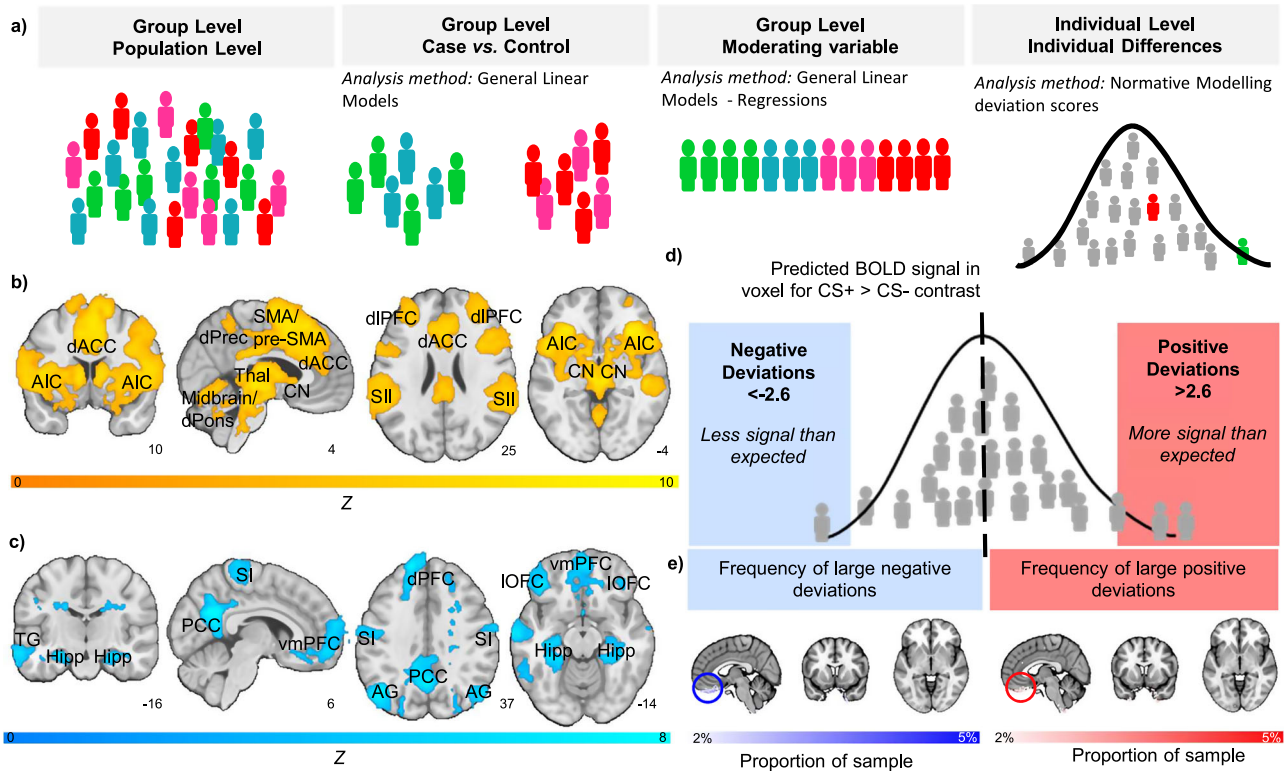


Fig. 1 | Neural correlates and individual-level heterogeneity in human fear conditioning. Schematic indicating the levels of analysis (**a**). Significant brain functional activation (**b**) and deactivation (**c**) to the CS+ versus CS− determined by mega-analysis ($n = 1888$ healthy controls). Schematic of normative modelling framework (**d**). Normative probability maps illustrate the percentage of participants in the healthy control test sample who had positive (hot colours -right) or negative deviations (cool colours - left) >±2.6 within each voxel. Circle highlights frequent large deviations (both positive and negative) within the most ventral region of the vmPFC (**e**). AIC anterior insular cortex, AG angular gyrus, CN caudate nucleus, dACC dorsal anterior cingulate cortex, dlPFC dorsolateral prefrontal cortex, dPFC dorsal prefrontal cortex, dPons dorsal pons, dPrec dorsal precuneus, Hipp hippocampus, HYP hypothalamus, lOFC lateral orbitofrontal cortex, PCC posterior cingulate cortex, SI primary somatosensory cortex, SII secondary somatosensory cortex, SMA supplementary motor area, TG temporal gyrus, Thal thalamus, vmPFC ventromedial prefrontal cortex. Source data are provided as a Source Data file.

cingulate cortex extending to the medial prefrontal cortex (Supplementary Fig. S4), controls most frequently had large deviations (both positive and negative) within the most ventral region of the vmPFC and inferior temporal pole. As this ventral region is notoriously prone to signal drop out, we interpret this result as most likely reflecting varying signal intensity rather than individual differences, and thus chose to interpret deviations within this region with caution (Fig. 1e).

Individual differences have small associations with conditioning

We examined the role of the following individual differences variables using LMMs and normative models (Fig. 1a): age, biological sex, and self-reported trait anxiety and depressive symptoms. In normative models, we analyzed both regression coefficients, reflecting each variable's contribution to the regression equation, and structure coefficients, indicating the direct bivariate relationship between a variable and brain activity without accounting for other predictors.

In LMMs, age ($n = 1884$ controls) and biological sex ($n = 1888$ controls) showed a significant association with brain activation or deactivation during fear conditioning (Supplementary Fig. S5). However, the effect sizes were small (Supplementary Discussion). Additionally, the age range was restricted (see Table 1). Regression and structure coefficients also showed minimal effects of age and biological sex ($n = 646$ controls) (Supplementary Fig. S5). Neither trait anxiety ($n = 1402$ controls), using either harmonised or non-harmonised scores (Supplementary Methods), nor depressive symptoms ($n = 213$ controls) were significantly associated with brain activation or deactivation during fear conditioning in LMMs. Similarly, elastic net regressions showed that whole-brain deviation scores derived from normative models could not explain the variance in individual levels of trait anxiety ($n = 751$ controls and cases; $r^2 = -0.095$, $p = 0.459$), nor depressive symptoms ($n = 152$ controls and cases; $r^2 = -0.257$, $p = 0.605$). See Methods for a note on negative r^2 values and Supplementary Table S1 for trait anxiety and depressive symptoms scores.

Task variables have a robust effect on conditioning

The influence of task variables on brain activation during fear conditioning was also examined using LMMs and structure coefficients from normative models in healthy controls. Several task variables were associated with robust effects across individuals, showing at least moderate effect sizes in LMMs and reaching significance in normative modeling analyses. These included instructions given to the participant about contingency prior to the task, the type of US, the use of a paradigm with multiple CSs (i.e., more than one CS+ or CS-), the pairing rate (i.e., percentage of CS+ followed by a US), and potential US confounding (i.e. whether trials followed by the US were included in the CS+ vs CS- contrast, and therefore the effects of the US may confound the effects of the CS+).

Partial instructions about CS-US contingency ($n = 1388$) were associated with significantly increased activation in the supplementary motor area and superior parietal lobule compared to no instructions about contingency ($n = 500$) in LMMs. Structure coefficients from the normative models ($n = 646$) showed that partial instructions (as compared to no instructions) produced a model predicting more activation in the bilateral anterior insula, the thalamus, the left caudate, clusters within the dorsomedial prefrontal cortex, the dorsolateral

https://doi.org/10.1038/s41467-025-63078-x

Table 1 | Descriptive statistics for all samples (N = 43) included in the analyses

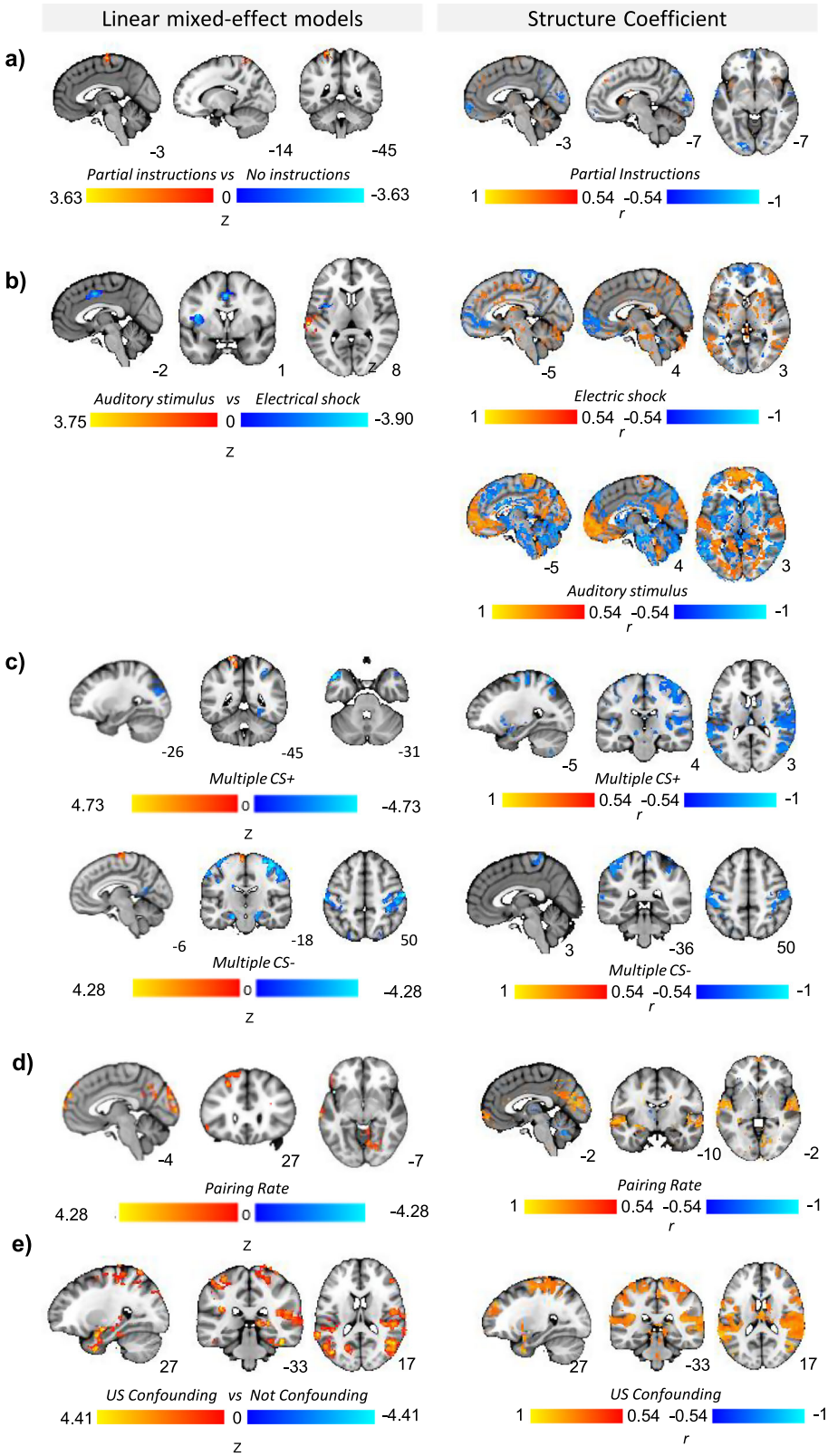
Sample	Country	*N*	Sex (%females)	Healthy Controls (n)	Patients (n)	Age M (SD) | Range (min-max)	Years of education M (SD) | Range (min-max)
Amsterdam_Visser/Kindt_sample_1	NL	18	72	18	0	22.06 (3.35) | 18–31	not available
Amsterdam_Visser/Kindt_sample_2	NL	41	73	41	0	20.56 (1.79) | 18–24	not available
Amsterdam_Visser/Kindt_sample_3	NL	12	75	12	0	21 (1.35) | 19-23	not available
Amsterdam_Visser/Kindt_sample_4	NL	10	80	10	0	22.8 (2.04) | 20–26	not available
Amsterdam_Visser/Kindt_sample_5	NL	13	85	13	0	22.23 (4.07) | 19–35	not available
Amsterdam_Visser/Kindt_sample_6	NL	14	79	14	0	23.43 (2.71) | 18–29	not available
Amsterdam_Visser/Kindt_sample_7	NL	16	44	16	0	24.06 (3.36) | 18–29	not available
Amsterdam_Visser/Kindt_sample_8	NL	9	100	9	0	20.33 (1.41) | 18–22	not available
Amsterdam_Visser/Kindt_sample_9	NL	38	58	38	0	23.66 (3.78) | 18–33	not available
Austin_Cisler	US	61	100	0	61	33.72 (8.48) | 21–50	15.46 (2.64) | 10–22
Barcelona_Cardoner	SP	71	66	45	26	22.66 (4.67) | 18-–40	14.49 (2.15) | 12–20
Barcelona_Soriano_sample_1	SP	35	51	17	18	37.43 (10.54) | 19–58	14.69 (3.72) | 6-18
Barcelona_Soriano_sample_2	SP	147	50	122	25	24.76 (4.22) | 19–36	18.63 (3.95) | 13–26
Bielefeld_Lonsdorf_sample_1	GE	116	66	116	0	24.61 (3.61) | 18–34	15.26 (2.14) | 1-16
Bielefeld_Lonsdorf_sample_2	GE	80	56	80	0	24.88 (3.51) | 19–34	not available
Bielefeld_Lonsdorf_sample_3	GE	28	64	28	0	24.68 (4.95) | 18–39	13.36 (1.75) | 11–20
Bochum_Elsenbruch	GE	29	48	29	0	26.45 (3.59) | 19–33	17.45 (4.02) | 3-23
Bochum_Merz_sample_1	GE	59	49	59	0	23.88 (4.17) | 18–34	16.07 (3.4) | 9-26
Bochum_Merz_sample_2	GE	59	47	59	0	24.39 (3.83) | 18–35	15.86 (3.72) | 5-23
Bochum_Merz_sample_3	GE	47	49	47	0	22.87 (2.61) | 19–30	not available
Bochum_Merz_sample_4	GE	29	0	29	0	24.21 (3.62) | 19–33	not available
Bochum_Merz_sample_5	GE	31	0	31	0	24.71 (3.87) | 20–34	not available
Bochum_Merz_sample_6	GE	60	50	60	0	23.57 (2.95) | 18–33	not available
Columbia_Neria	US	95	46	65	30	35.65 (12.26) | 18–60	15.11 (2.45) | 10–24
Duke_LaBar_sample_1	US	38	47	38	0	24.68 (4.2) | 19–35	not available
Duke_LaBar_sample_2	US	37	49	37	0	29.16 (11.07) | 19–66	not available
Florida_Keil	US	14	36	14	0	19.79 (2.08) | 18–26	14 (0) | 14–14
Harvard_McLaughlin	US	89	55	75	14	13.06 (2.6) | 8–17	7.04 (2.32) | 2–10
Manitoba_Greening_sample_1	CA	13	38	13	0	24 (5.07) | 19–36	17.15 (3.02) | 14–23
Manitoba_Greening_sample_2	CA	31	55	31	0	24.23 (4.56) | 17–33	not available
Melbourne_Harrison	AU	112	61	75	37	20.88 (2.34) | 16–25	15.02 (2.21) | 11–21
Munich_Koch	GE	45	56	23	22	34.47 (12.39) | 20-63	not available
Munster_Moeck_sample_1	GE	42	69	42	0	26.02 (6.22) | 19–51	12.33 (1.37) | 7-15
Munster_Moeck_sample_2	GE	29	52	29	0	15.79 (0.98) | 14–17	10.64 (0.99) | 8-12
Reading_Reekum_sample_1	UK	21	57	21	0	24 (2.59) | 21–31	not available
Reading_Reekum_sample_2	UK	50	60	50	0	17.8 (3.46) | 12–25	11.34 (1.82) | 8–14
MGH_Tuominen_sample_1	US	14	0	14	0	36.36 (9.61) | 22–49	15.69 (1.84) | 12–19
MGH_Tuominen_sample_2	US	37	43	37	0	28.51 (5.81) | 19–42	17.08 (2.27) | 12–23
USP_Diniz	BR	55	58	27	28	35.56 (10.97) | 19–63	13.13 (4.1) | 1–17
Texas_Dunsmoor	US	45	64	23	22	23.47 (4.51) | 18–37	NA
Ulm_Abler	GE	50	0	50	0	22.6 (2.92) | 18–29	NA
Uppsala_Ahs	SW	278	58	278	0	33.87 (10) | 20–58	14.16 (1.65) | 9–15
Vanderbilt_Kaczkurkin	US	81	0	53	28	33.47 (9.7) | 19–61	15.74 (2.18) | 13–20
Total n/Mean (SD)/Range		**2199**	**52.69**	**1888**	**311**	**25.26 (5.47) | 8–66**	**14.53 (2.56) | 1–26**

AU Australia, BR Brazil, *CA* Canada, *GE* Germany, *NA* Not available, *NL* The Netherlands, *SP* Spain, *SW* Sweden, *UK* United Kingdom, *US* United States. Note: To be included in the normative modelling analysis each participant had to have all essential data (age, sex) available, samples had to have control participants and larger samples required both genders available. These reasons lead to the exclusion of the entire Austin_Cisler and Vanderbilt_Kaczkurkin datasets, as well as 7 additional participants. The Bielefeld_Lonsdorf_sample_3 was not approved for inclusion in the normative modelling analysis. Thus, a total of 177 fewer participants were included in the normative modelling analys.

precuneus, and in the posterior region of the vmPFC. The model also predicted less activation within the bilateral visual cortex, the anterior medial temporal gyrus, and in the anterior vmPFC with the use of partial instructions (Fig. 2a). Note that we excluded instructed conditioning studies (Materials and Methods).

Compared with an auditory US ($n = 337$), a tactile electric shock US ($n = 1472$) produced significantly greater activation in bilateral dorsal mid-insula, dorsal medial thalamus, and pre-supplementary motor area, extending to the dACC ($n = 337$) in LMMs. In normative modelling analyses, a tactile electric shock US predicted increased activation within the dACC extending to the pre-supplementary motor area, the dorsal precuneus, secondary somatosensory cortex, the bilateral dorsal mid- to- posterior insula, the midbrain and pons, and the superior cerebellum, and less activation (i.e., more deactivation) within an expanse of the vmPFC, and SI. Moreover, the use of an auditory US was significantly associated with increased activation in the left auditory cortex and was predictive of increased activation in the bilateral auditory cortex (superior temporal lobe) and less

https://doi.org/10.1038/s41467-025-63078-x



deactivation (i.e., more differential activation) within an expanse of the vmPFC extending to the dorsomedial prefrontal cortex, PCC, angular gyrus, and SI (Fig. 2b).

In LMMs, compared to paradigms with a single CS+ ($n = 1283$), paradigms with multiple CS+ ($n = 605$) produced increased activation in the left supplementary motor area (SMA) and left dorsal precuneus and widespread increased deactivation in regions including the bilateral temporal poles, the right parahippocampal gyrus extending to the fusiform gyrus, the left visual association cortex extending to the angular gyrus, and the right primary motor and somatosensory cortex. Comparing paradigms with multiple CS- ($n = 302$) and those with a single CS- ($n = 1586$) revealed identical regions with increased activation to paradigms with multiple CS+. Conversely, increased deactivation was shown in the bilateral anterior hippocampus, ventral

Fig. 2 | Robust influence of task variables on brain activation during fear conditioning. Maps show the influence of pre-task instructions about CS-US contingency (**a**), type of US (**b**), number of CS used in paradigm (i.e. multiple CS+ or CS- or single CS+ or CS-) (**c**), pairing rate (**d**), and potential US confounding in CS + > CS- contrast (**e**) on mean activation (left; mega-analysis linear mixed-effects models) and relation to predicted activation (right; normative model structure coefficients). For the mega-analysis, warm colours indicate positive correlations (i.e., higher variable values associated with greater activation), while cool colours indicate negative correlations (i.e., higher variable values associated with reduced activation). For normative modeling, structure coefficient maps show the correlation coefficients (rho) thresholded by their respective coefficients of determination (rho2 > 0.3) of selected task variables. This can be interpreted as showing how predicted activation to the CS + > CS- contrast relates to the task variables included in the building of the normative models. Positive correlations (warm colours) indicate greater activation for higher values of the input variable and negative correlations (cool colours) greater activation for lower values of the input variable (note that some variables are dummy coded, e.g., pre-task instructions, type of US). CS Conditioned Stimulus; US Unconditioned Stimulus. For Pairing Rate (RR) in linear mixed-effects models, the figure shows significant results in the ANOVA comparing four categories (RR30, RR50, RR62, RR100). For the results of post-hoc tests, see Supplementary Figs. S6 and S7. Source data are provided as a Source Data file.

PCC, primary motor and somatosensory cortex, precuneus, and right mid-insula. In normative models, this was modelled using two variables (multiple CS+ and multiple CS-). Multiple CS+ predicted less activation within the bilateral amygdala, a bilateral expanse of SI the angular gyrus, the PCC, the bilateral putamen and caudate, and the lingual gyrus. Similarly, multiple CS- predicted decreased activation within a bilateral expanse of SI and the lingual gyrus (Fig. 2c).

Pairing rate, treated as a continuous variable, did not relate to brain activation during conditioning in LMMs. However, due to the non-normal distribution of pairing rates across studies and individuals, we categorized pairing rates (e.g., 30, 50, and 100%) and conducted ANOVA-like LLMs followed by pairwise comparisons with Holm-Bonferroni correction, which revealed significant effects (Fig. 2d). In particular, the comparisons involving the 50% pairing rate category was the category where significant differences between categories occurred most frequently. The significant differences between the pairing rate categories occurred both with (Supplementary Fig. S6) and without (Supplementary Fig. S7) US confounding. The structure coefficients for pairing rate (as a linear association) showed that a higher pairing rate predicted greater activation within visual regions (calcarine, lingual gyrus and cuneus), the precuneus, the left dorsolateral prefrontal cortex, the superior gyrus of the temporal lobe, and (less deactivation of) an anterior region of the vmPFC. Conversely, a higher pairing rate predicted less activation within the mid-cingulate cortex, the bilateral anterior insula, a posterior region of the vmPFC as well as the thalamus and caudate (Fig. 2d).

Finally, potential US confounding ($n = 997$), compared to no confounding ($n = 891$), was associated with significantly increased widespread activation during fear conditioning (CS + > CS- contrast). This activation was observed across the bilateral fusiform and lingual gyri, temporal poles, angular gyri, posterior insula, primary motor cortex, retrosplenial cortex (extending to the posterior hippocampus), and right amygdala, predominantly in the superficial amygdala, in linear mixed models (LMMs). Similarly, structure coefficients from the normative models showed that the model predicted greater activation within the bilateral mid-cingulate cortex extending to the dorsomedial prefrontal cortex and pre-supplementary motor area, angular gyri, mid-to-posterior insula, superior temporal gyrus and temporal poles, fusiform gyri and lateral mid-occipital gyrus, amygdala, caudate, dorsal thalamus, and dorsolateral cerebellum with potential US confounding (Fig. 2e).

None of the above results were affected by excessive multicollinearity, except for the association between pairing rate and the potential US confound (Supplementary Tables S5–S8). We identified six small clusters where the effects of both variables overlapped. To further disentangle their contributions, we conducted mixed-effects models within these clusters, including both variables as predictors. Results indicated that both variables exerted statistically significant effects in all clusters except for one small cluster in the right middle occipital region, where the effect of the US confound was no longer significant. Given the absence of multicollinearity (Variance Inflation Factor = 1.8), we concluded that activation in this region is specifically modulated by the pairing rate, rather than by the US confound.

The remaining task variables (for example, the number of trials during preconditioning) showed weaker effects or were not significantly associated with brain (de)activation during conditioning in the mega-analysis or normative modelling analyses (Supplementary Figs. S8 and S9 and Supplementary Discussion).

Cases and controls show differences in conditioning

In the mega-analysis, individuals with anxiety-related and depressive disorders (cases, $n = 311$) showed significantly increased activation in the right ventrolateral prefrontal cortex (anterior pars orbitalis), dorsal frontal pole, PCC, left temporal pole, and bilateral primary motor areas compared to controls ($n = 1888$) (Fig. 3a). Similar results were found when comparing individuals with anxiety-related disorders (i.e., excluding major depressive disorder; remaining $n = 297$) and controls, with additional clusters observed in the dorsal prefrontal cortex, visual association cortex, and primary somatosensory cortex (Supplementary Fig. S10). After excluding individuals who were taking medication at the time of the scan, those with anxiety-related and depressive disorders ($n = 221$) still showed significantly increased activation in the dorsal medial prefrontal cortex, dorsal PCC extending to the superior parietal lobule, left medial TG and bilateral ventrolateral prefrontal cortex compared to controls (Supplementary Fig. S11).

In normative modelling, we tested our clinical test sample (260 controls + 222 cases) against our reference normative models. This analysis compared the individuals' deviation scores (z-score) within each voxel, and quantified, as a percentage of the sample, the frequency of participants with large positive or large negative deviations (Fig. 3b). We compared the frequency of extreme deviations throughout the whole brain (Normative Probability Maps thresholded at > ±2.6), and found that cases had, on average, a greater frequency of extreme deviations than controls (Mann–Whitney U test = 111,167.5, $p = 0.014$; Fig. 3c). Qualitatively, cases showed a different pattern of deviation frequency than controls. Large deviations (i.e., more activity than would be predicted by the model) were common across cases within the dorsomedial prefrontal cortex, the primary somatomotor cortex, precuneus, the bilateral primary visual cortex (medial occipital lobe extending to the inferior medial and inferior lateral lobe) extending to the lingual and fusiform gyrus. As with controls, cases frequently had large negative deviations within the most ventral region of the vmPFC.

PTSD and OCD show distinct activation patterns and deviations

We divided our patient sample by primary diagnosis (PTSD, $n = 141$; OCD, $n = 68$; GAD, $n = 48$; and SAD, $n = 31$; other diagnoses were not included due to small sample size). ANOVA-like LMMs indicated that there were significant differences in brain activation during conditioning among patient groups. Post-hoc pairwise comparisons corrected for multiple comparisons showed that the most significant differences occurred between individuals with PTSD and OCD with respect to individuals with GAD and SAD (Supplementary Fig. S12).

Similarly, normative modelling analyses identified a significant difference in the frequency of large deviations among patient groups (Kruskal–Wallis H test = 71.529, $p = 1.98 \times 10^{-13}$; Fig. 3c). Follow-up

https://doi.org/10.1038/s41467-025-63078-x

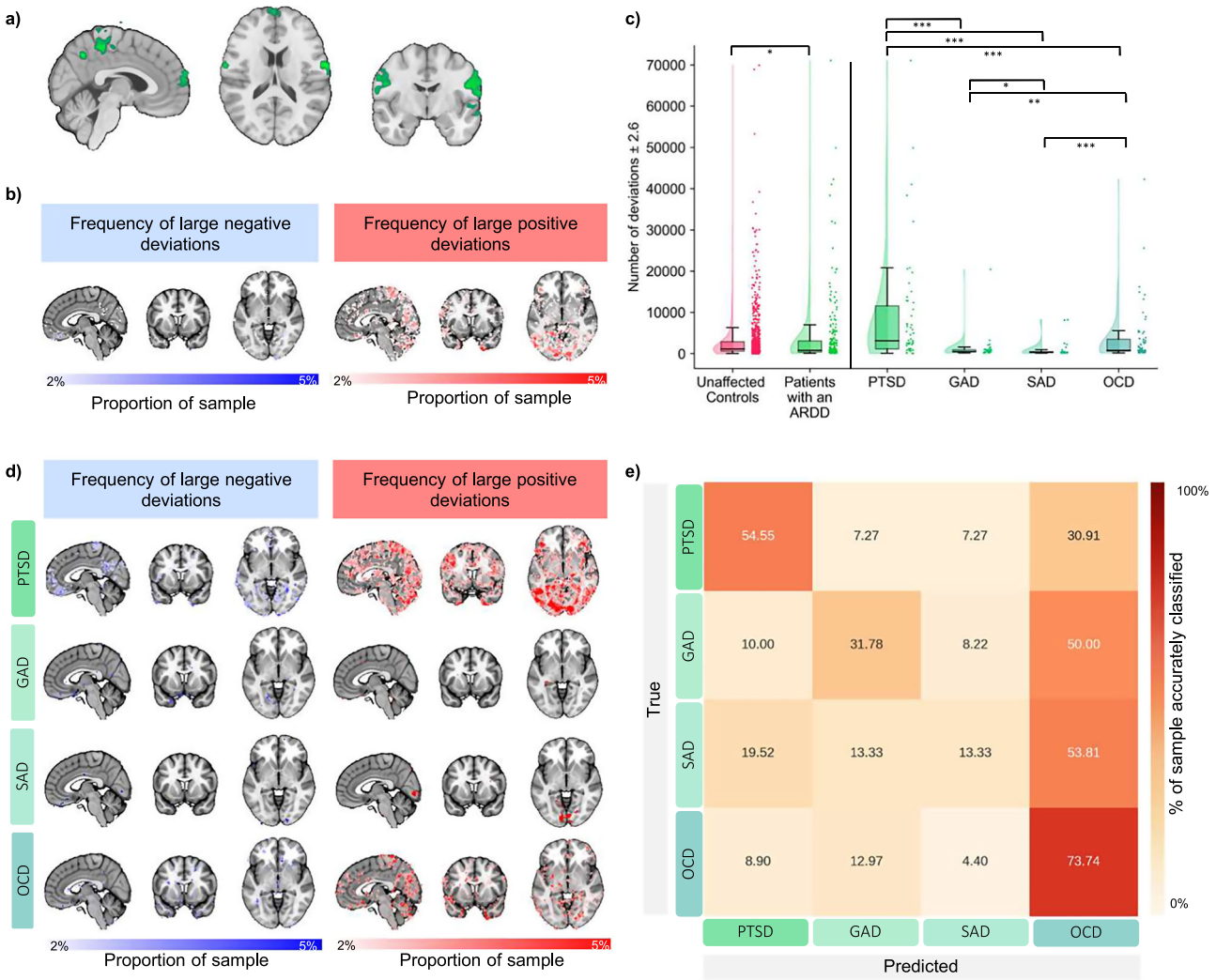


Fig. 3 | Differences between individuals with anxiety-related and depressive disorders and healthy controls in human fear conditioning. Regions wherein individuals with anxiety-related and depressive disorders ($n = 311$) showed significantly increased functional activation to the CS+ versus CS−, as compared to healthy controls (**a**). Normative probability maps illustrate the percentage of participants in the sample (test controls – top; individuals with anxiety-related and depressive disorders – bottom) who had positive (hot colours – right) or negative deviations (cool colours – left) > ±2.6 within each voxel (**b**). Box plots show the distribution of the total number of large deviations (> ±2.6) per group. The centre line indicates the median; box bounds represent the 25 and 75th percentiles (interquartile range, IQR); whiskers extend to the smallest and largest values within 1.5 × IQR from the lower and upper quartiles. Sample sizes: control group $n = 646$; PTSD $n = 55$; OCD $n = 68$; GAD $n = 48$; SAD $n = 31$; total clinical group $n = 202$ (**c**). Normative probability maps illustrate the percentage of each clinical group who had positive (hot colours – right) or negative deviations (cool colours – left) > ±2.6 within each voxel (**d**). Confusion matrix for multi-class support vector machine differentiating patterns of deviations among clinical groups (**e**). ARDD anxiety-related and depressive disorders, GAD Generalised Anxiety Disorder, OCD Obsessive Compulsive Disorder, PTSD Post-traumatic Stress Disorder, SAD Social Anxiety Disorder. * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.0001$. Kruskal–Wallis H-tests were used to test for main group effects (cases vs controls), with follow-up Mann–Whitney *U* tests false discovery rate (FDR) corrected at $\alpha = 0.05$. Source data are provided as a Source Data file.

Mann-Whitney *U* tests (FDR corrected for multiple comparisons) clarified, for example, that extreme deviations occurred most frequently in individuals with PTSD, as compared to other disorders, followed by OCD. We then illustrated the location of these extreme deviations at the voxel level to determine whether they were spatially overlapping within and between patient groups (Fig. 3d and Supplementary Fig. S13). Individuals with PTSD showed frequent large positive deviations within the bilateral medial occipital lobe extending to the inferior temporal lobe and lingual gyrus, bilateral vlPFC, an expanse of the dmPFC, precuneus, and bilateral amygdala. They also showed frequent large negative deviations within an expanse of the vmPFC (posterior vmPFC focus), precuneus, and a focus of the lingual gyrus and fusiform gyrus. There were very few regions wherein individuals with GAD showed overlapping large deviations, and similarly for SAD except for a small region of the bilateral lingual gyrus frequently found to have large positive deviations. Individuals with OCD showed frequent large deviations within the inferior parietal cortex, and temporal pole. A support vector machine could not classify cases from controls better than chance using whole-brain deviation maps (*mean* AUC = 0.44, $p = 1.0$, 95% CI [0.30, 0.58]). However, a multi-class support vector classifier confirmed a unique pattern of deviations among cases (Fig. 3e). More specifically, PTSD, on average, was accurately classified 54.55% of the time (*mean* F1 score = 0.58; $p = 2.06 \times 10^{-23}$; balanced accuracy = 0.43, where chance level across four classes = 0.25). Interestingly, despite fewer overlapping extreme deviations within the OCD group, the classifier was able to accurately label individuals with OCD 73.74% of the time (*mean* F1 score: 0.57; $p = 1.71 \times 10^{-7}$, 95% CI [0.54, 0.60]). GAD and SAD were only accurately classified 31.78% (*mean* F1 score: 0.35) and 13.33% (*mean* F1 score: 0.17) of the time, respectively, and were often misclassified as OCD. The

mean voxel-wise coefficient weights and frequency of contribution (in penalised permutations) to this classification are displayed in Supplementary Fig. S14.

Sample size for future studies

We conducted a series of sample size analyses to guide the design of future fMRI fear-conditioning studies (Supplementary Methods). To detect activation or deactivation in 50% of the brain regions identified in the mega-analysis (based on the AAL atlas[39]), a sample size of 122 was required, while detecting 80% of these regions required 243 participants (Supplementary Fig. S15). When considering activations only, the required sample sizes were slightly smaller: 100 participants to detect 50% and 199 participants to detect 80% of the mega-analytical findings. In contrast, substantially larger samples were needed to detect deactivations: 263 for 50% detection and 522 for 80%. The overall false positive rate was 9, 8, and 3% when activations and deactivations were assessed separately, averaging 5%. Additional sample sizes results are presented in Supplementary Figs. S16–S18.

Early and late conditioning

Given the importance of accounting for temporal dynamics in brain activity during human fear conditioning[8], we compared neural activation during the early and late phases of conditioning in a subset of participants ($n = 634$ controls; Supplementary Table S2). Consistent with the effects observed across the entire task, both phases showed significant activation in the CS + > CS- contrast across several brain regions. These included the insular cortices, SII, dlPFC, lateral premotor cortices, dorsal and lateral cerebellum, dACC extending to the pre-supplementary motor area and SMA, and the dPrec (Supplementary Fig. S19). Notably, there were several significant differences between the phases. The early phase showed greater activation in the bilateral fusiform gyrus, SMA, right amygdala, left frontal eye fields, and left motor cortex compared to the late phase (Supplementary Fig. S19). Additionally, significant differences were also observed in the left angular gyrus; dorsal, medial, and ventral anterior prefrontal cortices; and lateral orbitofrontal cortex. However, as these regions were implicated in the CS + < CS- contrast, this suggests that they exhibited reduced deactivation during the late phase.

Discussion

We compiled the largest ($n = 2199$) sample of individual-level fear conditioning fMRI data to date to comprehensively delineate the neural correlates of human fear conditioning, to assess the influence of several relevant sources of variation - including individual differences and task variables- and to evaluate potential differences in fear conditioning at the neural level between individuals with anxiety-related and depressive disorders and controls.

Our individual-level mega-analysis mapped fear conditioning activation to the “central autonomic–interoceptive” or “salience” network. As hypothesised, fear conditioning was associated with robust activations in the anterior insula, ventral striatum, pre-supplementary /supplementary motor areas, dorsal anterior cingulate cortex, and dorsolateral prefrontal cortex. It was also associated with activation in several subcortical regions, particularly the thalamus and basal ganglia. While many of the observed effects replicated previous findings[3], the increased statistical power provided by our analyses suggests that peak effects in the dorsal midbrain may originate in the substantia nigra/red nucleus and pretectal nuclei. Future work with a specific focus on these nuclei may aid in disentangling their specific contributions to fear conditioning. Also, as hypothesised, fear conditioning was associated with robust deactivations in the ventromedial prefrontal cortex and hippocampus. Other brain regions that were deactivated during conditioning included primarily regions of the default mode network (e.g., PCC and precuneus).

By incorporating a large sample from multiple laboratories worldwide, our study underscores the generalizability of the neural correlates of conditioning at the population level. At the same time, the methodological diversity across laboratories and studies suggests that our findings extend beyond specific experimental conditions, reinforcing their relevance to the broader fear conditioning process. Notably, at a time when neuroimaging research is increasingly emphasizing sample sizes in the thousands[40], our analyses show that studies with 100 participants can still reliably detect the neural correlates of fear conditioning, at least when considering activations only. Furthermore, our findings highlight that a significant source of variability in neural responses during fear conditioning stems from differences in task design. This insight is crucial for future human fMRI studies, as it enables researchers to anticipate the expected effects of various task and contrast design choices, along with their magnitudes, at the neural level. By making these adjustments in advance, researchers can strike a balance between the advantages of large, standardized studies and those of smaller studies with greater methodological diversity. Moreover, our normative modeling results underscore the potential of fear conditioning paradigms for informing the development of targeted interventions. Specifically, normative models can identify brain regions with atypical activation during conditioning, providing valuable guidance for interventions, such as neuromodulatory treatments aimed at these regions[41]. Additionally, by pinpointing abnormal activation patterns, normative models enable clinicians to tailor treatments more precisely to address these specific neural deviations.

The amygdala was not robustly activated during fear conditioning in either our mega-analysis or ROI-based mega-analysis for the contrast averaging across all trials, consistent with our previous group-level meta-analysis[3]. However, and in line with a recent study by Wen and colleagues[8] ($n = 601$, including individuals with anxiety-related disorders and controls), our analysis of early versus late trials in a large subsample of participants ($n = 634$ controls) revealed significantly greater activation in the right amygdala during early compared to late trials.

Inconsistencies regarding amygdala involvement in human fMRI conditioning studies have been attributed to several factors, including small sample sizes and limited anatomical specificity. The amygdala consists of distinct subregions, such as the basolateral (BLA) and centromedial (CMA) amygdala, and averaging responses may mask specific activations[8,10]. Moreover, the amygdala’s subcortical and ventral location, its small size, and the susceptibility artifacts and low signal-to-noise ratio associated with traditional imaging techniques can further hinder detection of significant effects[42]. Ultra-high field imaging has been shown to reduce these limitations and allows for more precise investigation of amygdala subnuclei[43,44], making it a valuable tool for future human fear conditioning studies.

Taken together with the findings of Wen and colleagues, our results highlight the importance of considering temporal dynamics when assessing amygdala activity during fear conditioning[8]. Specifically, they confirm that amygdala activation is strongest during early trials and habituates thereafter[45,46], suggesting that averaging across all conditioning trials may obscure these effects. In the current study, we also identified specific task features- such as the use of paradigms with multiple CS+ stimuli or US-related confounds- and diagnostic categories (e.g., PTSD; see also ref. 36) that modulate amygdala activity during conditioning. These findings underscore that both clinical and task-related variables may also contribute to the inconsistencies observed in the literature.

Biological sex had only minor effects, suggesting that fear conditioning mechanisms are relatively stable at the neural level between sexes. Additionally, none of our analyses found significant associations between brain activation during conditioning and levels of trait anxiety or depressive symptoms. While some mental health frameworks

suggest that dimensional constructs of psychopathology, like trait anxiety, may better reflect neural activation patterns[47], the variability and complexity in the neural states underlying these constructs and their lack of direct mapping to neural processes makes it challenging to identify clear linear relationships[48,49].

The brain activation differences during conditioning between individuals with anxiety-related and depressive disorders and healthy controls, observed in the mega-analysis, aligned with normative modeling results, showing a higher frequency of large deviations in cases. Importantly, these differences remained significant even after excluding medicated cases, suggesting that the observed effects are not due to medication. This is crucial, as commonly used treatments like selective serotonin reuptake inhibitors (SSRIs) can influence brain activation patterns observed with fMRI[50]. When the analysis was limited to anxiety-related disorders, significant differences in brain activation persisted, indicating that individuals with pathological anxiety are characterized by abnormal neural responses during fear conditioning. These findings suggest that such abnormalities could eventually serve as neural markers for anxiety-related disorders[51,52].

Among individuals with anxiety-related disorders, those with PTSD and OCD showed distinct patterns of brain activation and had distinct patterns of voxel-wise deviations that can be used to distinguish them from other anxiety-related disorders. This provides neurobiological support for the decision of current diagnostic classifications to separate these conditions[53]. In addition, it may provide valuable insights into the underlying mechanisms of psychopathology. The sample of individuals with PTSD was still relatively heterogeneous, with data from three independent samples, and yet there were often overlapping extreme positive deviations. Furthermore, using the derived deviation scores we were able to differentiate and classify individuals with PTSD and OCD with striking precision, compared to GAD and SAD. This is consistent with the previous literature that used mean averaging methods and reported differences in activation levels between groups of individuals with PTSD, compared to controls[36,54]. Taken together, this suggests that the neural mechanisms engaged during a fear conditioning paradigm are specifically relevant to the psychopathology of, and to some extent, similarly altered across individuals with PTSD; reinforcing the notion that fear conditioning is a foundational process in PTSD psychopathology, and as such, related tasks are a useful clinical model[31]. The accurate differentiation of OCD, despite few regions of overlapping large deviations, appeared to be driven by consistent coefficient weights with a region of the bilateral superior temporal gyrus and right vlPFC. Combined with no strong behavioural evidence[55], mixed imaging evidence of differences in fear conditioning tasks in this population[56–59], and evidence of altered baseline activity within the superior temporal region[60], this finding may be interpreted as capturing compensatory mechanisms that individuals with OCD engage to overcome obsessions and achieve the same behavioural output[55,60,61]. Despite significant differences in the frequency of extreme deviations between individuals with GAD and SAD compared to controls, their limited spatial overlap and less accurate classifications, suggest that there is significant heterogeneity in fear conditioning among individuals with these diagnoses. Thus, while we suggest that the psychopathology of PTSD is uniquely related to fear or threat processing as captured by fear conditioning tasks, we propose that other anxiety-related disorders, particularly GAD and SAD are less so.

Our study has several limitations. First, despite using harmonized pre-processing pipelines and statistical models to account for site differences, variations in diagnostic routines and imaging acquisition contributed to sample heterogeneity, particularly among individuals with anxiety and depressive disorders (a label that includes already heterogenous disorders). Second, mega-analyses may have limited power to detect effects in small subgroups (e.g., SAD patients). Third, for participants with a mental health diagnosis, we focused on primary diagnoses and we could not assess (or control for) comorbidity. Fourth, while our normative models adjusted for site, age, biological sex, and task influences on brain activity, future studies should explore the impact of adding more variables in the model construction. It is possible that alternative model structures could have increased the explained variance in the relatively noisy BOLD activation (where other literature has explained up to 51.3% of the variance[25]). However, care must be taken not to overfit or reduce the generalisability of models to ensure their wider utility. Fifth, we were unable to include other individual-level measures of conditioning (e.g., psychophysiological data) in our analyses, as this would have required separate collection and harmonization procedures. Finally, cross-sectional data on brain activation during fear conditioning raises concerns about the reliability of outcome measures. Although fMRI-based fear conditioning shows limited test-retest reliability at the whole-brain level, significant within-subject similarity across repeated time points has been observed[62], suggesting that large test-retest samples could help further validate the normative modeling approach, as demonstrated in other tasks[25].

With this work, we provide the largest analysis of the neural correlates of human fear conditioning and potential sources of variation to date. Our results confirm that human fear conditioning is a robust phenomenon at the neural level, consistently engaging multiple brain regions within the central autonomic-interoceptive or salience network. Our comprehensive review of the influence of task design choices on elicited and predicted brain activation can be used to help interpret differences in the previous literature and should remind researchers of the potentially significant influence of task design choices. Finally, we found that there are overall differences in fear conditioning at the neural level between individuals with anxiety-related and depressive disorders and controls, and that a unique mechanism of PTSD psychopathology is well captured by fear conditioning paradigms, supporting the use of these models to study this disorder.

Methods

The current manuscript combines two pre-registered analyses of individual-level fear conditioning fMRI data (https://osf.io/7n953; https://osf.io/w74bt). Data were collated from 43 samples originating from 23 sites in nine countries. Collation was coordinated by the lead group (IDIBAPS Barcelona). ENIGMA Fear Conditioning is part of the larger ENIGMA-Anxiety Working Group[63]. Table 1 summarizes the descriptive information on the samples. Informed consent was obtained from all participants by the sites providing their data. Some site-specific data have been reported previously, but no reports have examined all individual data together.

Inclusion and ethics

This study involved secondary analyses of previously collected human neuroimaging datasets. No new data were acquired specifically for the purposes of this study. All original studies received approval from their respective institutional ethics committees and were conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants in the original studies. The following Ethics Committees approved the individual studies: Ethics Committee of Ulm University, Ulm, Germany; Regional Ethics Review Board in Uppsala; Institutional Review Board of University Hospital of Bellvitge, Barcelona, Spain and Institutional Review Board of Hospital del Mar, Barcelona, Spain; Institutional Review Board University of Wisconsin; Institutional Review Board University of Arkansas for Medical Sciences; Comissão de Análise de Projetos de Pesquisa do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (CAPPesq); Institutional Review Board University of Texas at Austin; Ethics Committee of the University Hospital Essen, Germany; University of Southern California Institutional Review Board; Louisiana State

University Institutional Review Board; The University of Melbourne Human Research Ethics Committee; Institutional Review Board University of Minnesota; Institutional Review Board University of Florida; Ethics Committee of Klinikum Rechts der Isar, Technische Universität München; Duke University Health System Institutional Review Board; Ethics Committee of the General Medical Council Hamburg; University of Washington and Harvard University Institutional Review Boards; Ethics Committee of the Faculty of Medicine of the Ruhr University Bochum; Ethics Committee of the Faculty of Psychology and Sport Science, University of Giessen; University Research Ethics Committee of the University of Reading, UK; Clinical Research Ethics Committee (CEIC) of the Bellvitge University Hospital; Ethics Committee of the German Psychological Society; Partners HealthCare Institutional Review Board; Ethics Review Board (ERB) University of Amsterdam; Institutional Review Board, New York State Psychiatric Institute.

Inclusion and exclusion criteria were pre-specified within each dataset and were applied consistently during data aggregation. Sex was either self-reported by participants or recorded as sex assigned at birth, depending on the protocol of the original study. Sex was considered in the study design and used for descriptive statistics and group-level comparisons. Gender identity was not systematically collected across datasets and was therefore not analysed.

Fear conditioning task

We included data from participants who completed a fear conditioning experiment during an fMRI scan. There are several human fear conditioning paradigms, which vary based on the time elapsed between the CS and the US (e.g., delay, trace, simultaneous, or backward conditioning), the use of one (single-cue) versus two or more (differential-cue) CSs, and the instructions given to participants[2]: 1) *No instructions*: For example, "During this experiment, you will see various images and might experience mild electric shocks at certain times"; 2) *Partial instructions*: For example, "During this experiment, you may see a particular image sometimes followed by a mild electric shock. However, the shock won't happen every time you see the image, and sometimes it might not appear at all. Pay attention to the images, as they might give you some indication of when the shock could occur"; 3) *Full instructions* (instructed conditioning): For example, "During this experiment, you will see the image X, which is always followed by a mild electric shock. Whenever this image appears, it will be followed by the shock shortly afterward. No other images will be associated with the shock".

We focused on delay differential cue-conditioning paradigms with no or partial instructions (i.e., excluded instructed conditioning studies), and focused our analysis on comparing the response to a CS+ compared to a CS-. Table 2 summarises information on the fear conditioning tasks included in this manuscript.

Non-imaging data: sociodemographics and individual differences

All sites were asked to provide information regarding sociodemographics (age, biological sex) and individual differences: trait anxiety, assessed with the Trait subscale of the State-Trait Anxiety Inventory (STAI-T)[64]; and depressive symptoms, assessed with the Beck Depression Inventory (BDI)[65] (Supplementary Table S1). For individuals with anxiety-related and depressive disorders, sites were asked about principal mental health diagnosis and psychotropic medication use at the time of the scan (Supplementary Table S3). Previous normative studies of trait anxiety (STAI-T) have shown additive and multiplicative differences across countries, for which we harmonised trait anxiety scores across countries using ComBat[14] (Supplementary Methods) and conducted subsequent analyses twice: once with the raw scores and once with the country-harmonised scores.

Non-imaging data: task-related variables

We collected information about the following task variables: instructions given to the participant about contingency prior to the task (partial versus no information); type of US (e.g., electric shock versus aversive sound); number of trials during pre-conditioning; use of a paradigm with multiple CSs (i.e., more than one CS+ or CS-) during conditioning; type of CS (e.g. geometrical figures, faces, etc); number of CS+ and CS- trials during conditioning; average ITI (inter-trial interval); average ISI (inter-stimulus interval, i.e., between the CS+ and the US); pairing rate (percentage of CS+ followed by a US); potential US confounding; and the number of CS+ trials and CS- trials included in the fMRI contrast. For studies assessing awareness (conscious recognition of the association between the CS+ and the US, after the task), we also asked about participant´s contingency awareness (yes vs. no). Task variables were not explicitly listed in the pre-registration. The decision to include these variables was based on previous research[2,13].

Processing of neuroimaging data

We included only neuroimaging data acquired with whole-brain coverage. Individual-level raw task-based fMRI data were processed using the Harmonized Analysis of Functional MRI pipeline (HALFpipe, version 1.2.2)[66], a tool developed within the ENIGMA consortium to harmonise fMRI analyses across sites and facilitate reproducible analyses. HALFpipe provides a standardised workflow that extends fMRIprep[67] with several additional preprocessing steps, including spatial smoothing, grand mean scaling, temporal filtering, and confound regression. Moreover, HALFpipe generates a standardised quality assessment of the preprocessing outputs and imaging raw data (Supplementary Table S4). We used HALFPIPE default parameters (smoothing using 6 mm FWHM; confound removals using ICA-AROMA; and a high-pass filter of 125 s).

For the current study, each site was provided with a comprehensive manual to perform image pre-processing and quality control with HALFpipe in a fully harmonised manner, and each group shared the HALFPIPE output files for each individual along with the non-imaging data for each individual. The lead group (IDIBAPS-Barcelona) processed five sites, aggregated all the data, and carried out additional quality control procedures and measures to ensure the comparability of the data, as described in the Supplementary Methods).

Statistical analyses

We conducted two types of statistical analyses: mega-analyses and normative modelling analyses.

Mega-analyses

Participants. We included data from 2199 participants (*M* Age = 25.26, SD = 5.47; 57.2% female) comprising 1888 healthy controls (*M* Age = 25.85, SD = 8.51; 51.53% female) and 311 individuals with a primary diagnosis of an anxiety-related or depressive disorder (*M* Age = 29.91, SD = 10.75; 58.84% female) (Table 1 and Table 3). Diagnoses were established with structured clinical interviews.

Pre-scaling. Although we used the exact same processing protocol and conducted extensive quality control (see above), we observed differences in the BOLD response between samples, most likely due to varying units of measurement (note that MRI scans are acquired in arbitrary units[68]. To address these differences, we pre-scaled the images for healthy controls so that, for each sample, the voxel-wise-median standard deviation (after removing the effects of covariates) was 1 (see Supplementary Methods). We then applied the pre-scaling parameters obtained from the healthy controls to the cases (individuals with a primary diagnosis of an anxiety-related or depressive disorder). This approach differs from using the individual *z*-statistic images (i.e., dividing the BOLD response by its standard error), which we did not adopt for the mega-analysis. The reason is that the standard

https://doi.org/10.1038/s41467-025-63078-x

Table 2 | Characteristics of the fear conditioning tasks for each sample

Sample	CS+/CS- (n/n)	CS+ trials (n)	CS- trials (n)	Average ITI (ms)	Average ISI (ms)	Pairing rate (%)	CS type	Type of US	US confound	Assessment of awareness	Preconditioning phase
Amsterdam_Visser/Kindt_sample_1	2/2	22	22	22,000	6000	55	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_2	2/2	22	22	20,000	4000	55	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_3	2/2	18	18	17,500	4000	56	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_4	2/2	18	18	17,500	4000	56	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_5	2/2	18	18	10,350	4000	56	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_6	2/2	18	18	10,350	4000	56	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_7	2/2	18	18	4650	4000	56	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_8	2/2	18	18	17,500	4000	56	Neutral faces & pictures	Electric shock	no	yes	yes
Amsterdam_Visser/Kindt_sample_9	2/2	22	22	20,000	4000	55	Neutral faces & pictures	Electric shock	no	yes	yes
Austin_Cisler	1/1	18	18	4000	2500	50	Neutral pictures	Electric shock	no	yes	yes
Barcelona_Cardoner	1/1	32	32	5891	1900	50	Neutral pictures	Auditory stimulus	no	yes	yes
Barcelona_Soriano_sample_1	2/1	16	16	15,000	5800	62.5	Neutral pictures	Electric shock	yes	yes	yes
Barcelona_Soriano_sample_2	1/1	15	10	12,000	1750	33	Neutral pictures	Electric shock	no	yes	yes
Bielefeld_Lonsdorf_sample_1	1/1	14	14	13,000	6800	100	Neutral pictures	Electric shock	yes	yes	yes
Bielefeld_Lonsdorf_sample_2	1/1	14	14	13,000	7000	100	Neutral pictures	Electric shock	yes	no	yes
Bielefeld_Lonsdorf_sample_3	2/2	18	18	10,000	7000	100	Grey fractals	Electric shock	yes	yes	yes
Bochum_Elsenbruch	1/1	8	8	25,000	9000	100	Neutral pictures	Other*	yes	yes	no
Bochum_Merz_sample_1	2/1	16	8	10,750	8000	62.5	Neutral pictures	Electric shock	no	yes	no
Bochum_Merz_sample_2	2/1	16	8	10,750	8000	62.5	Neutral pictures	Electric shock	no	yes	no
Bochum_Merz_sample_3	1/1	21	21	12,000	8000	100	Neutral pictures	Electric shock	yes	yes	no
Bochum_Merz_sample_4	2/1	16	8	10,062	6000	62.5	Neutral pictures	Electric shock	no	yes	no
Bochum_Merz_sample_5	1/1	16	16	10,750	8000	62.5	Neutral pictures	Electric shock	no	yes	no
Bochum_Merz_sample_6	2/1	16	8	10,062	6000	62.5	Neutral pictures	Electric shock	no	yes	no
Columbia_Neria	1/2	15	30	3600	4000	80	Neutral pictures	Electric shock	yes	no	yes

https://doi.org/10.1038/s41467-025-63078-x

Table 2 (continued) | Characteristics of the fear conditioning tasks for each sample

Sample	CS+/CS- (n/n)	CS+ trials (n)	CS- trials (n)	Average ITI (ms)	Average ISI (ms)	Pairing rate (%)	CS type	Type of US	US confound	Assessment of awareness	Preconditioning phase
Duke_LaBar_sample_1	2/2	20	20	5750	6000	50	Avatars with neutral faces	Electric shock	yes	no	yes
Duke_LaBar_sample_2	1/1	16	16	15,900	4000	31	VR affective pictures	Electric shock	yes	no	yes
Florida_Keil	1/1	29	20	7000	5100	25	Gabor patches	Electric shock	yes	yes	yes
Harvard_McLaughlin	1/1	8	4	20,000	1500	40	Neutral pictures	Auditory stimulus	no	no	no
Manitoba_Greening_sample_1	1/1	24	24	12,000	6000	50	Gabor patches	Electric shock	no	no	yes
Manitoba_Greening_sample_2	1/1	24	24	12,000	3995	50	Gabor patches	Electric shock	no	no	yes
Melbourne_Harrison	1/1	15	10	12,000	1950	33	Neutral pictures	Auditory stimulus	no	yes	yes
Munich_Koch	1/1	8	8	12,000	12000	50	Affective faces and pictures	Electric shock	yes	no	no
Munster_Moeck_sample_1	1/1	27	27	5750	300	33	Neutral faces	Auditory stimulus	no	yes	yes
Munster_Moeck_sample_2	1/1	27	27	5750	300	33	Neutral faces	Auditory stimulus	no	yes	yes
Reading_Reekum_sample_1	1/1	12	12	10,530	500	100	Neutral pictures	Auditory stimulus	yes	no	no
Reading_Reekum_sample_2	1/1	12	12	10,530	500	100	Neutral pictures	Auditory stimulus	yes	no	no
MGH_Tuominen_sample_1	2/1	16	16	15,000	6000	62.5	Neutral pictures	Electric shock	yes	no	no
MGH_Tuominen_sample_2	1/1	8	8	15,000	6000	62.5	Neutral faces	Electric shock	yes	no	no
USP_Diniz	2/1	16	16	15,000	3000	62.5	Neutral pictures	Electric shock	yes	yes	no
Texas_Dunsmoor	1/1	24	24	6000	5000	50	Other**	Electric shock	yes	no	no
Ulm_Abler	2/1	80	20	variable	2500	50	Neutral pictures	Thermal stimulus	no	no	no
Uppshala_Ahs	1/1	16	16	14,000	6000	50	Humanoid characters	Electric shock	yes	yes	yes
Vanderbilt_Kaczkurkin	2/1	15	30	3600	3900	80	Neutral pictures	Electric shock	yes	yes	yes

CS conditioned stimulus, *CS+* CS followed by unconditioned stimulus, *CS−* CS not followed by unconditioned stimulus, *CS+/CS-* Number of different CS+ and CS-, *ITI* intertrial interval, *ISI* inter-stimulus interval, *US* Unconditioned stimulus. All samples used visual conditioned stimuli. All samples included an independent assessment of conditioning (e.g., skin conductance responses) except Amsterdam_Visser/Kindt_1. For all samples, the fMRI contrast (CS + > CS-) included either all CS+ trials (with US present) or all CS+ trials without the US, along with all CS- trials. Exceptions included Barcelona_Cardoner, Duke_LaBar_sample_1, and Duke_LaBar_sample_2, which only included trials from an early conditioning phase (n = 4CS + /4CS-, 5CS + /5CS-, and 8CS + /8CS- trials, respectively). *Rectal distension. ** Typical exemplars.

https://doi.org/10.1038/s41467-025-63078-x

Table 3 | Characteristics of individuals with anxiety-related and depressive disorders included in the analyses

Sample	N	Age M (SD)	Females (%)	Medication (%)	Comorbidity (%)	GAD (n)	MDD (n)	OCD (n)	PTSD (n)	SAD (n)	PD (n)	SP (n)
Austin_Cisler	61	33.72 (8.48)	100	59.02	67.21	0	0	0	61	0	0	0
Barcelona_Cardoner	26	23.88 (4.78)	61.54	3.85	11.54	26	0	0	0	0	0	0
Barcelona_Soriano_sample_1	18	40.56 (11.91)	61.11	88.89	50	0	0	18	0	0	0	0
Barcelona_Soriano_sample_2	25	25.56 (3.68)	64	0	16	21	0	0	0	4	0	0
Columbia_Neria	30	35.07 (13.82)	33.33	0	80	0	0	0	30	0	0	0
Harvard_McLaughlin	14	14.57 (2.14)	50	0	0	1	0	0	3	1	2	7
Melbourne_Harrison	37	19.89 (2.31)	51.35	0	56.76	0	11	0	0	26	0	0
Munich_Koch	22	33.55 (13.59)	59.09	54.55	27.27	0	0	22	0	0	0	0
USP_Diniz	28	33.68 (8.09)	53.57	0	71.43	0	0	28	0	0	0	0
Texas_Dunsmoor	22	25.95 (5.04)	68.18	NA	0	0	0	0	22	0	0	0
Vanderbilt_Kaczkurkin	28	34.57 (9.36)	0	3.57	32.14	0	3	0	25	0	0	0
Total n/M	311	29.91 (10.75)	58.84	21.22	44.05	48	14	68	141	31	2	7

Data refer to primary mental health diagnoses. "'Comorbidity' refers to the presence of at least one additional mental disorder. Data on comorbidity were not included in the analyses. *GAD* Generalized Anxiety Disorder, *MDD* Major Depressive Disorder, *NA* Not available, *OCD* Obsessive-Compulsive Disorder, *PD* Panic Disorder, *PTSD* Post-traumatic Stress Disorder, *SAD* Social Anxiety Disorder; *SP* Specific Phobia.

error may differ between cases and controls, and thus, differences in *z*-statistics between groups could reflect differences in the standard error rather than in the BOLD response (for more details, see Supplementary Methods).

Analyses. Differences in brain coverage across sites prevented us from using the standard ComBat method, which determines the harmonisation parameters using all voxels[14]. Additionally, there was no need to remove differences in scaling because we had already pre-scaled the images as described above. Thus, we used LMMs (with the sample as a random intercept) to investigate: 1st the pattern of brain activation during fear conditioning in healthy controls and in cases (individuals with anxiety-related and depressive disorders), which tested whether the mean activation in each voxel was non-null; 2nd the pattern of differences in brain activation during fear conditioning between cases and controls, which tested whether activation in each voxel was different between cases and controls; 3rd the pattern of differences in brain activation during fear conditioning among patient groups (PTSD, OCD, GAD, SAD), testing whether activation in each voxel differed among patient groups; 4th the potential influence of individual differences and task variables (see above) on brain activation during fear conditioning in healthy controls, which tested whether activation in each voxel was significantly associated with each individual differences or task variable. In all models, we incorporated age and sex as covariates. Significant LMMs comparing three or more groups (analog to ANOVAs) were followed by pairwise comparisons with Holm-Bonferroni correction.

We also conducted an ROI mega-analysis focusing on the amygdala. For this analysis, we extracted the pre-scaled BOLD response in the left and right amygdala based on the Automated Anatomical Labeling atlas[39]. We used an LMM, with age and sex as covariates, to test whether the mean activation significantly differed from zero. Potential differences between early and late conditioning were also analyzed using an LMM, with age and sex as covariates in a subsample of controls ($n = 679$; Supplementary Table S2).

We fitted the LMMs using custom functions (included in 'combat.enigma' R package) that call the 'nlme' R package voxel-wise and address voxel-specific details (e.g., varying collinearity due to differing brain coverage; see Supplementary Methods). FSL was then used to derive cluster-based corrected p-values using Gaussian Random Field (GRF) theory.

Analyses of multicollinearity. Given the diverse range of variables examined in this study—many of which may be influenced by methodological factors (e.g., pairing rate, type of conditioned stimuli) or sample characteristics (e.g., patient vs. control group)—there is a potential risk of confounding. That is, the observed effects attributed to one task variable may partially or wholly reflect the influence of another. To address this possibility, we systematically assessed interrelationships among all methodological and clinical variables using correlation analysis and evaluated multicollinearity using variance inflation factors (VIF)(Supplementary Tables S5-S8). For pairs of variables with correlation coefficients exceeding 0.5 (or η and Cramér's V when involving categorical variables), we further examined whether their associated activation maps exhibited spatial overlap. Overlap was defined as clusters of at least 10 contiguous voxels showing significant activation for both task contrasts. This approach was guided by the rationale that classical confounding requires both variables to be associated with activation in the same brain region. For any pair of correlated variables with overlapping activation, we re-estimated the mixed-effects linear models within the overlapping clusters, this time including both variables as predictors, to determine whether their effects remained statistically significant. A reduction to non-significance upon joint inclusion could indicate either collinearity (as suggested by the VIF) or potential confounding.

Effect sizes. To compare the effect sizes of different variables and to exclude findings with negligible or very small effects, we converted the regression coefficients of the peaks into correlation coefficients (Pearson r). For variables comparing two groups (e.g., cases vs. controls), we also calculated the corresponding standardised mean differences (Cohen's d). We considered effects with $r < 0.2$ (roughly equivalent to $d < 0.4$ for balanced binary variables) to be small, and only highlighted larger effects (i.e., $r > 0.2$, i.e., at least moderate) in the main text. It is important to note that peak effect sizes should be interpreted with caution, as they correspond to the peaks of clusters of statistical significance and are, therefore, larger than those obtained by other methods. Effect sizes for all the LMMs can be found at https://figshare.com/s/d44cc1390711bad3c147

Normative modelling analyses

Participants. We included data from 2022 participants; 1800 healthy controls (age range 8–66 years, M Age: 25.66 ± 8.4, 53.05% female) and 222 individuals with anxiety-related and depressive disorders (age range 9–63, M Age: 28.27 ± 11.06, 54.95% female) to build and test the normative models. See Table 1 note to explain discrepancy in participant numbers from mega-analysis.

Generating Normative Models of Activation to the CS + > CS- contrast. The z-statistic maps (files) from the CS + > CS- contrast for each participant were used as inputs (response variables) for the normative models. We created a normative model of fear-related activation per voxel, as a function of age, sex, and task variables (the same reported in the Non-imaging data: task-related variables section, except contingency awareness) by training a Gaussian Bayesian Linear Regression (BLR) model to predict activation for the CS + > CS- contrast[22]. We included dummy-coded site-related variables (sample, and MR strength) and a b-spline basis expansion as additional covariates of no-interest. This was performed in the Predictive Clinical Neuroscience toolkit (PCNtoolkit) software v0.26 (https://pcntoolkit.readthedocs.io/en/latest) implemented in Python 3.8. Generalisability was assessed by using a stratified train-test sample (train: 894, control test sample: 646).

Quantifying voxel-wise deviations from the normative model. To estimate a pattern of regional deviations from typical brain function for each participant in the control test sample ($n = 646$, mean age: 25.45 ± 7.19 years, 52.16% female), we derived a normative probability map (NPM) that quantifies the voxel-wise deviation from the normative model. The subject-specific Z-score indicates the difference between the predicted activation and true activation scaled by the prediction variance. This was repeated for the clinical test sample ($n = 482$, 260 controls + 222 cases, mean age: 26.76 ± 10.94 years, 54.97% female). We thresholded participants' NPM at $Z = \pm 2.6$ (i.e., $p < 0.005$) as in previous work[69–71] and summed the number of significantly deviating voxels for each participant. Kruskal-Wallis H-tests were used to test for group (cases or controls) and diagnosis effects and, when applicable, follow-up Mann–Whitney U tests were False Discovery Rate (FDR) corrected at $\alpha = 0.05$[72].

Normative models: individual differences and task variables

Model Coefficients. To probe the magnitude of the influence of individual differences (sociodemographics) and task variables on the predicted brain activation, we examined both the regression coefficients and the structure coefficients (correlation coefficients) of all sociodemographic and task variables input variables. Structure coefficients are preferable to regression coefficients when variables are collinear[73]. Note that negative r^2 values("negative" explained variance) are a possible outcome when the model fails to generalize effectively to new data, despite in-sample performance yielding non-negative explained variance (which is always positive or zero by construction). This phenomenon is not uncommon and arises when the model's predictions result in a residual sum of squares that exceeds the variance of the true values.

Linear Regression (Elastic Net) and Support Vector Classification (SVC). We applied an elastic net linear regression as implemented in the scikit-learn package (version 1.0.2)[74] with 10 repeats of nested 5-fold cross validation [alphas = 0.0001, 0.001, 0.01, 0.1, 0.3, 0.5, 0.7, 1; 90% train, 10% test split] to predict trait anxiety as measured by the STAI-T ($n = 751$), or depressive symptoms as measured by the BDI ($n = 440$) from participants' whole brain (unthresholded) deviation maps. The mean coefficient values and the frequency of the voxel's contribution (in other words, how many of the cross-folds split found this voxel to be important) indicate which voxel contributed to the prediction. The statistical significance of these results was tested against a 1000-fold nested 5-fold test for each variable. To classify participants ($n = 703$) who were contingency aware from those who were not based on their unthresholded whole-brain deviation maps, we used an SVC model with a linear kernel, regularisation parameter set to 1.0, and balanced class weights as implemented in the scikit-learn package (version 1.0.2).

Quantifying patterns of deviations between cases and controls. To classify individuals with anxiety-related or mood disorders and controls based on their whole brain unthresholded deviation maps, we used a SVC model with a linear kernel, regularisation parameter set to 1.0, as is common in neuroimaging, and balanced class weights (i.e. adjusted inversely proportional to class frequencies in the input data) as implemented in the scikit-learn package (version 1.0.2)[74]. The evaluation metric for the classification is area under the receiving operator curve (AUC) averaged across all folds within a 10-fold cross validation framework.

Quantifying patterns of deviations among patient groups. We used a one versus rest support vector classifier (SVC OvR) model as implemented in the scikit-learn package (sklearn.multiclass.OneVsRestClassifier version 1.0.2) to determine if there were quantifiably differentiable patterns within the whole brain unthresholded deviation maps among patient groups. Due to the small number of individuals with major depressive disorder ($n = 11$), specific phobia ($n = 7$), and panic disorder ($n = 2$), this analysis only included individuals diagnosed with PTSD ($n = 55$), OCD ($n = 68$), GAD ($n = 48$), and SAD ($n = 31$) (total $n = 202$). The model classes corresponded to the participants' diagnoses. The model classes were the participants' diagnosis. The evaluation metric for the classification was the F1-metric (the harmonic mean of precision and recall, also known as the balanced F-score, where values closer to 1 indicate greater classification success) per class within a 5-fold cross-validation framework, and the statistical significance was tested against a 1000-fold nested 5-fold test.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The results data generated in this study have been deposited at: https://doi.org/10.6084/m9.figshare.28540580.v1 The individual-level fMRI processed data (HALFIPE results files) are available for secondary data analysis. Access can be obtained by submitting an analysis plan to the ENIGMA-Anxiety Working Group (http://enigma.ini.usc.edu/ongoing/enigma-anxiety/). Data access is contingent on approval by PIs from contributing samples. The raw individual fMRI data are protected and are not available due to data privacy laws. Source data are provided with this paper.

Code availability

All code to reproduce the analyses in this manuscript is available at: https://doi.org/10.6084/m9.figshare.28540580.v1 The functions needed to conduct the mega-analysis are also included in the 'combat.enigma' R package.

References

1. Beckers, T. et al. Understanding clinical fear and anxiety through the lens of human fear conditioning. *Nat. Rev. Psychol.* **2**, 233–245 (2023).
2. Lonsdorf, T. B. et al. Don't fear 'fear conditioning': methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neurosci. Biobehav. Rev.* **77**, 247–285 (2017).
3. Fullana, M. A. et al. Neural signatures of human fear conditioning: an updated and extended meta-analysis of fMRI studies. *Mol. Psychiatry* **21**, 500–508 (2016).
4. Tovote, P., Fadok, J. P. & Lüthi, A. Neuronal circuits for fear and anxiety. *Nat. Rev. Neurosci.* **16**, 317–331 (2015).
5. LeDoux, J. The amygdala. *Curr. Biol.* **17**, R868–R874 (2007).
6. Johansen, J. P., Cain, C. K., Ostroff, L. E. & LeDoux, J. E. Molecular mechanisms of fear learning and memory. *Cell* **147**, 509–524 (2011).
7. Bechara, A. et al. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science* **269**, 1115–1118 (1995).
8. Wen, Z. et al. Temporally and anatomically specific contributions of the human amygdala to threat and safety learning. *Proc. Natl. Acad. Sci. USA* **119**, e2204066119 (2022).
9. Visser, R. M., Bathelt, J., Scholte, H. S. & Kindt, M. Robust BOLD responses to faces but not to conditioned threat: challenging the amygdala's reputation in human fear and extinction learning. *J. Neurosci.* **41**, 10278–10292 (2021).
10. Bach, D. R., Weiskopf, N. & Dolan, R. J. A stable sparse fear memory trace in human amygdala. *J. Neurosci.* **31**, 9383–9389 (2011).
11. Fullana, M. A. et al. Amygdala where art thou? *Neurosci. Biobehav Rev.* **102**, 430–431 (2019).
12. Morriss, J., Hoare, S. & van Reekum, C. M. It's time: a commentary on fear extinction in the human brain using fMRI. *Neurosci. Biobehav Rev.* **94**, 321–322 (2018).
13. Lonsdorf, T. B. & Merz, C. J. More than just noise: Inter-individual differences in fear acquisition, extinction and return of fear in humans - biological, experiential, temperamental factors, and methodological pitfalls. *Neurosci. Biobehav. Rev.* **80**, 703–728 (2017).
14. Radua, J. et al. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage* **218**, 116956 (2020).
15. Müller, V. I. et al. Ten simple rules for neuroimaging meta-analysis. *Neurosci. Biobehav. Rev.* **84**, 151–161 (2018).
16. Thompson, P. M. et al. ENIGMA and the individual: Predicting factors that affect the brain in 35 countries worldwide. *Neuroimage* **145**, 389–408 (2017).
17. Sjouwerman, R., Scharfenort, R. & Lonsdorf, T. B. Individual differences in fear acquisition: multivariate analyses of different emotional negativity scales, physiological responding, subjective measures, and neural activation. *Sci. Rep.* **10**, 15283 (2020).
18. Smith, M. C. CS-US interval and US intensity in classical conditioning of the rabbit's nictitating membrane response. *J. Comp. Physiol. Psychol.* **66**, 679–687 (1968).
19. Garcia, J. & Koelling, R. A. Relation of cue to consequence avoidance learning. *Psychon. Sci.* **4**, 123–124 (1966).
20. Segal, A. et al. Embracing variability in the search for biological mechanisms of psychiatric illness. *Trends Cogn. Sci.* **29**, 85–99 (2025).
21. Marquand, A. F. et al. Conceptualizing mental disorders as deviations from normative functioning. *Mol. Psychiatry* **24**, 1415–1424 (2019).
22. Fraza, C. J., Dinga, R., Beckmann, C. F. & Marquand, A. F. Warped Bayesian linear regression for normative modelling of big data. *Neuroimage* **245**, 118715 (2021).
23. Marquand, A. F., Rezek, I., Buitelaar, J. & Beckmann, C. F. Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biol. Psychiatry* **80**, 552–561 (2016).
24. Rutherford, S. et al. Evidence for embracing normative modeling. *Elife* **12**, e85082 (2023).
25. Savage, H. S. et al. Dissecting task-based fMRI activity using normative modelling: an application to the emotional face matching task. *Commun. Biol.* **7**, 814–888 (2024).
26. Wolfers, T. et al. Individual differences v. the average patient: mapping the heterogeneity in ADHD using normative models. *Psychol. Med.* **50**, 314–323 (2020).
27. Zabihi, M. et al. Fractionating autism based on neuroanatomical normative modeling. *Transl. Psychiatry* **10**, 384 (2020).
28. Tian, Y. E., Cole, J. H., Bullmore, E. T. & Zalesky, A. Brain, lifestyle and environmental pathways linking physical and mental health. *Nat. Ment. Health* **2**, 1250–1261 (2024).
29. Kjelkenes, R. et al. Deviations from normative brain white and gray matter structure are associated with psychopathology in youth. *Dev. Cogn. Neurosci.* **58**, 101173 (2022).
30. Pittig, A., Treanor, M., LeBeau, R. T. & Craske, M. G. The role of associative fear and avoidance learning in anxiety disorders: gaps and directions for future research. *Neurosci. Biobehav. Rev.* **88**, 117–140 (2018).
31. Fullana, M. A. et al. Human fear conditioning: from neuroscience to the clinic. *Behav. Res. Ther.* **124**, 103528 (2020).
32. Greco, J. A. & Liberzon, I. Neuroimaging of fear-associated learning. *Neuropsychopharmacology* **41**, 320–334 (2016).
33. Craske, M. G. et al. Anxiety disorders. *Nat. Rev. Dis. Prim.* **3**, 17024 (2017).
34. Marin, M.-F., Hammoud, M. Z., Klumpp, H., Simon, N. M. & Milad, M. R. Multimodal categorical and dimensional approaches to understanding threat conditioning and its extinction in individuals with anxiety disorders. *JAMA Psychiatry* **77**, 618–627 (2020).
35. Kausche, F. M., Carsten, H. P., Sobania, K. M. & Riesel, A. Fear and safety learning in anxiety- and stress-related disorders: an updated meta-analysis. *Neurosci. Biobehav. Rev.* **169**, 105983 (2024).
36. Suarez-Jimenez, B. et al. Neural signatures of conditioning, extinction learning, and extinction recall in posttraumatic stress disorder: a meta-analysis of functional magnetic resonance imaging studies. *Psychol. Med.* **50**, 1442–1451 (2020).
37. Savage, H. S., Davey, C. G., Fullana, M. A. & Harrison, B. J. Threat and safety reversal learning in social anxiety disorder - an fMRI study. *J. Anxiety Disord.* **76**, 102321 (2020).
38. Savage, H. S. et al. Neural mediators of subjective and autonomic responding during threat learning and regulation. *Neuroimage* **245**, 118643 (2021).
39. Tzourio-Mazoyer, 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).
40. Marek, S. et al. Reproducible brain-wide association studies require thousands of individuals. *Nature* **603**, 654–660 (2022).
41. Likhtik, E. & Johansen, J. P. Neuromodulation in circuits of aversive emotional learning. *Nat. Neurosci.* **22**, 1586–1597 (2019).
42. Boubela, R. N. et al. fMRI measurements of amygdala activation are confounded by stimulus correlated signal fluctuation in nearby veins draining distant brain regions. *Sci. Rep.* **5**, 10499 (2015).
43. Sladky, R. et al. High-resolution functional MRI of the human amygdala at 7 T. *Eur. J. Radio.* **82**, 728–733 (2013).

44. Kirstein, C. F., Güntürkün, O. & Ocklenburg, S. Ultra-high field imaging of the amygdala – A narrative review. *Neurosci. Biobehav. Rev.* **152**, 105245 (2023).
45. Blackford, J. U., Allen, A. H., Cowan, R. L. & Avery, S. N. Amygdala and hippocampus fail to habituate to faces in individuals with an inhibited temperament. *Soc. Cogn. Affect Neurosci.* **8**, 143–150 (2013).
46. Bas-Hoogendam, J. M. et al. Impaired neural habituation to neutral faces in families genetically enriched for social anxiety disorder. *Depress. Anxiety* **36**, 1143–1153 (2019).
47. Morris, S. E. et al. Revisiting the seven pillars of RDoC. *BMC Med.* **20**, 220 (2022).
48. Pessoa, L. How many brain regions are needed to elucidate the neural bases of fear and anxiety?. *Neurosci. Biobehav. Rev.* **146**, 105039 (2023).
49. Poldrack, R. A. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron* **72**, 692–697 (2011).
50. Armand, S. et al. Functional brain responses to emotional faces after three to five weeks of intake of escitalopram in healthy individuals: a double-blind, placebo-controlled randomised study. *Sci. Rep.* **14**, 3149 (2024).
51. Shackman, A. J. & Fox, A. S. Two decades of anxiety neuroimaging research: New insights and a look to the future. *Am. J. Psychiatry* **178**, 106–109 (2021).
52. Wen, Z., Marin, M.-F., Blackford, J. U., Chen, Z. S. & Milad, M. R. Fear-induced brain activations distinguish anxious and trauma-exposed brains. *Transl. Psychiatry* **11**, 46 (2021).
53. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. (American Psychiatric Association, 2013).
54. Kaczkurkin, A. N. et al. Neural substrates of overgeneralized conditioned fear in PTSD. *Am. J. Psychiatry* **174**, 125–134 (2017).
55. Cooper, S. E. & Dunsmoor, J. E. Fear conditioning and extinction in obsessive-compulsive disorder: a systematic review. *Neurosci. Biobehav. Rev.* **129**, 75–94 (2021).
56. Milad, M. R. et al. Deficits in conditioned fear extinction in obsessive-compulsive disorder and neurobiological changes in the fear circuit. *JAMA Psychiatry* **70**, 608–554 (2013).
57. Apergis-Schoute, A. M. et al. Neural basis of impaired safety signaling in obsessive compulsive disorder. *Proc. Natl. Acad. Sci. USA* **114**, 3216–3221 (2017).
58. Cano, M. et al. Neural correlates of fear conditioning and fear extinction and its association with cognitive-behavioral therapy outcome in adults with obsessive-compulsive disorder. *Behav. Res. Ther.* **144**, 103927 (2021).
59. Hearne, L. J. et al. Revisiting deficits in threat and safety appraisal in obsessive-compulsive disorder. *Hum. Brain Mapp.* **44**, 6418–6428 (2023).
60. Fan, J. et al. Spontaneous neural activity in the right superior temporal gyrus and left middle temporal gyrus is associated with insight level in obsessive-compulsive disorder. *J. Affect. Disord.* **207**, 203–211 (2017).
61. Westlin, C. et al. Improving the study of brain-behavior relationships by revisiting basic assumptions. *Trends Cogn. Sci.* **27**, 246–257 (2023).
62. Klingelhöfer-Jens, M., Ehlers, M. R., Kuhn, M., Keyaniyan, V. & Lonsdorf, T. B. Robust group- but limited individual-level (longitudinal) reliability and insights into cross-phases response prediction of conditioned fear. *Elife* **11**, e78717 (2022).
63. Bas-Hoogendam, J. M. et al. ENIGMA-anxiety working group: Rationale for and organization of large-scale neuroimaging studies of anxiety disorders. *Hum. Brain Mapp.* **43**, 83–112 (2022).
64. Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R. & Jacobs, G. A. Manual for the State-Trait Anxiety Inventory. (Consulting Psychologists Press, 1983).
65. Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. & Erbaugh, J. An inventory for measuring depression. *Arch. Gen. Psychiatry* **4**, 561–571 (1961).
66. Waller, L. et al. ENIGMA HALFpipe: interactive, reproducible, and efficient analysis for resting-state and task-based fMRI data. *Hum. Brain Mapp.* **43**, 2727–2742 (2022).
67. Gorgolewski, K. et al. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. *Front. Neuroinform.* **5**, 13 (2011).
68. Shinohara, R. T. et al. Statistical normalization techniques for magnetic resonance imaging. *Neuroimage Clin.* **6**, 9–19 (2014).
69. Wolfers, T. et al. Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry* **75**, 1146–1155 (2018).
70. Holz, N. E. et al. A stable and replicable neural signature of lifespan adversity in the adult brain. *Nat. Neurosci.* **26**, 1603–1612 (2023).
71. Floris, D. L. et al. Atypical brain asymmetry in autism—a candidate for clinically meaningful stratification. *Biol. Psychiatry Cogn. Neurosci. Neuroimag.* **6**, 802–812 (2021).
72. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57**, 289–300 (1995).
73. Kraha, A., Turner, H., Nimon, K., Zientek, L. R. & Henson, R. K. Tools to support interpreting multiple regression in the face of multicollinearity. *Front Psychol.* **3**, 44 (2012).
74. Pedregosa, F. et al. Scikit-learn: machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2012).

Acknowledgements

The research was supported by Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement de la Generalitat de Catalunya (No. 2021 SGR 1128, J.R.; M.A.F.), the Swedish Research Council (No. 2014-01160, F.Å.), the European Research Council (No. 648176, T.B.; ERC-2018 CoG-816564, D.B.; 101001118, A.F.M.), the National Institute of Mental Health (No. R01MH119132 and MH108753, J.C.; R01MH095904 and K23MH076054, D.J.H.; R01MH125615, A.K.; R01-MH103291, K.A.M.; 1R61MH129559-02, Y.N.; R01-AA031261, A.S.; K01MH122774, X.Z.; R01MH131532, B.S.J; R01MH125615, L.A.), the German Research Foundation (No. 316803389 - SFB 1280, S.E., C.J.M. and A.I.; LO 1980/1-1 and CRC-TRR58 subproject B07, T.B.L.; LO 1980/4-1, T.B.L. and M.R.E.; 521379614 – TRR393, B.S., U.D. and T.K.; 44541416 CRC-TRR58 T.S., K.R., K.D. and M.J.; 442075332 (RU 5187) and 461947532 (RU 5389), U.L.; WA 1539/11-1, H.W.;), the Economic and Social Research Council UK (No. ES/W000776/1, D.B.), the NSERC Discovery (No. RGPIN-2021-02906, S.G.), the National Institute of Health (No. R01 MH106574, C.L.; R01 MH122387, J.D.), the United states-Israel binational science foundation and NIA (No. 2P30AG064198, Y.N.), the National Institute of Mental Health Intramural Research (No. ZIA-MH-002781, D.P.), the Biotechnology and Biological Sciences Research Council (No. BB/L02697X/1, C.vR.), NIAAA (No. R01-AA030042, A.S.), the Instituto de Salud Carlos III (ISCIII) (No. PI16/00889 and PI19/01171, C.S.M.; FI22/00219, M.O.), Fundació Marató de TV3 (No. 202201-31, C.S.M.), Agència de Gestió d'Ajuts Universitaris i de Recerca (No. 2021SGR01017, C.S.M.), the Ministerio de Ciencia, Innovación y Universidades, Spain (No. PID2022-139081OB-C22, C.S.M.), the Department of Veteran Affairs (No. I01-CX002760, D.M.S.), the German Federal Ministry of Education and Research (No. 01EE1402E, B.S., T.K., A.P., U.D., K.D. and H-U.W.; 01EE2101, U.L.), Brain and Behavior Research Foundation NARSAD Young Investigator Award (X.Z.), "la Caixa" Foundation (No. LCF/BQ/IN17/11620071, V.P.A.), Spanish Ministry of Science, Innovation and Universities (No. JDC2022–048445-I, V.P.A), the National Eye Institute Core Grant (No. P30 EY001319, B.S.J.), the Medical Research Council (No. MR/J003980/1, J.M.), National Science Foundation Graduate Research Fellowship (No. DGE1745303, S.N.D.), the South African Medical Research Council (D.J.S.), the EU Innovative Medicines Initiative (IMI) 2 Joint Undertaking (IMI2 JU) (No. 101034377, NJVW), the Dutch

Research Agenda (NeuroLabNL–Small Projects for NWA routes 21/22) (No. NWA.1418.22.025, JM.B-H.), the Talent Acceleration grant (Medical Delta) (JM.B-H.), the NIH Big Data to Knowledge (BD2K) award (No. U54 EB020403 and R01MH131806, P.M.T.), the National Health and Medical Research Council of Australia (NHMRC) Project Grants (No. 1161897 and 1145010, B.J.H.).

Author contributions

J.R., H.S.S., E.V., and M.A.F. designed and performed analyses. J.R., H.S.S., A.J., and. M.A.F. wrote the manuscript. J.M.B-H., N.A.G., D.J.S., N.J.V.W., J.D., A.F.M., and. B.J.H. discussed the results. M.A.F. supervised research. J.R., H.S.S., E.V.,. A.J., B.A., F.Å., T.B., N.C., J.M.C., J.B.D., D.R.B., S.E., S.G., D.J.H., A.N.K., A.K., M.K., K.K., K.S.L., C.L.L., Cr.L.L., T.B.L., C.J.M., K.A.M., Y.N., D.S.P., C.R., A.J.S.,. C.S-M., V.I.S., D.M.S., B.S., T.S., L.T., R.M.V., L.A., V.A., M.C.B., P.R.B., E.E.B.,. M.C., P.C-E., S.E.C., U.D., V.P-A., S.N.D., K.D., M.R.E., J.L.G., A.O.H., M.J.H.,. A.A.H., A.I., A.J-S., M.J., T.K., K.K., M.K., F.L., S.M.L., M.L., U.L., J.M., I.M.-Z., R.M., J.M., M.O., A.P., D.P-C., J.R., I.C.R., W.R., K.R., J.R., A.N.R., R.S., J.S., A.S., B.S-J., M.U., H.U.W., X.Z., L.W., H.W., P.M.T., J.M.B-H., N.A.G., D.J.S., N.J.V.W.,. J.E.D., A.F.M., B.J.H., and M.A.F. commented on the manuscript.

Competing interests

Dr Stein has received consultancy honoraria from Discovery Vitality, Johnson & Johnson,. Kanna, L'Oreal,Lundbeck, Orion, Sanofi, Servier, Takeda and Vistagen. The other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-63078-x.

Correspondence and requests for materials should be addressed to Miquel A. Fullana.

Peer review information *Nature Communications* thanks Ajay Satpute, Candace Raio, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Joaquim Radua[1,2,3,96], Hannah S. Savage[4,5,6,96], Enric Vilajosana[1,2], Alec Jamieson[7], Birgit Abler[8], Fredrik Åhs[9], Tom Beckers[10,11], Narcís Cardoner[12,13,14], Josh M. Cisler[15], Juliana B. Diniz[16], Dominik R. Bach[17,18,19], Sigrid Elsenbruch[20], Steven G. Greening[21], Daphne J. Holt[22,23,24], Antonia N. Kaczkurkin[25], Andreas Keil[26], Merel Kindt[27], Kathrin Koch[28], Kevin S. LaBar[29], Charlene L. Lam[30,31], Christine L. Larson[32], Tina B. Lonsdorf[33,34], Christian J. Merz[35], Katie A. McLaughlin[36,37], Yuval Neria[38], Daniel S. Pine[39], Carien M. van Reekum[40], Alexander J. Shackman[41,42], Carles Soriano-Mas[3,43,44], Victor I. Spoormaker[45], Daniel M. Stout[46,47,48], Benjamin Straube[49,50], Thomas Straube[51], Lauri Tuominen[52], Renée M. Visser[27], Laura Ahumada[26], Volker Arolt[53,54], Marcelo C. Batistuzzo[16,55], Paulo R. Bazán[56], Emma E. Biggs[10], Marta Cano[12,13], Pamela Chavarría-Elizondo[44,57], Samuel E. Cooper[58], Udo Dannlowski[53], Víctor de la Peña-Arteaga[12,13,43], Stephanie N. DeCross[36], Katharina Domschke[59,60], Mana R. Ehlers[33,34], John L. Graner[29], Alfons O. Hamm[61], Martin J. Herrmann[62], Ashley A. Huggins[63], Adriane Icenhour[20], Asier Juaneda-Seguí[44], Markus Junghoefer[54,64], Tilo Kircher[49,50], Katja Koelkebeck[65], Manuel Kuhn[34,66,67], Franziska Labrenz[20], Shmuel M. Lissek[68], Martin Lotze[69], Ulrike Lueken[60,70], Jürgen Margraf[71], Ignacio Martínez-Zalacaín[44,72], Robert Moeck[51], Jayne Morriss[73], María Ortuño[1], Andre Pittig[74], Daniel Porta-Casteras[12], Jan Richter[75], Isabelle C. Ridderbusch[49,50], Winfried Rief[76], Kati Roesmann[64,77], Jörgen Rosén[78], Alena N. Rußmann[34,79], Rachel Sjouwerman[34,80], Jennifer Spohrs[81,82], Andreas Ströhle[83], Benjamin Suarez-Jimenez[84,85], Martin Ulrich[8], Hans-Ulrich Wittchen[86], Xi Zhu[38,87], Lea Waller[83], Henrik Walter[60,83], Paul M. Thompson[88], Janna Marie Bas-Hoogendam[89,90,91], Nynke A. Groenewold[92], Dan J. Stein[93], Nic J. Van der Wee[90,91], Joseph E. Dunsmoor[94,97], Andre F. Marquand[4,5,97], Ben J. Harrison[7,97] & Miquel A. Fullana[1,95,97] ✉

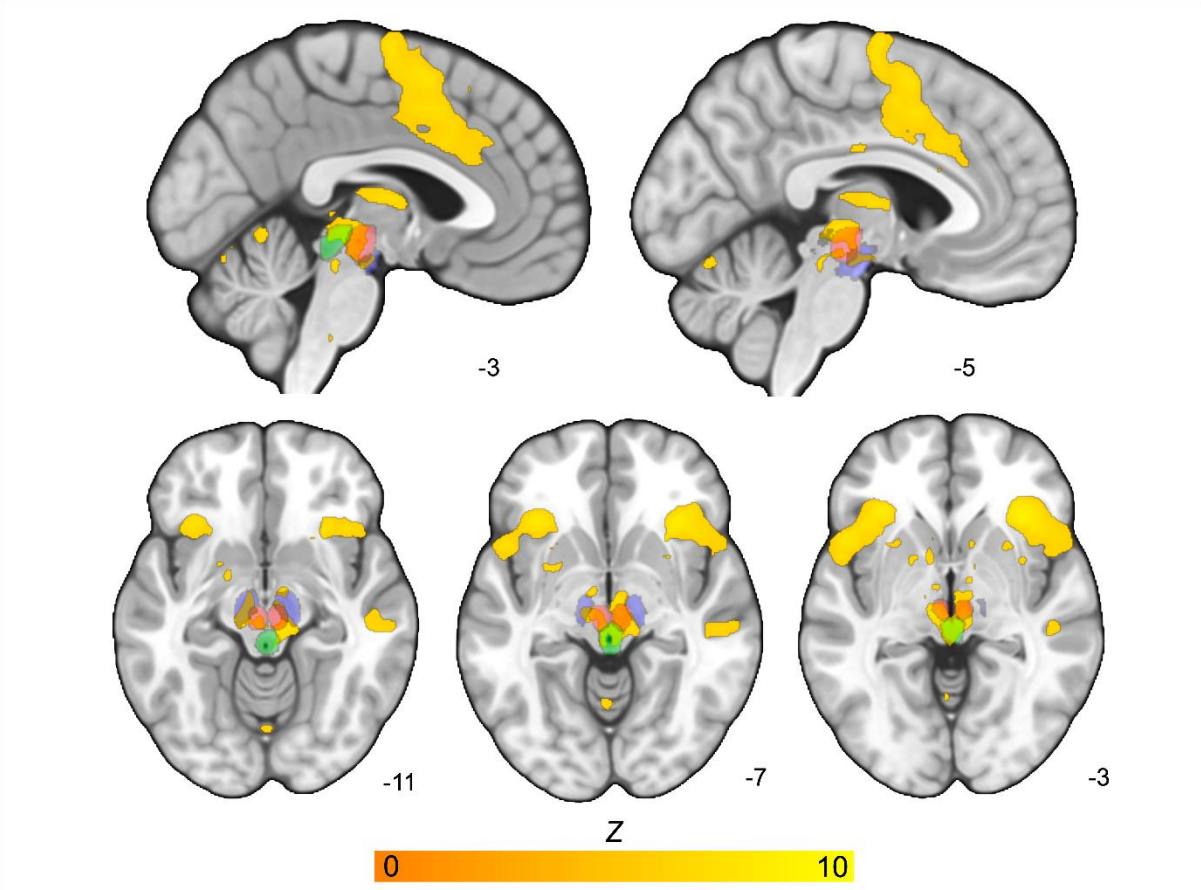
[1]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain. [2]University of Barcelona, Barcelona, Spain. [3]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Instituto de Salud Carlos III (ISCIII), Madrid, Spain. [4]Donders Institute of Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands. [5]Department of Cognitive Neuroscience, Radboud University Medical Centre, Nijmegen, The Netherlands. [6]Institute of Cognitive Neuroscience, UCL, London, UK. [7]Department of Psychiatry, The University of Melbourne, Parkville, VIC, Australia. [8]Department of Psychiatry and Psychotherapy III, Ulm University Clinic, Ulm, Germany. [9]Department of Psychology and Social Work, Mid Sweden University, Östersund, Sweden. [10]Faculty of Psychology and Educational Sciences, KU Leuven Leuven, Belgium. [11]Leuven Brain Institute, KU Leuven Leuven, Belgium.

[12]Sant Pau Mental Health Group, Institut d'Investigació Biomèdica Sant Pau (IBB-Sant Pau), Barcelona, Spain. [13]Hospital de la Sant Creu i Sant Pau, Barcelona, Spain. [14]Universitat Autònoma de Barcelona (UAB), Barcelona, Spain. [15]Department of Psychiatry and Behavioral Sciences, Dell Medical School, University of Texas at Austin, Austin, TX, USA. [16]Department of Psychiatry, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil. [17]University of Bonn, Transdisciplinary Research Area "Life and Health", Hertz Chair for Artificial Intelligence and Neuroscience, Institute of Computer Science, Medical Faculty, Institute of Experimental Epileptology and Cognition Research, Bonn, Germany. [18]University Hospital Bonn, Department of Psychiatry and Psychotherapy, Bonn, Germany. [19]Department for Imaging Neuroscience, UCL Queen Square Institute of Neurology, University College, London, UK. [20]Department of Medical Psychology and Medical Sociology, Medical Faculty, Ruhr University, Bochum, Germany. [21]Brain & Cognitive Sciences, Department of Psychology, University of Manitoba, Winnipeg, Canada. [22]Department of Psychiatry, Massachusetts General Hospital, Charlestown, MA, USA. [23]Harvard Medical School, Boston, MA, USA. [24]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, MA, USA. [25]Department of Psychology, Vanderbilt University, Nashville, TN, USA. [26]Department of Psychology, University of Florida, Gainesville, FL, USA. [27]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands. [28]School of Medicine and Health, Technical University of Munich, Munich, Germany. [29]Center for Cognitive Neuroscience, Duke University, Durham, NC, USA. [30]Laboratory of Clinical Psychology & Affective Neuroscience, Department of Psychology, University of Hong Kong, Hong Kong, China. [31]State Key Laboratory of Brain and Cognitive Sciences, Hong Kong, China. [32]University of Wisconsin-Milwaukee, Department of Psychology, Milwaukee, WI, USA. [33]Biological Psychology and Cognitive Neuroscience, Bielefeld University, Bielefeld, Germany. [34]Institute for Systems Neuroscience, University Medical Center Hamburg Eppendorf, Hamburg, Germany. [35]Department of Cognitive Psychology, Institute of Cognitive Neuroscience, Ruhr University Bochum, Bochum, Germany. [36]Department of Psychology, Harvard University, Cambridge, MA, USA. [37]Ballmer Institute for Children's Behavioral Health, University of Oregon, Eugene, OR, USA. [38]Department of Psychiatry, Columbia University Irving Medical Center, New York, NY, USA. [39]Section on Development and Affective Neuroscience, National Institute of Mental Health, Bethesda, MD, USA. [40]School of Psychology and Clinical Language Sciences, University of Reading, Reading, UK. [41]Department of Psychology, Neuroscience and Cognitive Science Program, University of Maryland, College Park, MD, USA. [42]Maryland Neuroimaging Center, University of Maryland, College Park, MD, USA. [43]Department of Social Psychology and Quantitative Psychology, Institute of Neurosciences, University of Barcelona, Barcelona, Spain. [44]Psychiatry and Mental Health Group, Neuroscience Program, Institut d'Investigació Biomèdica de Bellvitge - IDIBELL, L'Hospitalet de Llobregat, Spain. [45]Max Planck Institute of Psychiatry, Munich, Germany. [46]University of California, San Diego, CA, USA. [47]VA San Diego Healthcare System, San Diego, CA, USA. [48]Center of Excellence for Stress and Mental Health, San Diego, CA, USA. [49]Department of Psychiatry and Psychotherapy, University of Marburg, Marburg, Germany. [50]Center for Mind, Brain and Behavior (CMBB), University of Marburg, Marburg, Germany. [51]Institute of Medical Psychology and Systems Neuroscience, University of Muenster, Muenster, Germany. [52]University of Ottawa Institute of Mental Health Research at the Royal, Ottawa, ON, Canada. [53]Institute for Translational Psychiatry, University of Münster, Münster, Germany. [54]Otto-Creutzfeld-Center of Cognitive Neuroscience, University of Münster, Münster, Germany. [55]Department of Methods and Techniques in Psychology, Pontifical Catholic University, São Paulo, Brazil. [56]Radiology Institute, Faculdade de Medicina, Universidade de São Paulo, São Paulo, Brazil. [57]Department of Clinical Sciences, School of Medicine, Universitat de Barcelona - UB, L'Hospitalet de Llobregat, Barcelona, Spain. [58]Department of Psychiatry and Behavioral Sciences, University of Texas at Austin, Austin, TX, USA. [59]Department of Psychiatry and Psychotherapy, Medical Center - University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany. [60]German Center of Mental Health (DZPG), partner site Berlin/Potsdam, Berlin, Germany. [61]Department of Psychology, University of Greifswald, Greifswald, Germany. [62]Department of Psychiatry, Psychosomatics and Psychotherapy, University of Würzburg, Würzburg, Germany. [63]Department of Psychology, University of Arizona, Tucson, AZ, USA. [64]Institute for Biomagnetism and Biosignal analysis, University of Münster, Münster, Germany. [65]Bielefeld University, Medical School and University Medical Center OWL, Protestant Hospital of the Bethel Foundation, Department of Psychiatry and Psychotherapy, Bielefeld, Germany. [66]Center for Depression, Anxiety, and Stress Research, McLean Hospital, Belmont, MA, USA. [67]Department of Psychiatry, Harvard Medical School, Cambridge, MA, USA. [68]Department of Psychology, University of Minnesota, Minneapolis, MN, USA. [69]Diagnostic Radiology, University Medicine Greifswald, Greifswald, Germany. [70]Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany. [71]Mental Health Research and Treatment Center, Ruhr-University Bochum (FBZ) and German Center for Mental Health (DZPG), partner site Bochum/Marburg, Bochum, Germany. [72]Radiology Department, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Barcelona, Spain. [73]School of Psychology, Faculty of Environmental and Life Sciences, University of Southampton, Southampton, UK. [74]Translational Psychotherapy, Institute for Psychology, University of Göttingen, Göttingen, Germany. [75]Department of Experimental Psychopathology, University of Hildesheim, Hildesheim, Germany. [76]Division of Clinical Psychology and Psychotherapy, University of Marburg, Marburg, Germany. [77]Institute of Psychology, Clinical Psychology in Childhood and Adolescence, University of Osnabrück, Osnabrück, Germany. [78]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden. [79]Department of Psychology, Clinical Psychology and Neuroscience, University of Hamburg, Hamburg, Germany. [80]Experimental Health Psychology, Maastricht University, Maastricht, The Netherlands. [81]Department of Child and Adolescent Psychiatry and Psychotherapy, Ulm University Medical Centre, Ulm, Germany. [82]Department of Psychiatry, Psychotherapy and Psychotraumatology Military Medical Centre, Ulm, Germany. [83]Department of Psychiatry and Psychotherapy, Charité - Universitätsmedizin Berlin, Berlin, Germany. [84]Department of Neuroscience, University of Rochester Medical Center, Rochester, NY, USA. [85]Center for Visual Sciences, University of Rochester, Rochester, NY, USA. [86]Psychiatric University Hospital, Ludwig-Maximilians-University München, München, Germany. [87]New York State Psychiatric Institute, New York, NY, USA. [88]Imaging Genetics Center, Stevens Neuroimaging & Informatics Institute, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA. [89]Developmental and Educational Psychology, Institute of Psychology, Leiden University, Leiden, The Netherlands. [90]Leiden Institute for Brain and Cognition, Leiden University Medical Center, Leiden, The Netherlands. [91]Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands. [92]Department of Psychiatry and Mental Health, Neuroscience Institute, University of Cape Town, Cape Town, South Africa. [93]SAMRC Unit on Risk & Resilience in Mental Disorders, Department of Psychiatry and Mental Health, Neuroscience Institute, University of Cape Town, Cape Town, South Africa. [94]Department of Neuroscience, University of Texas at Austin, Austin, TX, USA. [95]Adult Psychiatry and Psychology Department, Institute of Neurosciences, Hospital Clinic, Barcelona, Spain. [96]These authors contributed equally: Joaquim Radua, Hannah S. Savage. [97]These authors jointly supervised this work; Joseph E. Dunsmoor, Andre F. Marquand, Ben J. Harrison, Miquel A. Fullana. ✉e-mail: mafullana@clinic.cat

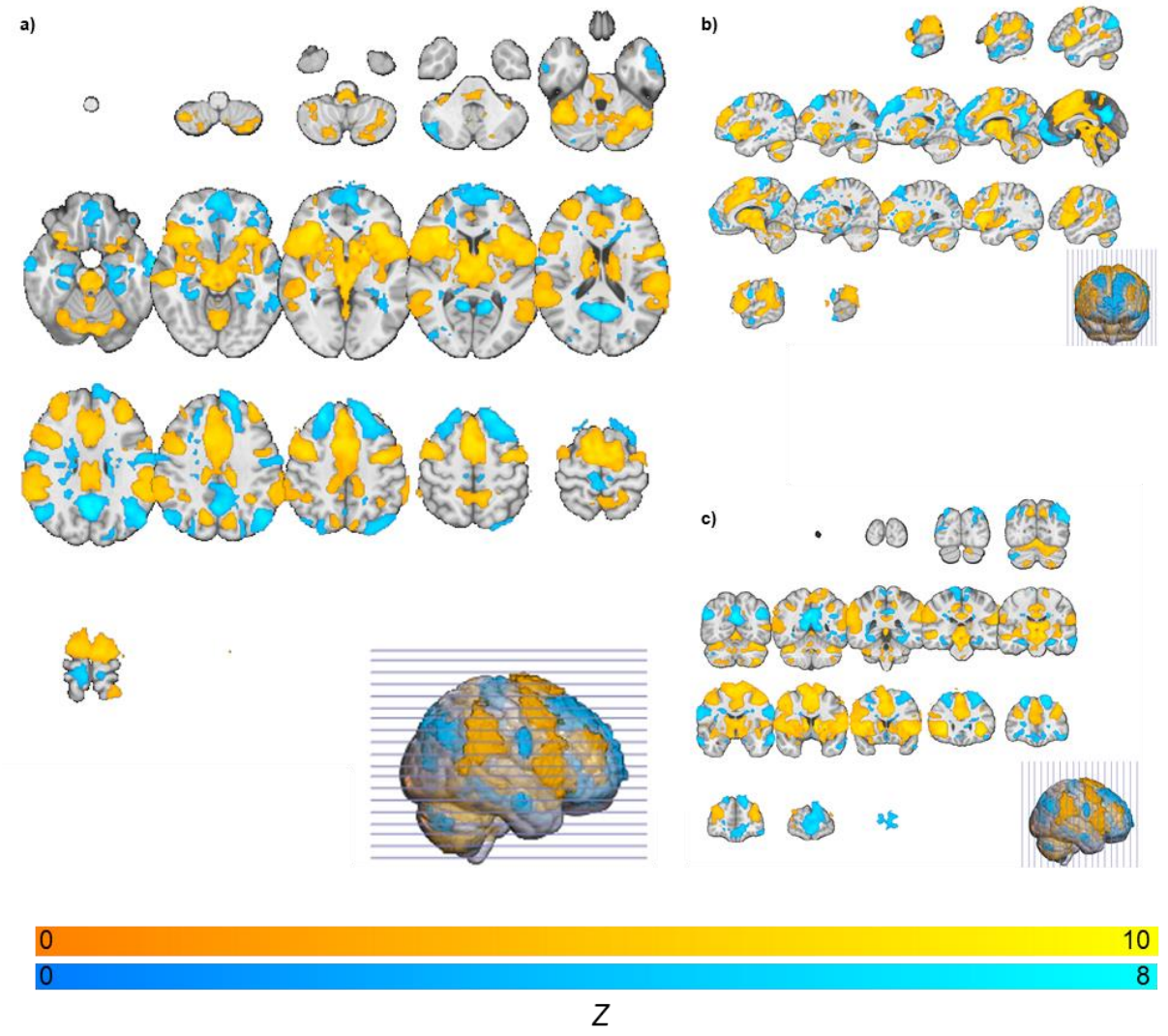
Neural correlates of human fear conditioning and sources of variability in 2199 individuals.

Supplementary Information

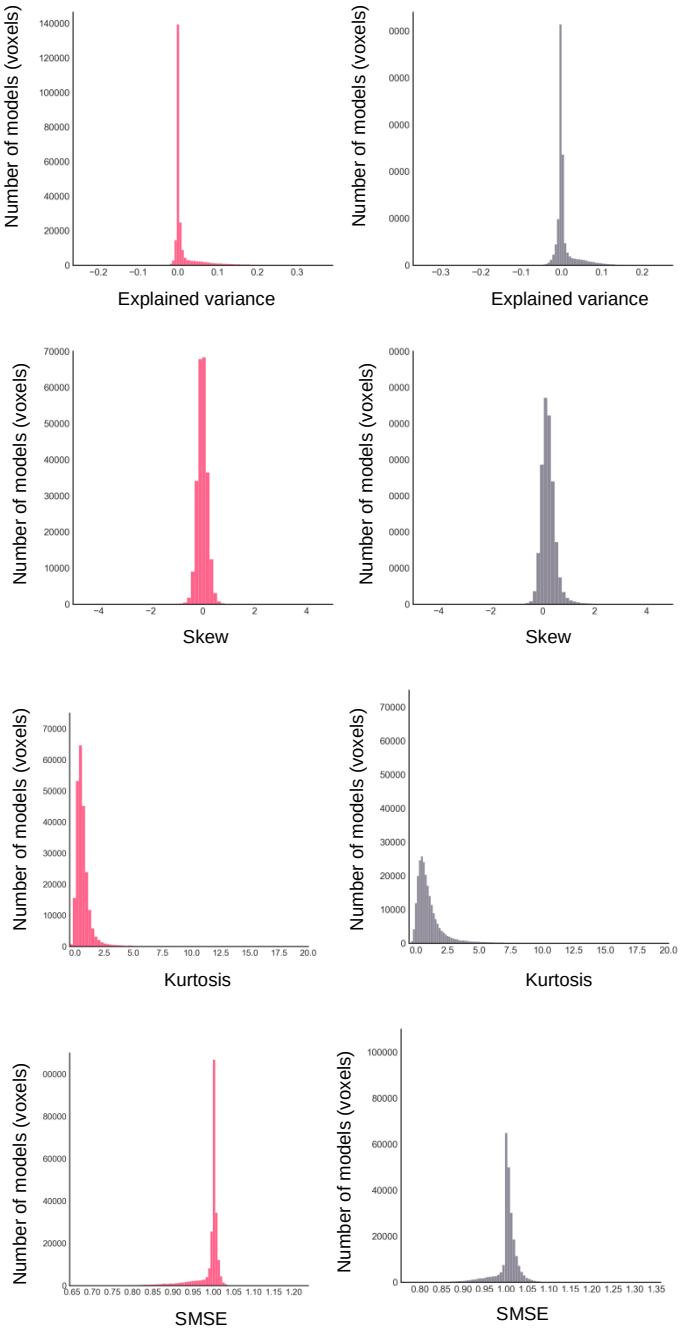
Supplementary Figures



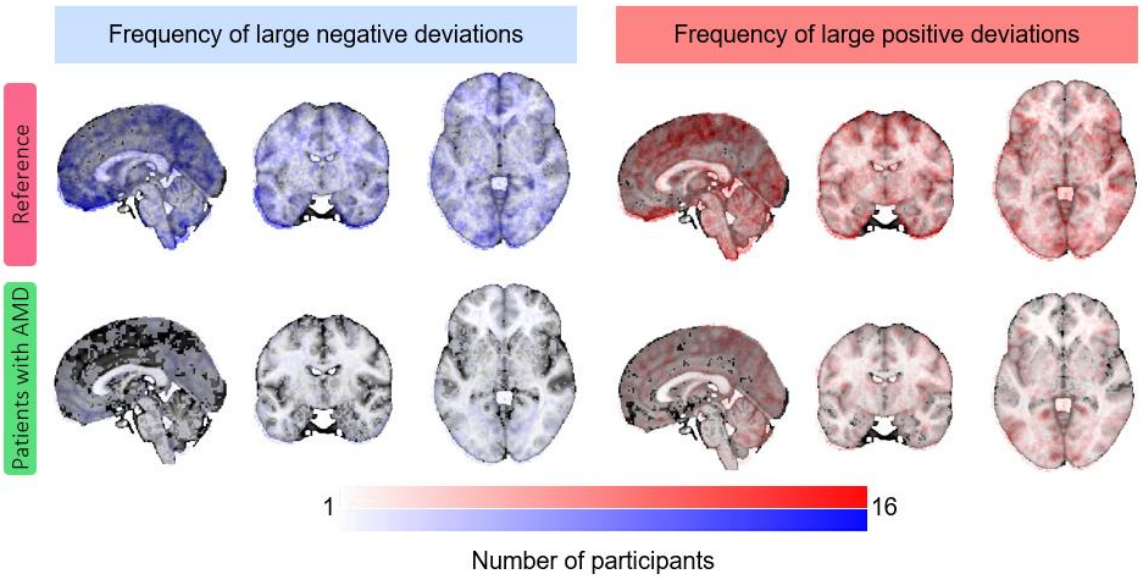
Supplementary Figure S1. CS+ versus CS− contrast (*n*=1888 controls) thresholded to the top 1% of activated voxels (Z > 6.06; warm colours) to highlight specific subcortical regions. Masks for the substantia nigra (blue), red nucleus (red), and periaqueductal grey area (green) are shown to aid in the localization of effects.



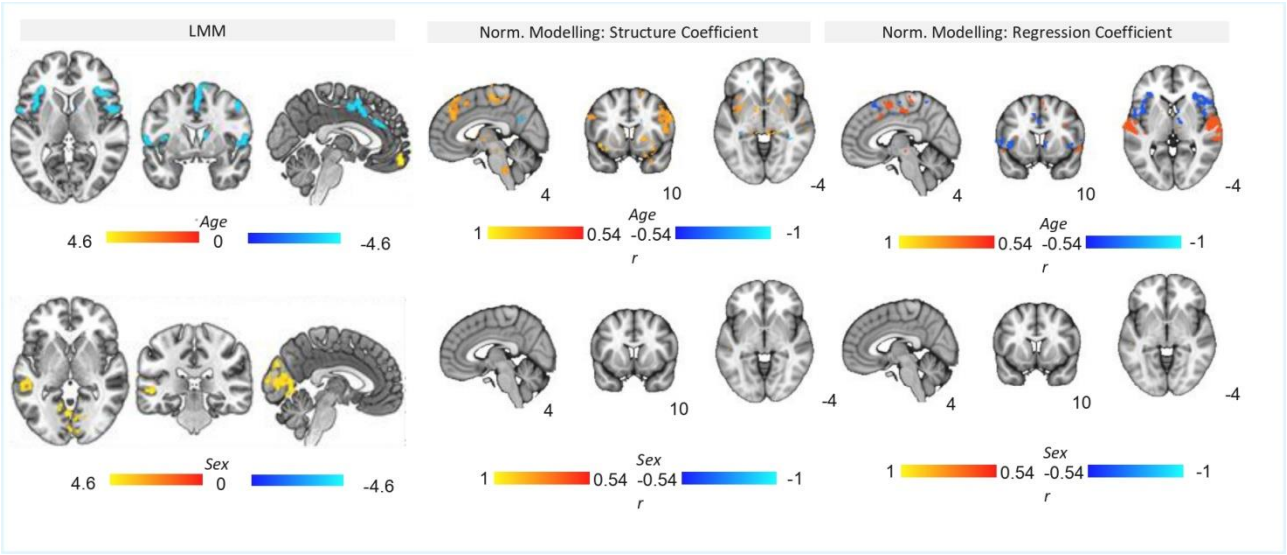
Supplementary Figure S2. Significant brain activation (warm colours) and deactivation (cool colours) to the CS+ versus CS- across axial (**a**; Z = -68 to 106), sagittal (**b**; X = -86 to 88) and coronal (**c**; Y = -120 to 86) slices (*n*=1888 controls).



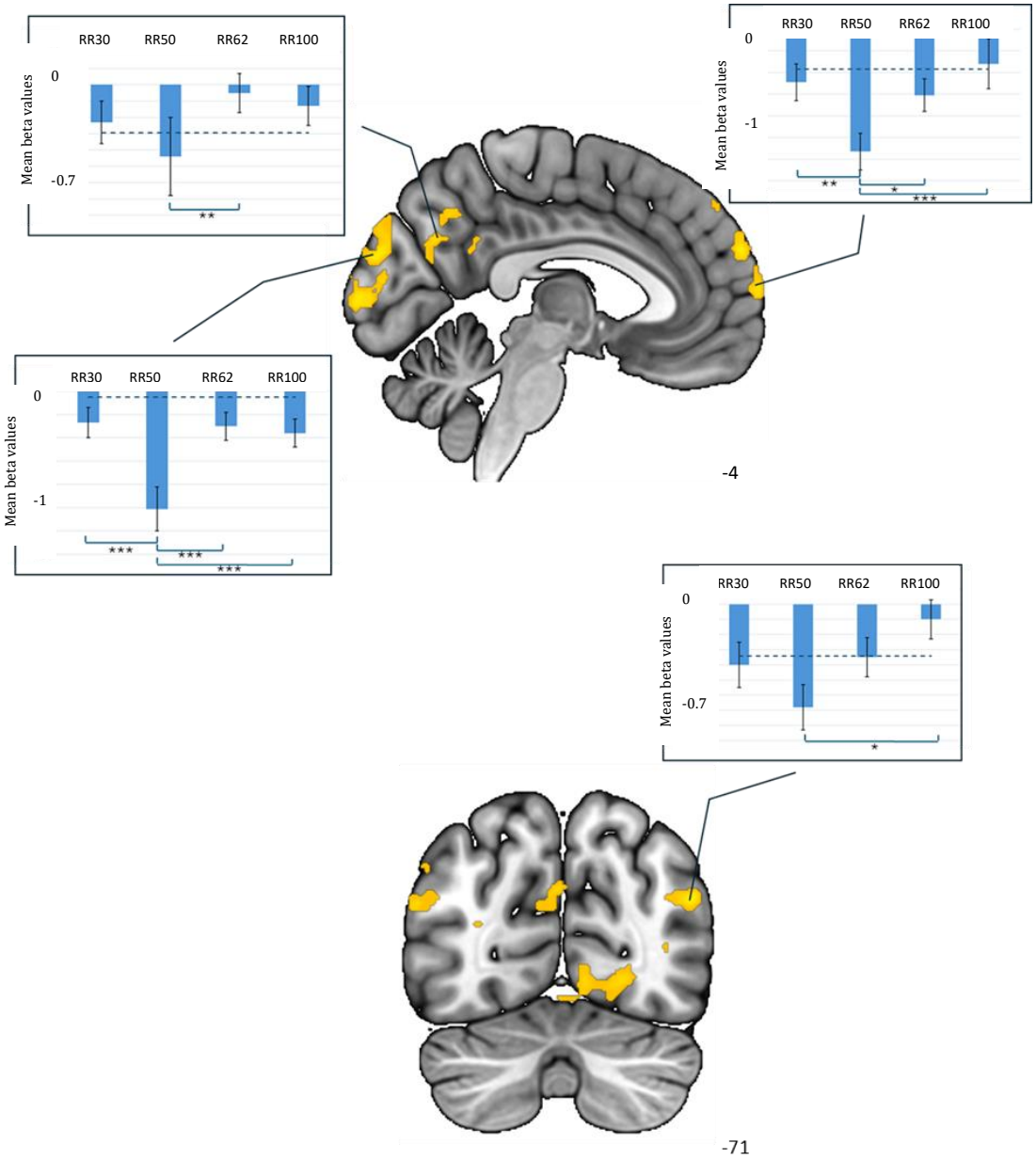
Supplementary Figure S3. Evaluation metrics of normative models**.** Explained variance (a), skew (b), kurtosis (c), and Standardized Mean Squared Error (SMSE) (d) for control test (*n* = 646 controls - left, pink) and clinical test (*n* = 260 controls + 222 individuals with anxiety-related or depressive disorders – right, grey).



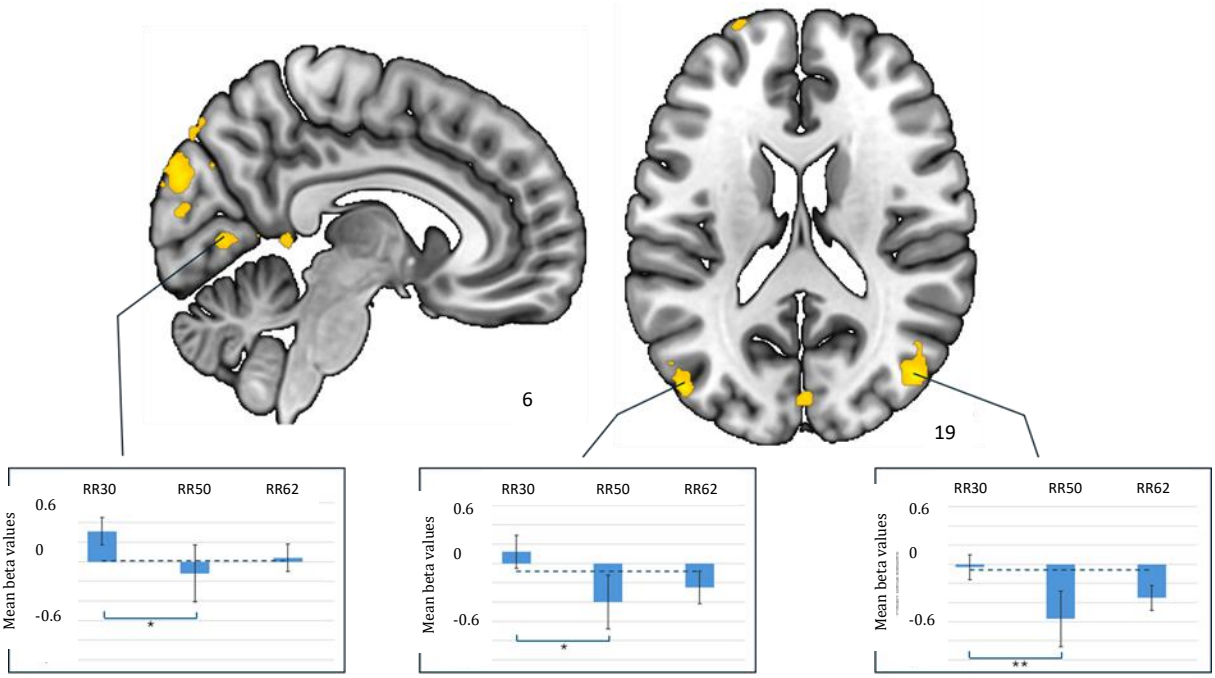
Supplementary Figure S4. Normative probability maps illustrate the number of participants in the sample (test controls - top; individuals with anxiety-related or depressive disorders (AMD) - bottom) who had positive (hot colours - right) or negative deviations (cool colours - left) >±2.6 within each voxel.



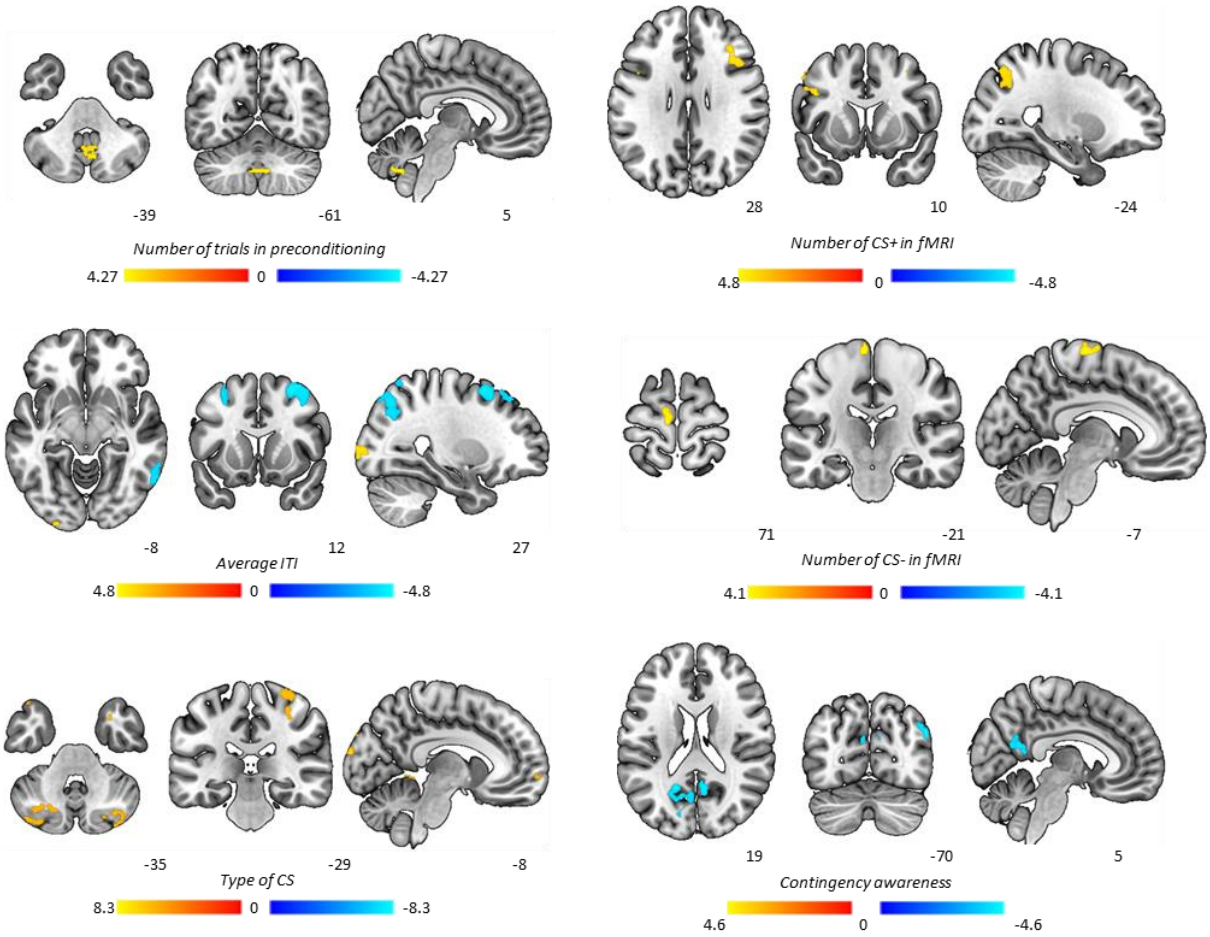
Supplementary Figure S5. Association of age and sex with brain (de)activation during fear conditioning (*n*=1884 for age, *n*=1888 for sex for linear models; *n*=646 for age and sex for normative models). Results from linear mixed-effect models and normative modelling. For the mega-analysis, warm colours indicate positive correlations (i.e., higher variable values associated with greater activation), while cool colours indicate negative correlations (i.e., higher variable values associated with reduced activation). For normative modelling, maps show the regression coefficient or structure coefficients (rho) from normative models for each task variable, thresholded by their respective coefficients of determination (rho^2 > 0.3). Positive correlations (warm colours) indicate greater activation for higher values of the input variable and negative correlations (cool colours) greater activation for lower values of the input variable.



Supplementary Figure S6. Differences in brain activation (warm colours) between different pairing rates (including participants with potential US confounding effect). RR30 (*n*=268); RR50 (*n*=501); RR62 (*n*=333); RR100 (*n*=371). Results of pairwise comparisons after significant ANOVAs. Asterisks indicate significant differences between groups with Bonferroni correction ($p < 0.05$, *$p < 0.01$, **$p < 0.001$). Dashed blue lines indicate mean brain activation for healthy controls. Error bars represent standard errors. RR=pairing rate

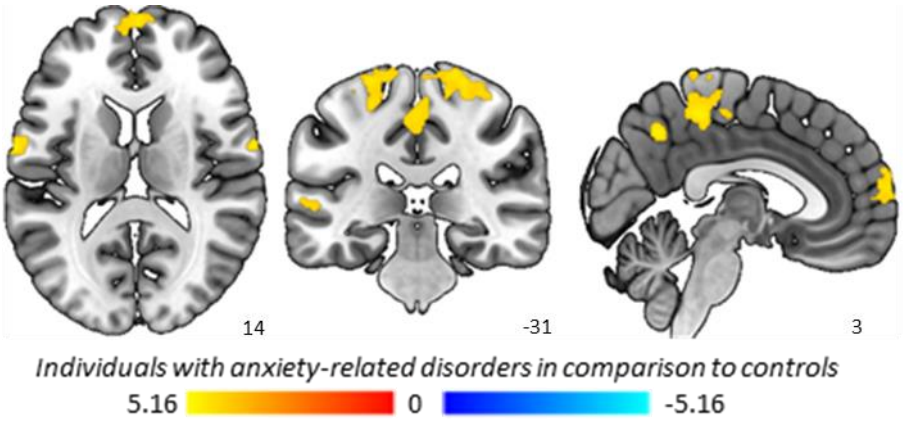


Supplementary Figure S7. Differences in brain activation (warm colours) between different pairing rates (excluding participants with potential US confounding effect). RR30 (*n*=268); RR50 (*n*=139); RR62 (*n*=238). Results of pairwise comparisons after significant ANOVAs. Asterisks indicate significant differences between groups with Bonferroni correction ($p < 0.05$, *$p < 0.01$, **$p < 0.001$). Dashed blue lines indicate mean brain activation for healthy controls. Error bars represent standard errors. RR=pairing rate

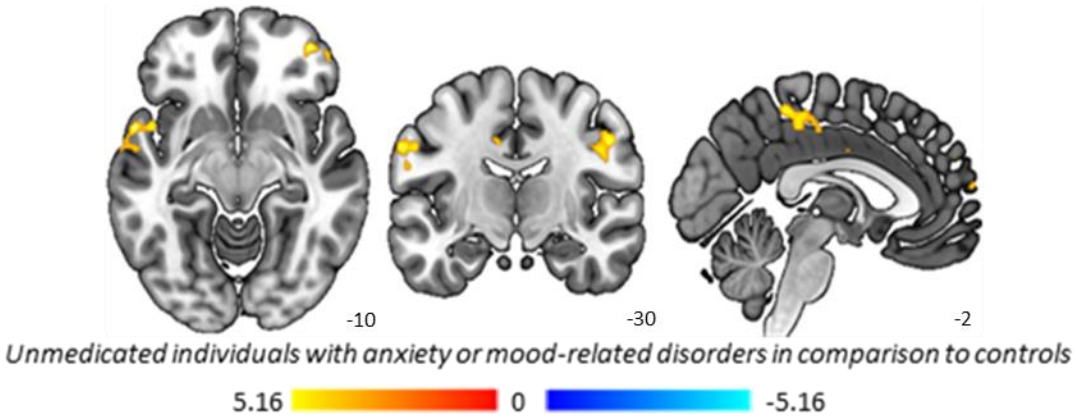


Supplementary Figure S8. Influence of task variables on brain activation (warm colours) and deactivation (cool colours) during fear conditioning. Results from linear mixed-effect models for task variables not presented in the main text (*n* number trials in preconditioning=1254; *n* average ITI=1838; *n* type of CS=1828; *n* number of CS+ in fMRI=1888; n number of CS- in fMRI=1888; *n* contingency awareness=1332). CS+=Conditioned Stimulus followed by the Unconditioned Stimulus. ITI= Intertrial Interval. Number of CS+ in fMRI=Number of CS+ included in fMRI contrast. For type of CS, the figure shows significant results in the ANOVA comparing three categories (humanoid, affective pictures, and neutral faces).

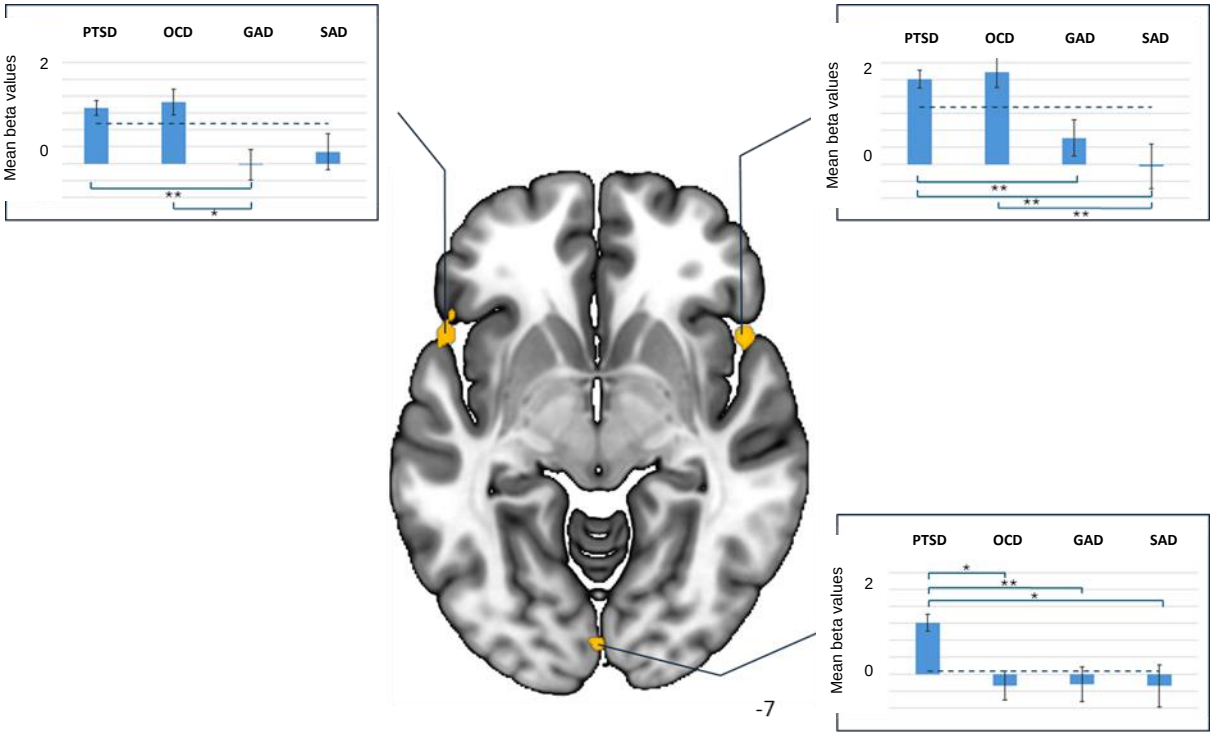
Supplementary Figure S9. Influence of task variables on brain activation during fear conditioning. Results from normative models (*n*=646). Maps show the regression coefficient or structure coefficients (rho) from normative models for each task variable, thresholded by their respective coefficients of determination (rho^2 > 0.3). Positive correlations (warm colours) indicate greater activation for higher values of the input variable and negative correlations (cool colours) greater activation for lower values of the input variable (note that some variables are dummy coded, e.g., instructions, type of US stimuli). CS=Conditioned Stimulus; US=Unconditioned Stimulus. Any task-related variable maps not shown in the main text or in this table did not contain any voxels exceeding the threshold (i.e., they were empty maps).



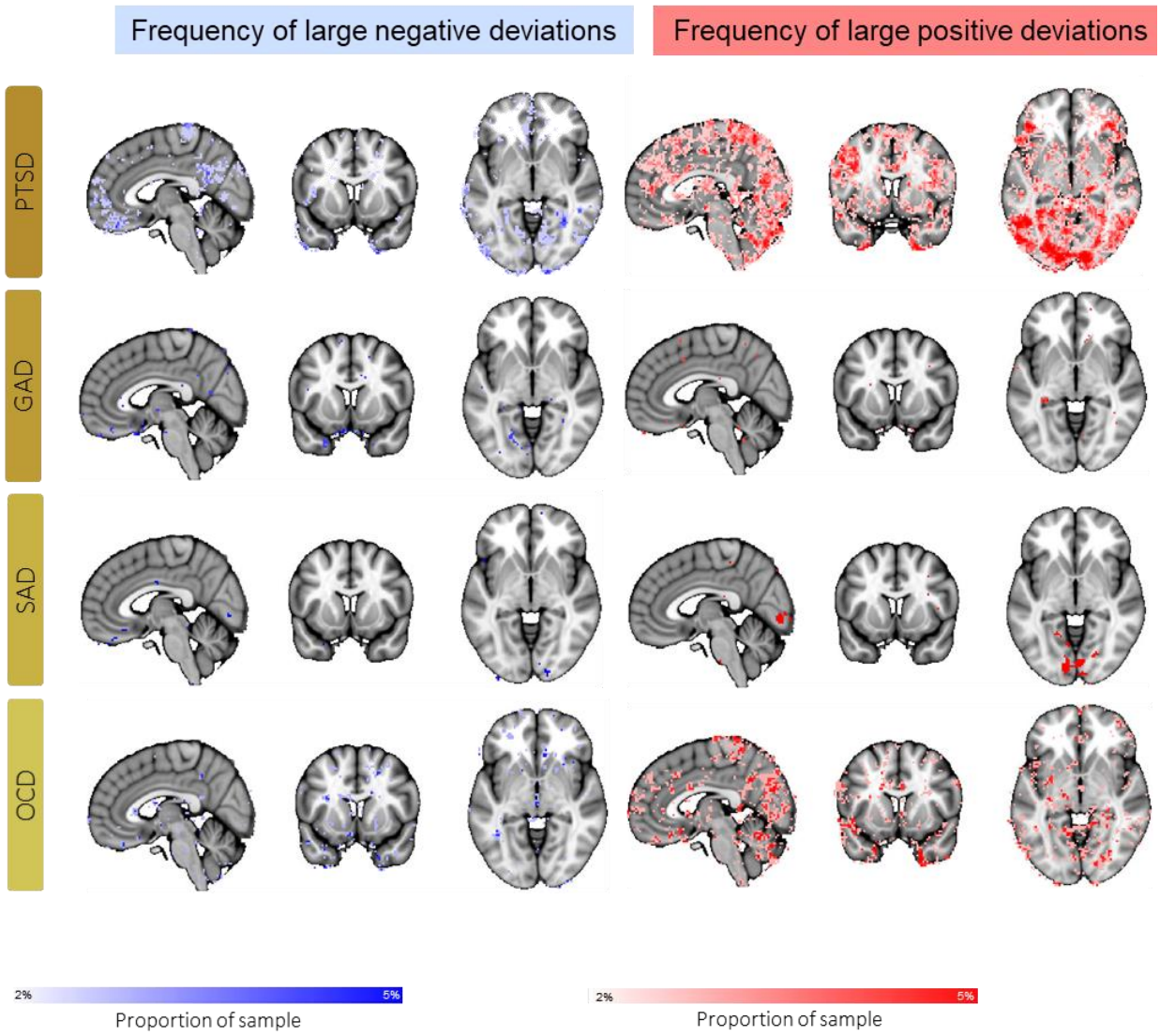
Supplementary Figure S10. Differences in brain activation (warm colours) and deactivation (cool colours) between individuals with anxiety-related disorders (*n*=297) and healthy controls (*n*=1888).



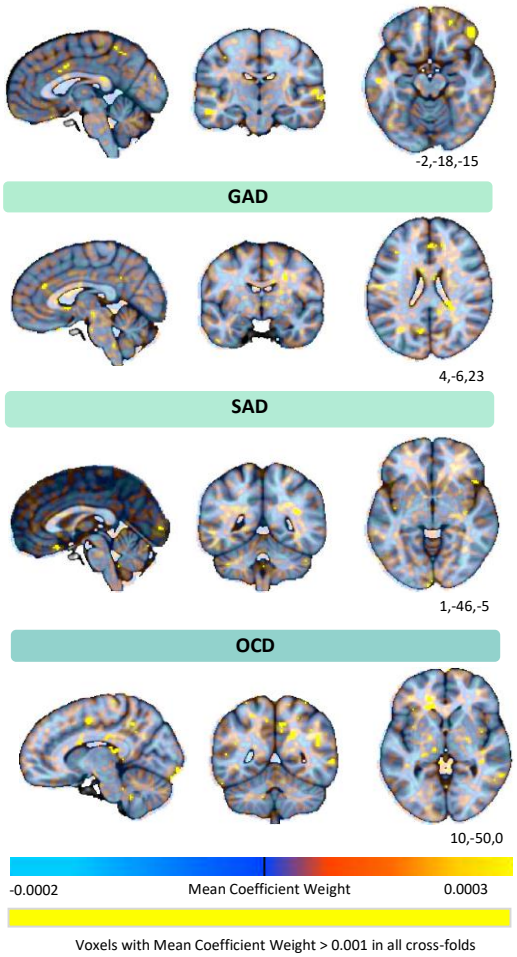
Supplementary Figure S11. Differences in brain activation (warm colours) and deactivation (cool colours) between unmedicated individuals with anxiety or mood-related disorders (*n*=221) and healthy controls (*n*=1859).



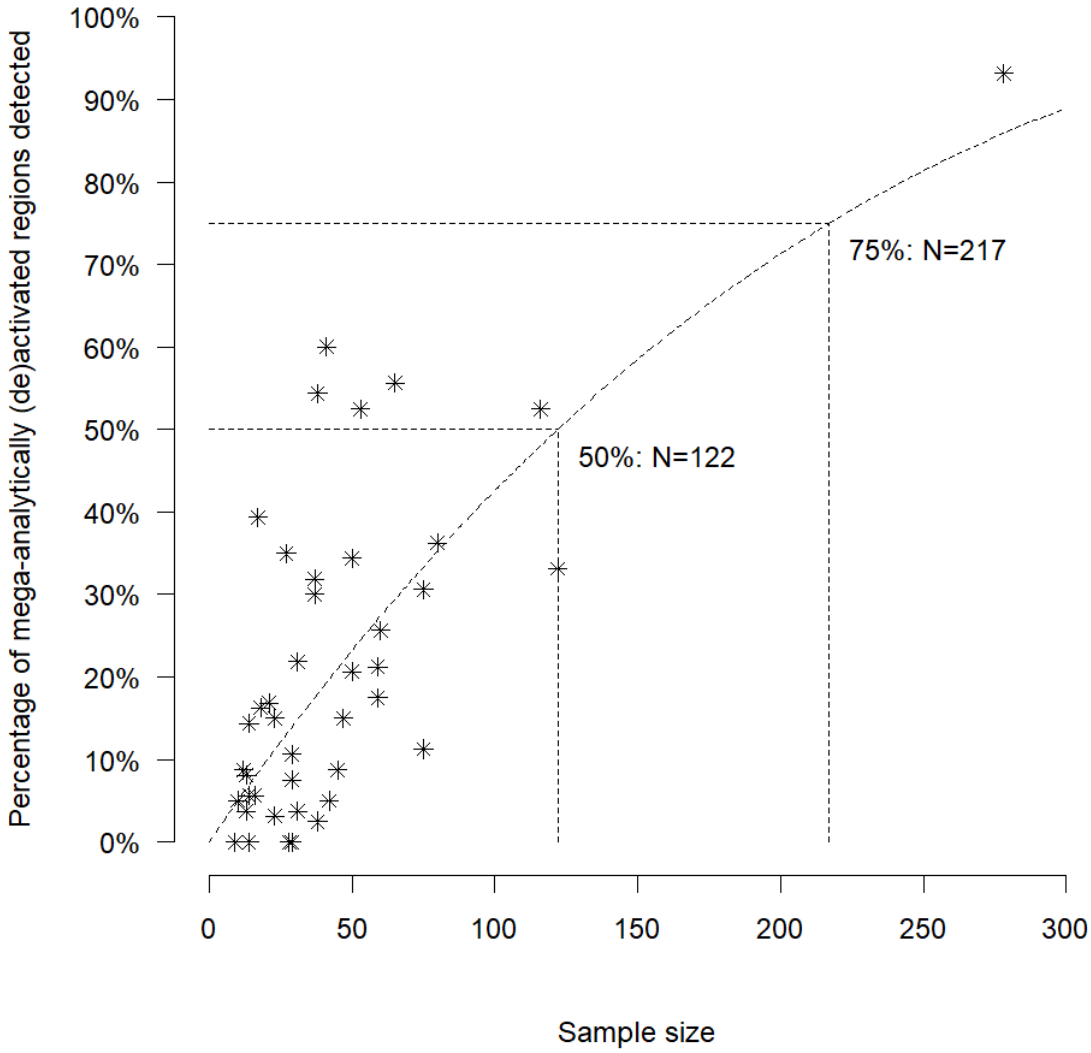
Supplementary Figure S12 Differences in brain activation (warm colours) between patient groups. PTSD=post-traumatic stress disorder (*n*=141); OCD=obsessive-compulsive disorder (*n*=68); GAD=generalized anxiety disorder (*n*=48); SAD=social anxiety disorder (*n*=31). Results of pairwise comparisons after significant ANOVAs. Asterisks indicate significant differences between groups with Bonferroni correction ($p < 0.05$, *$p < 0.01$, **$p < 0.001$). Dashed blue lines indicate mean brain activation for healthy controls. Dashed blue lines indicate mean brain activation for healthy controls. Error bars represent standard errors.



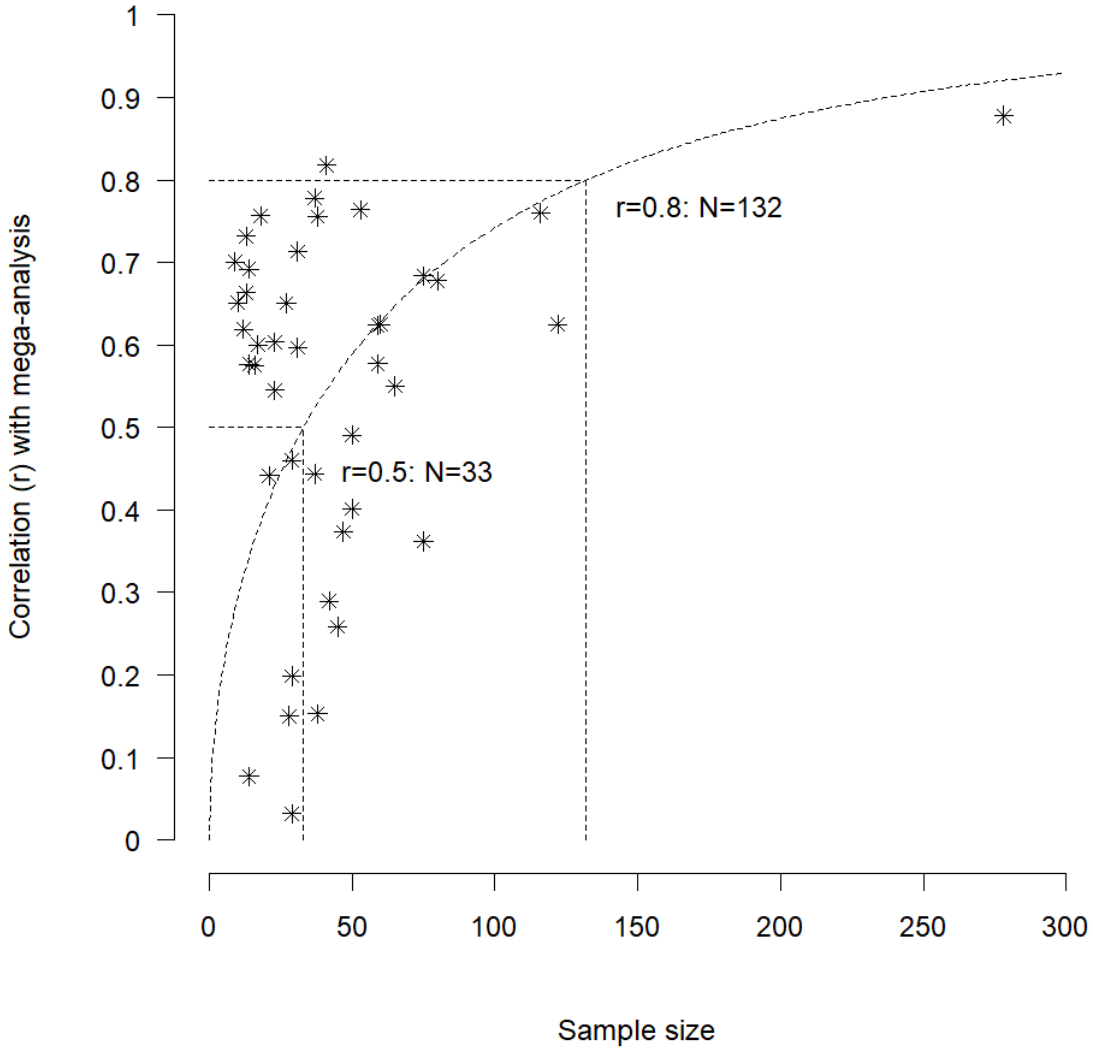
Supplementary Figure S13. Expanded version of Figure 3d from the main manuscript with enhanced visualization for improved clarity and detail. Normative probability maps illustrate the percentage of each clinical group who had positive (hot colours - right) or negative deviations (cool colours - left) >±2.6 within each voxel.



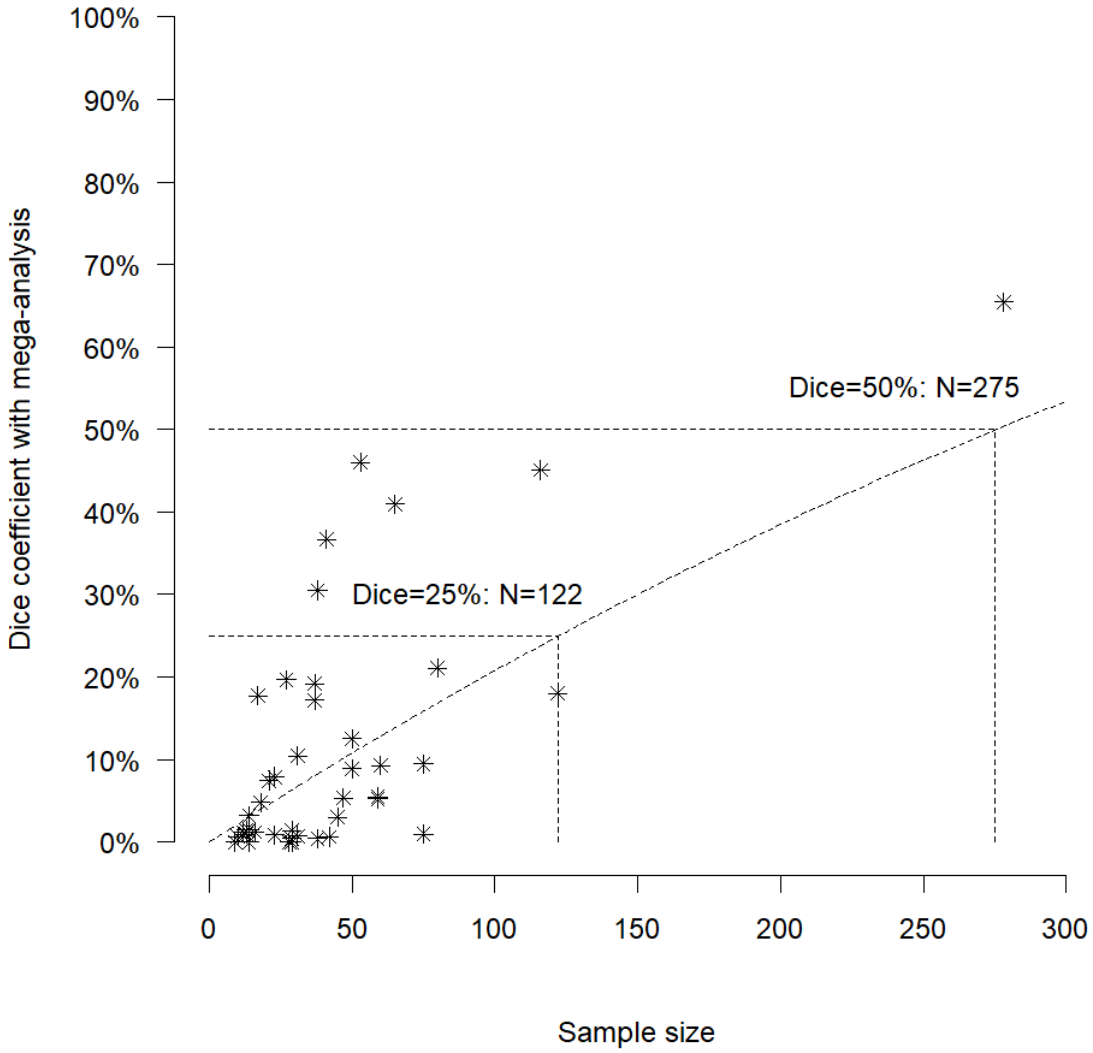
Supplementary Figure S14. Mean coefficient weights from multi-class support vector classifier, used to differentiate whole-brain unthresholded deviation maps between patient groups (PTSD *n* = 55; OCD *n* = 68; GAD *n* = 48; SAD *n* = 31). Yellow indicates voxels that had a mean coefficient weight > 0.001 in all cross-folds (i.e. were frequently used to inform classification).



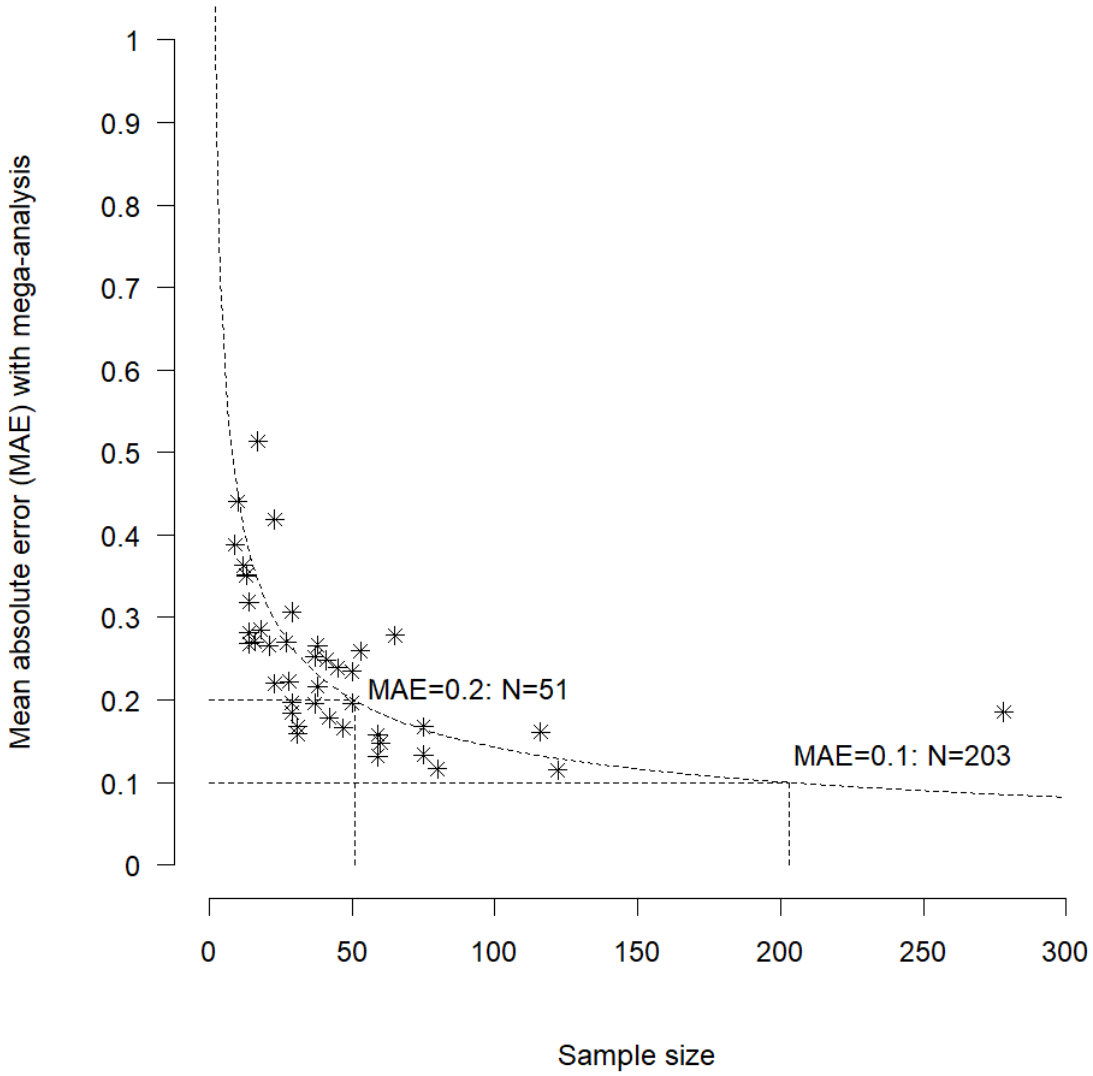
Supplementary Figure S15. Sample size analyses (*n* datasets=42). Percentage of (de)activated brain regions detected in the mega-analysis according to sample size.



Supplementary Figure S16. (Fisher-transformed) correlation between each study (*n* datasets=42) and the mega-analysis. Sample sizes of 33 and 132 were required to achieve correlations of 0.5 and 0.8, respectively.

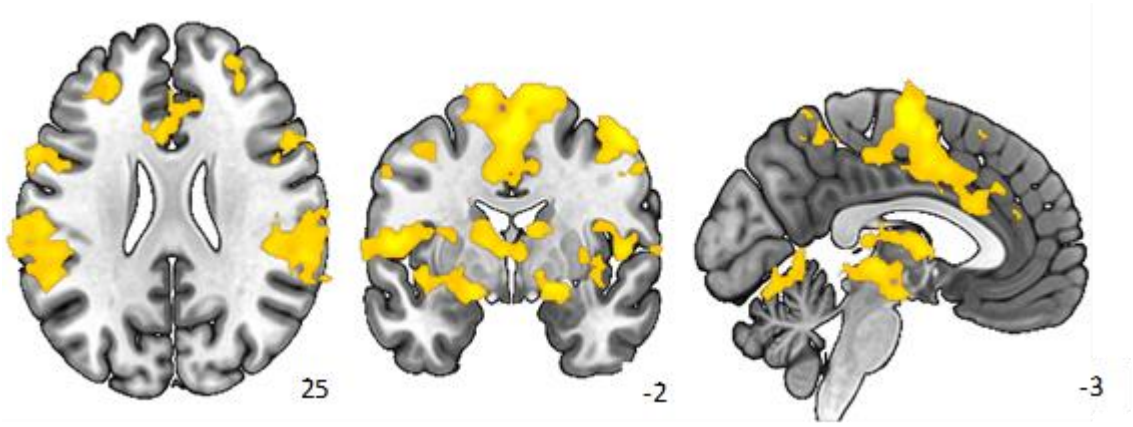


Supplementary Figure S17. (Arcsine-transformed) Dice coefficient between each study (*n* datasets=42) and the mega-analysis. Sample sizes of 122 and 275 were required to attain Dice coefficients of 25% and 50%, respectively.

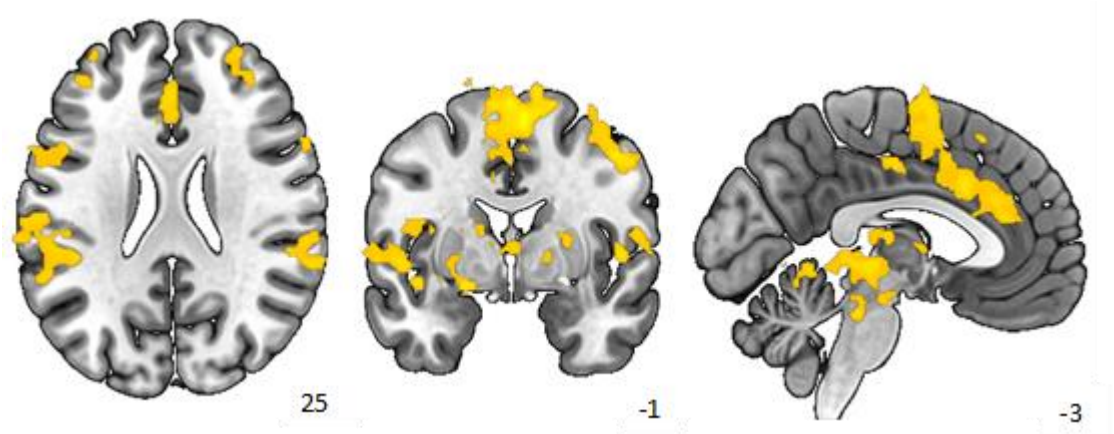


Supplementary Figure S18. Sample size analyses. Mean absolute error (MAE) between each study (n datasets=42) and the mega-analysis. Sample sizes of 51 and 203 were required to achieve MAE values of 0.2 and 0.1, respectively.

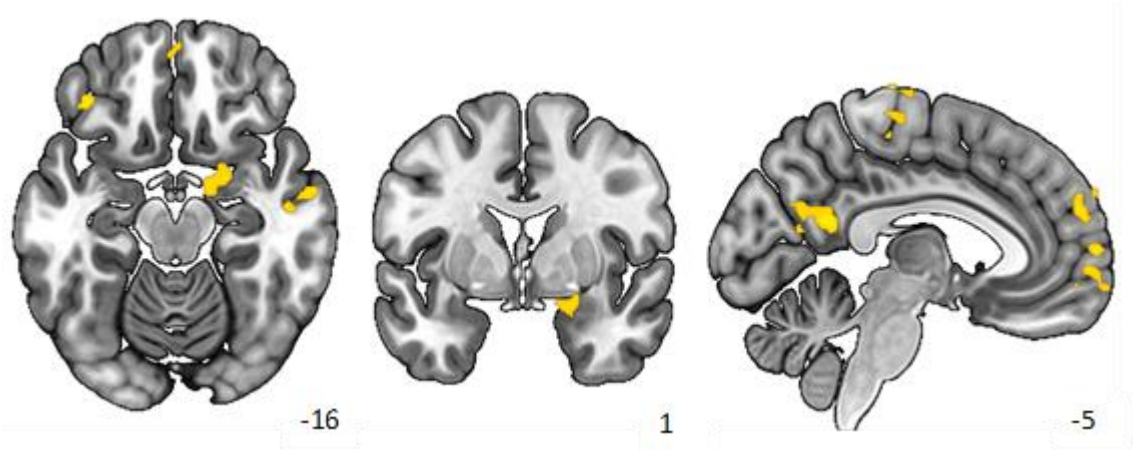
Early conditioning



Late conditioning



Early versus late conditioning



Supplementary Figure S19. Significant brain activation (warm colours) in response to CS+ versus CS− during early (*n*=679) and late (n=634) phases of conditioning, and brain regions showing significant differential activation (warm colours) between early and late conditioning in healthy controls (*n* = 634). Samples included in the analysis are reported in Supplementary Table S2.

Supplementary Tables

Supplementary Table S1. Descriptive statistics for STAI-T and BDI across samples.

Sample	STAI-T (*n*)	STAI-T M (SD)	STAI-T range	BDI (*n*)	BDI M (SD)	BDI range
Amsterdam_Visser_sample_1	18	35.33 (10.39)	22 - 59	NA	NA	NA
Amsterdam_Visser_sample_2	41	34.66 (8.84)	22 - 53	NA	NA	NA
Amsterdam_Visser_sample_3	12	32.67 (5.82)	23 - 44	NA	NA	NA
Amsterdam_Visser_sample_4	10	35.3 (5.38)	29 - 46	NA	NA	NA
Amsterdam_Visser_sample_5	13	37.46 (9.47)	26 - 60	NA	NA	NA
Amsterdam_Visser_sample_6	14	35.29 (9.71)	21 - 58	NA	NA	NA
Amsterdam_Visser_sample_7	16	33.5 (6.04)	25 - 46	NA	NA	NA
Amsterdam_Visser_sample_8	9	36.44 (8.14)	27 - 52	NA	NA	NA
Amsterdam_Visser_sample_9	38	35.03 (8.63)	20 - 52	NA	NA	NA
Austin_Cisler	NA	NA	NA	61	22.57 (12.51)	0 - 55
Barcelona_Cardoner*	71	25.49 (13.49)	1 - 53	71	14 (11.87)	0 - 46
Barcelona_Soriano_sample_2*	147	20.47 (10.73)	1 - 52	NA	NA	NA
Bielefeld_Lonsdorf_sample_1	116	34.86 (7.36)	24 - 55	NA	NA	NA
Bielefeld_Lonsdorf_sample_2	80	35.37 (10)	20 - 59	NA	NA	NA
Bielefeld_Lonsdorf_sample_3	28	35.93 (6.96)	24 - 52	NA	NA	NA
Bochum_Elsenbruch	29	33.03 (6.51)	21 - 44	NA	NA	NA
Bochum_Merz_sample_5	31	33.32 (6.82)	20 - 52	NA	NA	NA
Bochum_Merz_sample_6	60	36.2 (6.88)	23 - 52	NA	NA	NA
Duke_LaBar_sample_1	38	32.39 (7.86)	21 - 53	NA	NA	NA
Duke_LaBar_sample_2	37	33.28 (6.55)	20 - 48	NA	NA	NA
Manitoba_Greening_sample_1	13	38.92 (9.3)	29 - 59	NA	NA	NA
Manitoba_Greening_sample_2	31	35.27 (10.45)	21 - 57	NA	NA	NA
Melbourne_Harrison	112	38.97 (13.05)	21 - 73	NA	NA	NA
Munster_Moeck_sample_1	42	34.19 (7.3)	22 - 50	42	3.62 (4.36)	0 - 16
Reading_Reekum_sample_1	21	41.62 (8.66)	27 - 59	NA	NA	NA
Reading_Reekum_sample_2	50	42.92 (9.82)	26 - 75	NA	NA	NA
Royal_Tuominen_sample_1	28	35.57 (13.83)	20 - 67	28	5.68 (7.98)	0 - 27
Royal_Tuominen_sample_2	71	34.97 (10.33)	20 - 68	71	5.15 (6.48)	0 - 23
USP_Diniz	NA	NA	NA	25	20.4 (11.47)	0 - 44
Texas_Dunsmoor	NA	NA	NA	45	15.68 (10.89)	0 - 41

Ulm_Abler	50	33.38 (6.13)	23 - 52	NA	NA	NA
Uppshala_Ahs	278	36.27 (11.44)	20 - 67	NA	NA	NA
Vanderbilt_Kaczkurkin	82	43.38 (12.14)	21 - 70	82	12.38 (8.62)	0 - 31
TOTAL	**1586**	**34.45 (11.56)**	**1 - 75**	**425**	**12.41 (11.48)**	**0 - 55**

BDI: Beck Depression Inventory; NA: Not available: STAI-T: State Trait Anxiety Inventory-Trait version. *These samples used the Spanish version of the STAI-T (scores range from 0 to 60)

Supplementary Table S2. Samples included in the early versus late analyses

Sample	*n*	CS+ trials included in early/late analysis (n/n)	CS- trials included in early/late analysis (n/n)	Pairing rate (%)	US confound in early/late analysis
Barcelona_Cardoner*	45	8/-	16/-	50	no
Barcelona_Soriano_sample_2	122	5/5	5/5	33	no
Bochum_Elsenbruch	29	4/4	4/4	100	yes
Bochum_Merz_sample_1	56	8/8	4/4	62.5	yes
Bochum_Merz_sample_2	58	8/8	4/4	62.5	yes
Bochum_Merz_sample_3	47	10/10	10/10	100	yes
Bochum_Merz_sample_4	28	8/8	4/4	62.5	yes
Bochum_Merz_sample_5	31	8/8	4/4	62.5	yes
Bochum_Merz_sample_6	60	8/8	4/4	62.5	yes
Duke_LaBar_sample_1	38	5/5	5/5	50	yes
Duke_LaBar_sample_2	37	8/8	8/8	31	yes
Harvard_McLaughlin	75	4/4	2/2	40	yes
Vanderbilt_Kaczkurkin	53	7/7	15/15	80	yes

CS, conditioned stimulus; CS+, CS followed by unconditioned stimulus; CS −, CS not followed by unconditioned stimulus; CS+/CS-, US=Unconditioned stimulus.
All samples used visual conditioned stimuli. All samples used an electric shock as US except Barcelona_Cardoner and Harvard_McLaughlin, which used an auditory stimulus.
*Only early trials were available.

Supplementary Table S3. Patient's medications.

Sample	Medicated (*n*)	SSRI or SNRI (*n*)	BZD (*n*)	Other* (*n*)
Austin_Cisler	36	2	0	34
Barcelona_Cardoner	1	0	1	0
Barcelona_Soriano_sample_1	16	10	0	6
Munich_Koch	12	7	0	5
Vanderbilt_Kaczkurkin	1	1	0	0
TOTAL	66	20	1	45

SSRI: Selective Serotonin Reuptake Inhibitors; SNRI: Selective Noradrenaline Reuptake Inhibitors; BZD: Benzodiazepines. *Includes other medications or combinations of medications.

Supplementary Table S4. Participants excluded after quality control (QC)

Sample	N collected	N excluded after HALFpipe QC	N excluded after manual QC	N included in analysis
Amsterdam_Visser_sample_1	19	0	1	18
Amsterdam_Visser_sample_2	41	0	0	41
Amsterdam_Visser_sample_3	12	0	0	12
Amsterdam_Visser_sample_4	11	1	0	10
Amsterdam_Visser_sample_5	13	0	0	13
Amsterdam_Visser_sample_6	14	0	0	14
Amsterdam_Visser_sample_7	16	0	0	16
Amsterdam_Visser_sample_8	10	1	0	9
Amsterdam_Visser_sample_9	38	0	0	38
Austin_Cisler	88	27	0	61
Barcelona_Cardoner	90	16	3	71
Barcelona_Soriano_sample_1	37	2	0	35
Barcelona_Soriano_sample_2	191	44	0	147
Bielefeld_Lonsdorf_sample_1	120	4	0	116
Bielefeld_Lonsdorf_sample_2	83	1	2	80
Bielefeld_Lonsdorf_sample_3	32	4	0	28
Bochum_Elsenbruch	30	1	0	29
Bochum_Merz_sample_1	60	1	0	59
Bochum_Merz_sample_2	60	1	0	59
Bochum_Merz_sample_3	48	1	0	47
Bochum_Merz_sample_4	33	4	0	29
Bochum_Merz_sample_5	32	1	0	31
Bochum_Merz_sample_6	64	4	0	60
Columbia_Neria	114	15	4	95
Duke_LaBar_sample_1	40	2	0	38
Duke_LaBar_sample_2	40	3	0	37
Florida_Keil	15	0	1	14
Harvard_McLaughlin	95	6	0	89
Manitoba_Greening_sample_1	13	0	0	13
Manitoba_Greening_sample_2	31	0	0	31
Melbourne_Harrison	154	40	2	112
Munich_Koch	52	4	3	45
Munster_Moeck_sample_1	44	2	0	42
Munster_Moeck_sample_2	31	2	0	29
Reading_Reekum_sample_1	22	1	0	21
Reading_Reekum_sample_2	52	2	0	50
Royal_Tuominen_sample_1	17	0	3	14
Royal_Tuominen_sample_2	37	0	0	37
Texas_Dunsmoor	48	3	0	45

Ulm_Abler	51	1	0	50
Uppsala_Ahs	306	28	0	278
USP_Diniz	56	1	0	55
Vanderbilt_Kaczkurkin	88	6	1	81
TOTAL	2448	229	20	2199

Supplementary Table S5. Pairwise Pearson correlation between sociodemographic, individual differences and task variables.

	Age	Sex	Anxiety score	Depression score	Instructions given about	Number of trials in	Paradigm with multiple CS	Paradigm with multiple CS	Type of CS	Number of CS plus during	Number of CS minus during	Average ITI	Average ISI	Pairing rate	Type of US used	Potential US confounding	Number of CS plus included	Number of CS minus included	Contingency awareness
Sex	0.03																		
Anxiety score	0.11	0.02																	
Depression score	0.12	0.19	0.73																
Instructions given about contingency prior to the task	0.06	0.08	0	0.33															
Number of trials in preconditioning	0.07	0.24	0.16	0.07	0.2														
Paradigm with multiple CS plus	0.03	0.12	0	0	0.23	0.16													
Paradigm with multiple CS minus	0.02	0.07	0.05		0.17	0.16	0.37												
Type of CS	0.33	0.05	0.14	0.41	0.3	0.31	0.27	0.16											
Number of CS plus during conditioning	0.06	0.09	0.02	0.11	0.26	0.18	0.26	0.03	0.07										
Number of CS minus during conditioning	0.23	0.02	0.05	0.1	0.09	0.58	0.08	0.38	0.08	0.42									
Average ITI	0.2	0.04	0.12	0.25	0.12	0.49	0.02	0.06	0.21	0.43	0.59								
Average ISI	0.26	0.04	0.14	0.14	0.11	0.07	0.16	0	0.26	0.2	0.15	0.13							
Pairing rate	0.34	0.17	0.13	0.44	0.71	0.26	0.68	0.18	0.46	0.36	0.53	0.3	0.62						
Type of US used	0.43	0.06	0.17	0.12	0.08	0.02	0.33	0.22	0.23	0.04	0.1	0.06	0.66	0.45					
Potential US confounding	0.41	0.04	0.12	0.15	0.13	0.09	0.18	0.03	0.39	0.25	0.18	0.05	0.44	0.6	0.38				
Number of CS plus included in the fMRI contrast	0.29	0.03	0.06	0.03	0.05	0.09	0.11	0.07	0.33	0.39	0.54	0.28	0.03	0.42	0.24	0.52			
Number of CS minus included in the fMRI contrast	0.22	0.03	0.03	0.07	0.02	0.43	0.2	0.17	0.11	0.4	0.79	0.4	0.11	0.36	0.17	0.17	0.7		
Contingency awareness	0.1	0.01	0.04	0.06	0.01	0.04	0.01	0.1	0.23	0.06	0.09	0.07	0.21	0.19	0.18	0.03	0.14	0.09	
Diagnosis of anxiety related disorder (PTSD,GAD,...)	0.49	0.06	0.38	0.16	0.31	0.4	0.58	0.35		0.6	0.62	0.87	0.66	0.78	0.78	0.72	0.42	0.65	0.18
Patient versus controls	0.16	0.05	0.47	0.65	0.07	0.15	0.06	0.06	0.2	0.04	0.12	0.25	0.16	0.23	0.05	0.04	0.01	0.09	0.16

Supplementary Table S6. Pairwise variance inflation factor (VIF) between sociodemographic, individual differences and task variables*

	Age	Sex	Anxiety score	Depression score	Instructions given about	Number of trials in	Paradigm with multiple CS plus	Paradigm with multiple CS	Type of CS	Number of CS plus during	Number of CS minus during	Average ITI	Average ISI	Pairing rate	Type of US used	Potential US confounding	Number of CS plus included in	Number of CS minus included	Contingency awareness
Sex	1.0																		
Anxiety score	1.0	1.0																	
Depression score	1.0	1.0	1.6																
Instructions given about contingency prior to the task	1.0	1.0	1.0	1.0															
Number of trials in preconditioning	1.0	1.0	1.0	1.1	1.1														
Paradigm with multiple CS plus	1.0	1.0	1.0	1.0	1.0	1.0													
Paradigm with multiple CS minus	1.0	1.0	1.0		1.0	1.0	1.2												
Type of CS	1.2	1.0	1.0	1.2	1.1	1.1	1.1	1.0											
Number of CS plus during conditioning	1.0	1.0	1.0	1.0	1.1	1.0	1.1	1.0	1.0										
Number of CS minus during conditioning	1.0	1.0	1.0	1.1	1.0	1.4	1.0	1.2	1.0	1.2									
Average ITI	1.0	1.0	1.0	1.2	1.0	1.3	1.0	1.0	1.1	1.2	1.3								
Average ISI	1.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.1	1.0	1.0							
Pairing rate	1.1	1.0	1.0	1.4	2.2	1.0	1.9	1.0	2.1	1.2	1.4	1.1	1.8						
Type of US used	1.2	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.1	1.0	1.0	1.0	2.0	1.2					
Potential US confounding	1.2	1.0	1.0	1.0	1.1	1.0	1.1	1.0	1.2	1.1	1.0	1.0	1.2	1.8	1.1				
Number of CS plus included in the fMRI contrast	1.1	1.0	1.0	1.1	1.0	1.0	1.1	1.0	1.2	1.2	1.4	1.1	1.0	1.4	1.0	1.3			
Number of CS minus included in the fMRI contrast	1.0	1.0	1.0	1.1	1.0	1.1	1.0	1.0	1.0	1.2	2.5	1.1	1.0	1.2	1.0	1.0	1.9		
Contingency awareness	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.0	1.1	1.0	1.1	1.0	1.0	1.0	
Diagnosis of anxiety related disorder (PTSD,GAD,...)	1.3	1.0	1.2	1.0	1.1	1.2	1.5	1.1		1.6	1.6	4.0	1.8	**7.8**	2.5	2.1	1.2	1.7	1.0
Patient versus controls	1.0	1.0	1.3	1.7	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.1	1.0	1.0	1.0	1.0	1.0

***** The VIF analysis revealed no collinearity concerns (VIF > 5), except for the pair of variables highlighted in red. However, their activation maps did not show any overlap.

Supplementary Table S7. Task variables showing high (>0.5) inter-correlations.

Variable 1	Variable 2	n	r	VIF	Overlap
Number of CS minus during conditioning	Number of CS minus included in the fMRI contrast	1884	0.774975	2.503673	NO
Instructions given about contingency prior to the task	Pairing rate	1506	0.736404	2.184792	NO
Average ISI	Type of US used	1805	0.698958	1.955196	NO
Number of CS plus included in the fMRI contrast	Number of CS minus included in the fMRI contrast	1884	0.695982	1.939456	NO
Paradigm with multiple CS plus	Pairing rate	1506	0.680728	1.863555	NO
Pairing rate	Potential US confound	1506	0.675691	1.840122	YES*
Average ISI	Pairing rate	1506	0.658202	1.764384	NO
Anxiety score	Depression score	189	0.625581	1.642984	NO
Pairing rate	Number of CS plus included in the fMRI contrast	1506	0.538111	1.407584	NO
Number of CS minus during conditioning	Pairing rate	1506	0.536643	1.404466	NO
Type of CS	Pairing rate	1446	0.533694	2.135761	YES**
Number of CS minus during conditioning	Number of CS plus included in the fMRI contrast	1884	0.525154	1.38081	NO
Number of trials in preconditioning	Number of CS minus during conditioning	1251	0.523602	1.377713	NO
Depression score	Pairing rate	160	0.519464	1.369569	NO

*(see main text)

** Pairing rate and CS type shared two small clusters. In one cluster, both variables remained statistically significant when included in the same model. In the other cluster, located in the left orbitofrontal cortex, only the pairing rate remained significant. With no collinearity concerns (VIF = 2.1), this suggests that activation in this region is modulated by the pairing rate rather than CS type.

Supplementary Table S8. Task and clinical variables showing high (>0.5) inter-correlations.

Variable 1	Variable 2	n	r	VIF	Overlap
Average ITI	Diagnosis of anxiety related disorder	288	0.866338	4.008687	NO
Pairing rate	Diagnosis of anxiety related disorder	228	0.784478	7.796471	NO
Type of US used	Diagnosis of anxiety related disorder	288	0.775372	2.507536	NO
Potential US confound	Diagnosis of anxiety related disorder	288	0.717325	2.059966	NO
Average ISI	Diagnosis of anxiety related disorder	288	0.656201	1.756235	NO
Depression score	Patient versus controls	375	0.652073	1.739734	NO
Number of CS minus included in the fMRI contrast	Diagnosis of anxiety related disorder	288	0.648278	1.724925	NO
Number of CS minus during conditioning	Diagnosis of anxiety related disorder	288	0.622024	1.631092	NO
Number of CS plus during conditioning	Diagnosis of anxiety related disorder	288	0.597379	1.554877	NO
Paradigm with multiple CS plus	Diagnosis of anxiety related disorder	288	0.580814	1.509081	NO

Supplementary Methods

Changes with respect to pre-registration

As noted in the main text, both the mega-analysis (https://osf.io/7n953) and normative modeling analysis (https://osf.io/w74bt) were pre-registered. The following changes were made after pre-registration:

1. At the time of pre-registration, we had collected data from 43 samples. We excluded one sample (*n*=22) because it employed a "multi-CS" conditioning paradigm (36 CS+, 18 CS-) which is difficult to compare with the other experiments included.
2. For the mega-analysis, we used pre-scaling instead of Combat to reduce site-related heterogeneity (see "Pre-scaling" in page 5).

The normative modelling analysis plan was updated to best complement the meta-analysis approach and thus the following changes were made after pre-registration:

1. Sample size. The participants included were a subset of the final sample used in the meta-analysis, for whom all required data were available.
2. Variables included. The variables used were matched to those included in the mega-analysis study to facilitate a better comparison between the results of these complementary methods
3. Analysis plan. Research question 1A. We chose not to create models for separate ROIs. Research question 1C. We did not perform whole-brain sparse canonical correlation analysis to determine how deviations in task activation predicted outcome measures, rather, we chose statistical approaches more appropriate to the type of data. Research question 2B. Again, we did not perform whole-brain sparse canonical correlation analysis, for the same reasons as mentioned above. We did not perform analyses on transdiagnostic scales with insufficient sample sizes (e.g., Beck Anxiety Inventory, Hamilton-Anxiety, Hamilton-Depression) and similarly excluded small diagnostic groups from relevant analyses. We did not use a clustering method.

Variables collected and not included in analyses

The following variables were collected but not included in the analyses because the data collected were insufficient, or too heterogeneous to be aggregated: IQ, comorbidity, ethnicity, years of education, use of a concurrent task during conditioning, and US aversiveness. We excluded the variable "use of preconditioning

phase" from the analyses because we already accounted for it by including "number of trials during preconditioning". Descriptive data on years of education and comorbidity for the samples with available data are reported in Tables 1 and 3 of the main manuscript.

Non-imaging data

Harmonization of trait anxiety scores

As noted in the main text, we conducted the analysis of the State-Trait Anxiety Inventory-Trait version (STAI-T) scores using both raw and harmonized scores.

To harmonize the STAI-T scores, we took the following steps, we first assessed the potential variability of STAI-T scores across versions, languages, or countries, by conducting a meta-analysis of the mean STAI-T scores reported in the normative studies[1-11] as well as a meta-analysis of the reported standard deviations. In both analyses, substantial heterogeneity between studies was observed (I^2 statistic for the mean: 99%; I^2 statistic for the standard deviation: 95%, Q test $p<0.001$ in both cases). This heterogeneity indicates significant differences in the reported means and standard deviations between studies. We then examined potential moderators of this heterogeneity, including the version of the STAI-T (X or Y), language, and country. The results revealed statistically significant differences in the mean and standard deviation across countries ($p=0.014$ and 0.001, respectively) and in the mean across languages ($p=0.012$) but not on the version of the STAI-T.

		Mean		Log SD	
		Estimate (95%CI)	*P*	Estimate (95%CI)	*P*
Version	X	41.2 (36.9-45.4)	n.s.	2.36 (2.31-2.41)	n.s.
	Y	39.2 (36.4-42.0)		2.22 (2.09-2.35)	
Language	Dutch	35.2 (33.0-37.5)	0.012	2.23 (1.97-2.48)	0.353
	English	38.0 (35.7-40.4)		2.17 (2.01-2.32)	
	French	41.9 (40.7-43.1)		2.15 (2.05-2.25)	
	German	43.0 (41.0-44.9)		2.39 (2.36-2.42)	
	Japanese	46.8 (44.6-49.1)		2.43 (2.29-2.57)	
	Spanish	46.2 (37.5-55.0)		2.32 (2.25-2.39)	
Country	America	36.5 (33.9-39.1)	0.014	2.13 (1.88-2.39)	0.001
	Australia	36.4 (35.8-37.0)		2.41 (2.37-2.45)	
	England	41.1 (36.1-46.2)		2.02 (1.79-2.25)	
	France	41.9 (40.7-43.1)		2.15 (2.05-2.25)	
	Germany	43.0 (41.0-44.9)		2.39 (2.36-2.42)	
	Japan	46.8 (44.6-49.1)		2.43 (2.29-2.57)	
	Netherlands	35.2 (33.0-37.5)		2.23 (1.97-2.48)	
	Spain	46.2 (37.5-55.0)		2.32 (2.25-2.39)	

These findings suggest that the observed heterogeneity in STAI-T scores is partly explained by country (or language) differences in the included studies. We could not separate the effects of “country” and “language” because each language corresponded to one country, except for English (which corresponded to America, Australia, and England). However, given that “country” better explained the heterogeneity and that we expected cultural differences among English-speaking countries, we decided to harmonize STAI-T scores based on country (rather than language). The harmonization was conducted with ComBat for ENIGMA[12] (see expanded code in the figshare repository):

```
i_controls = which(X$patient == 0)
age_sex = cbind(X$age, X$sex)
combat = combat_fit(X$stai[i_controls],
                    site = X$country[i_controls], cov = age_sex[i_controls,],
                    n.min = 8, impute_missing_cov = TRUE)
X$stai  = combat_apply(combat, X$stai, site = X$country, cov = age_sex)$dat
```

Quality control

Three investigators (EV, HS, MAF) independently performed quality control of the non-imaging data and contacted the sites for additional information when necessary.

Neuroimaging data

Quality control

Data were collected from 2448 participants. In addition to quality control using HALFpipe, which excluded 229 individuals (Supplementary Table S4), two investigators (EV, HS) independently reviewed all neuroimaging data, which excluded 20 additional participants. Two of the included samples (Manitoba_Greening_sample_1 and Manitoba_Greening_sample_2) were analyzed in different runs. For these samples, we used the average of all runs to obtain the main contrast. One sample (Harvard_McLaughlin) was analyzed using blocks; due to the short interval-stimulus-interval (ISI), individual events could not be reliably obtained.

Statistical analyses. Mega-analyses

Pre-scaling

As noted in the main text, after processing with HALFpipe, we observed differences in the BOLD response between sites. Such variability exceeded the expected small normally distributed differences typically addressed by site-harmonizing mixed-effects models such as ComBat[12]. To remove these differences, we performed a pre-scaling step that consisted of dividing the BOLD response of individuals from each site by their standard deviation. The use of such standardized scores is common in many areas of psychology and neuroscience. Specifically, for each voxel with brain

coverage across all sites, we estimated the standard deviation using linear models with appropriate covariates (see below). We then calculated the median of the standard deviations across these voxels and divided all images in the sample by this standard deviation. We have included this step in the "combat.enigma" package[12] in R for use by other groups. Following recommendations for between-site harmonization (see below), we estimated the standard deviations exclusively using data from healthy controls.

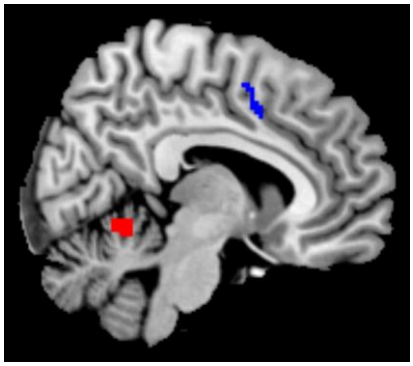
A note about the use of z-statistics in mega-analyses

HALFpipe generates "*z*-statistic images", and one may (wrongly) assume that these *z*-statistic images are equivalent to *z*-scores. However, *z*-statistic images are calculated by dividing each participant´s mean BOLD response (to different trials) by its standard error rather than by the standard deviation across participants. Thus, critically, these *z*-statistic images mix the task-related BOLD response with its standard error. This is not inherently wrong, but it means that differences in *z*-statistics between cases and controls may be due not only to differences in the task-related BOLD response but also to differences in its standard error.

These differences in standard error could be unrelated to the task, for example, due to differences in the amplitude of BOLD signal fluctuations. In the following R code, we simulated a study comparing the task-related BOLD response between cases and controls, with no actual differences in the task-related BOLD response but differences in its standard error. As expected, the t-tests comparing the groups show no differences in the task-related BOLD response. However, they do show statistically significant differences in within-subject *z*-scores.

```
# Create a task time-series design matrix
design = rep(c(rep(0:1, 20), 0), each = 8)
dat = NULL

# For each group
for (group in c("patient", "control")) {

  # For each individual in the group
  for (i in 1:30) {

    # Simulate the BOLD signal with the same BOLD response but more noise
    # in patients
    ts = rnorm(length(design), design, ifelse(group == "patient", 1.2, 1))

    # Simplified analysis to estimate the task-related BOLD response
    m = summary(lm(ts ~ design))$coefficients[2,]

    # Save the individual task-related BOLD response and z-statistic
    dat = rbind(dat, data.frame(
      group,
      bold_response = m[1],
      z_statistic = m[1] / m[2]
    ))
  }
}

# Conduct t-tests to compare patients and controls
t.test(dat$bold_response ~ dat$group)
t.test(dat$z_statistic ~ dat$group)
```

In other words, we do not know whether differences in *z*-statistics are related to differences in the task-related BOLD response or to differences in other aspects of the BOLD signal that may be unrelated to the task. Indeed, we examined whether cases (individuals with anxiety-related and depressive disorders) and controls in this study might have different standard errors of the fear conditioning-related BOLD response and found that they might. For each sample containing cases and controls, we calculated the standardized mean difference (Cohen's *d*) in standard error and then averaged d across the samples. At a descriptive level, using a threshold of $d \geq 0.2$, cases showed larger standard errors in the cerebellum, but smaller in the mid-cingulum (see figure).



Linear mixed-effects models

To fit the models, we created a new function that, for each voxel, performs the following steps:

1) Assesses which participants and sites have information, taking into account the specific brain coverage of each individual fMRI scan;

2) Detects and discards collinear or constant covariates, which can vary depending on the participants with information in that voxel;

3) Fits a linear mixed-effects model using the "lme" function from the “nlme” R package[13]:

m = lme(y ~ x, random = ~ 1 | sample)

or a simple linear model if the participants are from only one sample:

m = lm(y ~ x)

Where “m” is the model, “y” is the voxel value, “x” is a matrix with the variables of interest and covariates, and “site” is a random intercept.

4) Tests the linear hypothesis if specified (e.g., for ANOVAs):

linearHypothesis(m, hypothesis)

where “m” is the model, and “hypothesis” is the hypothesis matrix.

5) Saves the results, including maps of sigma (the standard deviation estimated in the model), the model coefficients, and z-statistics. We have included this function in the “combat.enigma” R package.

We used cluster-based inference to correct for multiple testing. Specifically, we created clusters of voxels with $Z \geq 3.1$ and converted cluster sizes to cluster-wise p-values using the Gaussian Random Field (GRF) theory, using the FSL utilities smoothest and cluster.

Normative modelling: Thresholding

We follow existing work in the field and apply a threshold equal to or greater than 2.6 ($Z \pm 2.6$) to determine ‘large’ deviation scores. With the normative modelling approach, we are not performing a group-level hypothesis test as would be done using a statistical parameteric mapping framework, rather we are aiming to statistically detect differences in individual subjects with respect to the common

reference model. For this type of analysis, it is not clear that how spatial comparisons, and multiple comparison correction ought to be done, nor that doing so is even appropriate because multiple comparisons correction obscures the degree of inter-individual differences that are detected. In prior work[14], a subject level FDR method was evaluated where the results did not differ from the conclusions made when using the original un-corrected input data. The interpretation of the single subject FDR correction method is not straightforward as a different threshold is estimated for every subject. In other words, an individual with widespread decreased or increased BOLD activation during the task (as quantified using a contrast z-statistic) may appear to have 'normal' findings using an FDR threshold as the overall distribution of their voxel intensities is shifted. Therefore, in this work we elected to use a single fixed threshold (Z±2.6) to determine the significance of a deviation, which simplifies the comparison across individuals and is in line with other work on normative models[14-16].

Sample size analyses

Leveraging data from the 43 samples included in the study, we conducted a series of sample size analyses to inform the design of future studies. Specifically, we treated the mega-analytic primary activation map as the ground truth and examined how activation patterns at varying sample sizes compared to this reference. To ensure harmonization across datasets, all analyses were re-run using the lmm_fit function and corrected for multiple comparisons using Gaussian random field theory. The primary objective was to examine how study sample size relates to the proportion of brain regions showing activation or deactivation in the mega-analysis. To this end, we parcellated the brain using the AAL atlas[17], and defined a region as activated (or deactivated) if it contained at least 10 statistically significant activating (or deactivating) voxels. For each study, we then calculated the percentage of AAL regions identified as activated (or deactivated) in the mega-analysis that were also detected in the individual study. The relationship between the arcsine-transformed percentage of detected regions and the square root of the study's sample size was subsequently assessed.

For completeness, we also computed the average false positive rate—defined as the percentage of regions not activated or deactivated in the mega-analysis but incorrectly identified as such in the individual studies. All analyses were conducted separately for activations and deactivations. Secondary analyses examined how the square root of the study's sample size related to several additional metrics:
a) the Fisher-transformed correlation between the study and mega-analytic maps;
b) the mean absolute error between the two; and
c) the arcsine-transformed Dice coefficient[18] quantifying their spatial overlap.

Supplementary Discussion

In the main text, we highlighted those individual differences or task variables with more robust effects. Here we discuss the remaining significant associations.

Sociodemographic variables

Older age was significantly associated with greater activation in the ventromedial prefrontal cortex and medial temporal gyrus, as well as significantly less activation in the anterior insula, pre-supplementary motor area extending to the dorsal anterior cingulate, dorsal caudate and bilateral supramarginal gyrus extending to the posterior insula. Female participants (*n*=973) showed greater activation across the visual cortex, and left medial/superior temporal gyrus than males (*n*=915). Regression coefficients from the normative models indicated a minimal effect of age on the predicted BOLD signal, but unthresholded effects largely replicated the findings of the mega-analysis. Structure coefficients from the normative models showed minimal relation between sex and predicted BOLD signal, with only a very small cluster in the mid-anterior cerebellum predicted to show heightened activation in females. These results are presented in **Sup. Figure S5**.

Task variables

The following task variables showed significant albeit small/weak associations with brain activation during conditioning (see **Sup. Figure S8** for the mega-analysis results and **Sup. Figure S9** for the structure coefficients of the normative modeling results). Normative modelling regression coefficient maps are also shown in **Sup. Figure S9** for completeness but are not discussed below.

The number of trials during preconditioning showed a significant positive association with activation in the inferior cerebellum in the mega-analysis. Structure coefficients did not show a relationship between the number of trials during preconditioning and predicted BOLD signal.

In the mega-analysis, the type of CS (categorized as humanoid, affective pictures, and neutral faces) revealed significant effects. See full results at https://figshare.com/s/d44cc1390711bad3c147.In normative modeling analyses, the use of a humanoid CS was predictive of increased activation in the cingulate cortex, extending to the dorsomedial prefrontal cortex and pre-supplementary motor area, secondary somatosensory cortex (SII), dorsal precuneus, dorsolateral prefrontal cortex, the bilateral insula, the bilateral temporoparietal junction, the thalamus, the caudate and the left anterior cerebellum, as well as decreased activation (i.e. more deactivation) in the anterior ventromedial prefrontal cortex and posterior cingulate cortex. Moreover, the use of neutral pictures as CS predicted more activation (i.e. less deactivation) in the anterior ventromedial prefrontal cortex and posterior cingulate cortex, and less activation within the cingulate cortex, extending to the dorsomedial prefrontal cortex and pre-supplementary motor area, dorsal precuneus, SII, the bilateral insula, the bilateral temporoparietal junction, the thalamus, the caudate and left anterior cerebellum Finally, the use of neutral faces as CS predicted more activation within the subgenual anterior cingulate cortex, and less activation within the bilateral fusiform face area and SII. The use of other types of CS (affective faces and pictures, a gabor patch, a neutral male avatar, images of animals or tools, or of snakes and spiders) did not have an influence on predicted BOLD signal.

Average intertrial-interval (ITI) length demonstrated a significant positive association with activation within the bilateral primary visual cortex and a significant negative association with the bilateral posterior parietal cortex, and superior frontal gyri extending to the supplementary motor area in the mega-analysis. Structure coefficients showed that increased average ITI was predictive of increased activation within the primary visual cortex, dorsomedial prefrontal cortex, extending to the pre SMA, the bilateral thalamus, caudate and putamen, the brainstem, and the anterior and medial cerebellum. Conversely, a longer ITI predicted less activation (i.e., more deactivation) within an expanse of the ventromedial prefrontal cortex, within the dorsolateral prefrontal cortex, primary somatosensory cortex (SI) the precuneus, the

lingual gyrus and fusiform face area extending into bilateral middle gyri of the temporal lobe, and bilateral hippocampus.

In the mega-analysis, the number of CS+ included in the fMRI contrast showed a significant positive association with activation in the left primary visual cortex, right orbitofrontal cortex, right precuneus, right superior parietal lobule, and right dorsolateral prefrontal cortex. Moreover, the number of CS- included in the fMRI contrast showed a significant positive association with activation in the left superior parietal lobule and the right dorsolateral prefrontal cortex.

Being unaware of the relationship between CS and US (i.e., contingency unawareness; *n*=72) showed a positive association with activation in the ventral posterior cingulate extending to the dorsal anterior cingulate/precuneus compared with being aware (*n*=1260). As contingency awareness was not available for all participants this variable was not included in the construction of the normative models, and therefore their relationship to predicted task (de)activation cannot be assessed using structure coefficients. Rather, for participants in the two test samples (controls + individuals with an anxiety or mood-related disorder) with these data available (n = 703) we used a support vector classifier and found whole-brain deviation score could not be used to predict whether a participant was contingency aware or not (mean accuracy = 50% +/- 16%; p = 0.426; 10-fold cross validation; 1000 permutations).

For the main results on type of US, please refer to the main text. In addition to these main results, in normative modeling analyses, the use of a thermal stimuli as US was predictive of decreased activation within the bilateral amygdala, the mid-cingulate cortex extending to the pre-supplementary motor area, the dorsomedial prefrontal cortex, a posterior region of the ventromedial prefrontal cortex, the cuneus, and (i.e., more deactivation) in the angular gyrus. The use of a visceral stimuli as US had no influence on predicted BOLD signal during CS+>CS-. These two variables were not investigated separately using linear models.

The following task variables were not significant in the mega-analysis nor in normative modelling analyses: number of CS+ trials during conditioning; number of CS- trials during conditioning; and average ISI (inter-stimulus interval, i.e., between the CS+ and the US).

Supplementary References

1. Crawford, J., Cayley, C., Lovibond, P. F., Wilson, P. H. & Hartley, C. Percentile norms and accompanying interval estimates from an Australian general adult population sample for self-report mood scales (BAI, BDI, CRSD, CES-D, DASS, DASS-21, STAI-X, STAI-Y, SRDS, and SRAS). *Aust. Psychol.* **46**, 3–14 (2011).
2. Guillén-Riquelme, A. & Buela-Casal, G. Psychometric revision and differential item functioning in the State Trait Anxiety Inventory (STAI). *Psicothema* **23**, 510–515 (2011).
3. Sandin, B., Chorot, P. & McNally, R. J. Anxiety sensitivity index: normative data and its differentiation from trait anxiety. *Behav. Res. Ther.* **39**, 213–219 (2001).
4. Spielberger, C. D. *Manual for the State-Trait Anxiety Inventory (self-evaluation questionnaire)*. (1970).
5. Caci, H., Baylé, F. J., Dossios, C., Robert, P. & Boyer, P. The Spielberger trait anxiety inventory measures more than anxiety. *Eur. Psychiatry* **18**, 394–400 (2003).
6. Kennedy, B. L., Schwab, J. J., Morris, R. L. & Beldia, G. Assessment of state and trait anxiety in subjects with anxiety and depressive disorders. *Psychiatr. Q.* **72**, 263–276 (2001).
7. Marteau, T. M. & Bekker, H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br. J. Clin. Psychol.* **31**, 301–306 (1992).
8. Iwata, N. & Higuchi, H. R. Responses of Japanese and American university students to the STAI items that assess the presence or absence of anxiety. *J. Pers. Assess.* **74**, 48–62 (2000).
9. Bieling, P. J., Antony, M. M. & Swinson, R. P. The State-Trait Anxiety Inventory, Trait version: structure and content re-examined. *Behav. Res. Ther.* **36**, 777–788 (1998).
10. Brand, S. *et al.* A multi-site German validation of the Interoceptive Accuracy Scale and its relation to psychopathological symptom burden. *Commun. Psychol.* **1**, 14 (2023).
11. Ryckewaert, R. Aandachtsbias voor doodsgerelateerde informatie bij ouderen versus jongeren. *Thesis*, Universiteit Gent (2008).
12. Radua, J. *et al.* Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *Neuroimage* **218**, 116956 (2020).
13. Pinheiro, J., Bates, D. & R Core Team. *nlme: Linear and Nonlinear Mixed Effects Models* (2024).
14. Wolfers, T. *et al.* Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. *JAMA Psychiatry* **75**, 1146–1155 (2018).
15. Holz, N. E. *et al.* A stable and replicable neural signature of lifespan adversity in the adult brain. *Nat. Neurosci.* **26**, 1603–1612 (2023).
16. Floris, D. L. *et al.* Atypical brain asymmetry in autism—a candidate for clinically meaningful stratification. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **6**, 802–812 (2021).

17. Tzourio-Mazoyer, 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).
18. Zou, K. H. *et al.* Statistical validation of image segmentation quality based on a spatial overlap index: Scientific reports. *Acad. Radiol.* **11**, 178–189 (2004).

nature portfolio

Corresponding author(s): Miquel A. Fullana

Last updated by author(s): Jul 23, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No specific software was used for data collection
Data analysis	The following software was used: 'nlme' R package; Predictive Clinical Neuroscience toolkit (PCNtoolkit) software v0.26 implemented in python 3.8; scikit- learn package (version 1.0.2). All code to reproduce the analyses in this manuscript is available at: https://doi.org/10.6084/m9.figshare.28540580.v1 The functions needed to conduct the mega-analysis are also included in the 'combat.enigma' R package.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The results data generated in this study have been deposited at: https://doi.org/10.6084/m9.figshare.28540580.v1 The individual-level fMRI processed data (HALFIPE results files) are available for secondary data analysis. Access can be obtained by submitting an analysis plan to the ENIGMA-Anxiety Working Group (http://enigma.ini.usc.edu/ongoing/enigma-anxiety/). Data access is contingent on approval by PIs from contributing samples. The raw individual fMRI data are protected and are not available due to data privacy laws. Source data are provided with this paper.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Sex (biological sex) was included as a main variable or covariate in some analyses. Information on sex distributions is reported for all main analyses.
Reporting on race, ethnicity, or other socially relevant groupings	We did not include race, ethnicity or other socially relevant groupings in our analysis.
Population characteristics	Healthy participants of both sexes and indidivualswith a diagnosis of anxiety-related or depressive disorders, as recruited in original studies.
Recruitment	Investigators of eligible studies were contacted and invited to share data
Ethics oversight	The manuscript reports on previously collected data from individual studies. All studies received ethical approval. The following Ethics Committees approved the individual studies: Ethics Committee of Ulm University, Ulm, Germany; Regional Ethics Review Board in Uppsala; Institutional Review Board of University Hospital of Bellvitge, Barcelona, Spain and Institutional Review Board of Hospital del Mar, Barcelona, Spain; Institutional Review Board University of Wisconsin; Institutional Review Board University of Arkansas for Medical Sciences; Comissão de Análise de Projetos de Pesquisa do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (CAPPesq); Institutional Review Board University of Texas at Austin; Ethics Committee of the University Hospital Essen, Germany; University of Southern California Institutional Review Board; Louisiana State University Institutional Review Board; The University of Melbourne Human Research Ethics Committee; Institutional Review Board University of Minnesota; Institutional Review Board University of Florida; Ethics Committee of Klinikum Rechts der Isar, Technische Universität München; Duke University Health System Institutional Review Board; Ethics Committee of the General Medical Council Hamburg; University of Washington and Harvard University Institutional Review Boards; Ethics Committee of the Faculty of Medicine of the Ruhr University Bochum; Ethics Committee of the Faculty of Psychology and Sport Science, University of Giessen; University Research Ethics Committee of the University of Reading, UK; Clinical Research Ethics Committee (CEIC) of the Bellvitge University Hospital; Ethics Committee of the German Psychological Society; Partners HealthCare Institutional Review Board; Ethics Review Board (ERB) University of Amsterdam; Institutional Review Board, New York State Psychiatric Institute.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	2199 individual participants. Sample size was not determined a-priori because this is a mega-analysis based on available data.
Data exclusions	Data were collected from 2448 participants. In addition to quality control using HALFpipe, which excluded 229 individuals (Sup. Table S3), two investigators independently reviewed all neuroimaging data, which excluded 20 additional participants. For some normative modeling analyses, there were less participants available than for linear model analyses (see Table 1 in the main text)

Replication	For the mega-analysis on trait anxiety, analyses were conducted with both harmonized and non-harmonized data.
Randomization	NA to mega-analysis or normative modeling analyses
Blinding	NA

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	NA
Study protocol	NA
Data collection	NA
Outcomes	NA

Plants

Seed stocks	NA
Novel plant genotypes	NA
Authentication	NA

Magnetic resonance imaging

Experimental design

Design type	Task-based (fear conditioning) fMRI
Design specifications	See Table 2 in the main manuscript
Behavioral performance measures	See Table 2 in the main manuscript

Acquisition

Imaging type(s)	Functional
Field strength	Tesla 1.5 and Tesla 3
Sequence & imaging parameters	Variable
Area of acquisition	Whole Brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	Various
Normalization	HALFpipe, version 1.2.2
Normalization template	HALPIPE default template: MNI152NLin2009cAsym
Noise and artifact removal	HALFPIPE default parameters (smoothing using 6mm FWHM; confound removals using 655 ICA-AROMA; and a high-pass filter of 125 s)
Volume censoring	See quality control (Sup.Mat)

Statistical modeling & inference

Model type and settings	We conducted mega-analysis and normative modeling analyses. For mega-analyses, we mostly used LMMs (with the sample as random intercept) with age and sex as covariates. For normative modeling, we used different models. We trained a Gaussian Bayesian Linear Regression (BLR) model to predict activation for the main contrast and assessed generalisability by using a stratified train-test sample. To assess the association with other variables, we examined the regression coefficients and the structure coefficients of all input variables.
Effect(s) tested	CS+>CS- (main contrast). Association of task variables with main contrast. Differences patients-controls. Differences among patient groups.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☒ Both

Anatomical location(s)	We conducted whole-brain mega-analyses and normative modeling analyses and one ROI mega-analysis focused on the amygdala region, using the automated Anatomical Labeling atlas, version 3 (AAL3).
Statistic type for inference (See Eklund et al. 2016)	FSL was used to derive cluster-based corrected p-values using Gaussian Random Field (GRF) theory.
Correction	MEGA-ANALYSIS. We used cluster-based inference to correct for multiple testing. Specifically, we created clusters of voxels with Z≥3.1 and converted cluster sizes to cluster-wise p-values using the Gaussian Random Field (GRF) theory, using the FSL utilities smoothest and cluster. For ANOVA-like LLMs, we conducted pairwise comparisons with Holm- Bonferroni correction, NORMATIVE MODELING. False Discovery Rate (FDR) corrected at $\alpha = 0.05$ to quantifyi voxel-wise deviations from the reference normative model. 1000-fold nested 5-fold test

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis