Supplementary Online Content

Marin M-F, Hammoud MZ, Klumpp H, Simon NM, Milad MR. Multimodal categorical and dimensional approaches to understanding threat conditioning and its extinction in individuals with anxiety disorders. *JAMA Psychiatry*. Published February 5, 2020. doi:10.1001/jamapsychiatry.2019.4833

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Participants, Paradigm, Data Processing, Analytic Approach, *DSM-IV* vs RDoC

Participants

The following exclusion criteria were used: history of seizures or significant head trauma, current substance abuse or dependence, metal implants, pregnancy, breastfeeding, or positive urine toxicology screen for drugs of abuse. For the anxiety groups, participants needed to be either medication-free for at least 8 weeks or on a stable medication regimen (only SSRI and atypical antipsychotics were allowed) for at least 8 weeks. They also needed to be free of benzodiazepines for at least 2 weeks. Some of the data used were published in a manuscript comparing individuals suffering from anxiety disorders (n=61) to healthy controls (n=21)[14].

Paradigm

The structure for each trial of this paradigm is the same across the different phases: a black screen ranging between 12 and 18 seconds (with an average of 15 seconds) (inter-trial interval), which is followed by a picture of a room with an unlit lamp for 3 seconds (context) after which the lamp turns on to blue, red or yellow (stimulus). On the first day, participants have to select their level of electric stimulation to be used during the experiment, so that it is highly annoying but not painful. Habituation occurs where three different neutral stimuli (colored lamps) are presented in two distinct contexts (either an office or a library). Following habituation, fear conditioning occurs in one context (e.g., the office) where two of the colored lamps (e.g., blue and red, CS+) and partially reinforced (62.5% reinforcement rate) with a mild electric shock (500ms, occurring at the offset of CS presentation) and the other colored lamp (e.g., yellow) is never paired with the shock (CS-). Fear conditioning consists of a total of 32 trials (8 presentations for each of the two CS+s and 16 presentations of the CS-). Following this, extinction learning occurs in a different context (e.g., the library), where one of the CS+ is presented 16 times without any shocks (CS+ extinguished; CS+E) intermixed with 16 presentations of the CS-. The next day, extinction memory recall is tested in the context of extinction learning (e.g., library), where the three colored lamps are presented (8 presentations of the extinguished CS+E, 8 presentations of the unextinguished CS+ (CS+U) along with 16 presentations of the CS-). For all phases of the paradigm, the order of stimulus presentation is pseudo-random. Note that this paradigm uses a A-B-B design, in which the conditioning and extinction contexts are different. One of the advantage of this design is that it parallels clinical implications where the initial fear memory is learned in a given context and therapy provides a different context to learn about safety.

Data Processing

Psychophysiological data: For fear conditioning, extinction learning, and extinction recall, conditioned responses were obtained for each trial by subtracting the average skin conductance response (SCR) levels exhibited during the last two seconds of context presentation from the maximum SCR level exhibited during CS presentation. For the fear conditioning phase, unconditioned responses were also calculated by subtracting the average skin conductance levels obtained during the last 2 seconds of the CS presentation from the maximal skin conductance levels obtained during the 6 seconds following CS offset. SCR analyses for the conditioned and unconditioned responses were performed on square-root transformed data (if negative values were obtained, square-root transformation was applied to the absolute value and a negative sign was then assigned to the square root transformed value). SCR were analyzed using mixed ANCOVAs. Greenhouse-Geisser corrected values are reported, when necessary. Significant main effects or interactions were decomposed using Tukey post-hoc t-tests.

Imaging data: Data were acquired in a Trio 3.0 Tesla whole-body MRI scanner (Siemens Medical Systems, Iselin, New Jersey) using a 32-channel head coil (TR: 2.56 seconds, TE: 30 ms, slice number: 48, voxel size: 3x3x3 mm). Matlab and SPM12 were used for all imaging data analysis. All coordinates reported are based on the Montreal Neurological Institute (MNI) system. Functional images were corrected for slice timing, realigned, co-registered with the structural image, normalized into MNI space, and smoothed with a 8mm full width half-maximum Gaussian kernel. Signal drift and biorhythms were modeled using high-pass temporal filtering (128 s), an autoregressive AR-1 model. Artifact detection toolbox (ART, http://gablab.mit.edu) was used and motion artifact data detected by ART were used in the first-level analysis as regressors with movement parameters (x, y, z, roll, pitch, and yaw) from the realignment process. Motion regressors generated from the ART-tool were then applied to all first-level analyses.

First-level contrast images were obtained for each subject and then modeled at the second level using a mixed linear model. During conditioning, the following contrasts were examined: 1) CS+ vs. CS- (contrasting the 16 CS+ trials to the 16 CS- trials) and 2) Shock vs. No Shock (using the CS offset for the 10 CS+ trials for which a shock was

administered and contrasting them to CS offset of the 16 CS- trials). For extinction learning, we examined 1) early extinction (contrasting the first four CS+E trials to the first four CS- trials) and 2) late extinction (contrasting the last four CS+E trials to the last four CS- trials). For extinction recall, the first four CS+E trials were contrasted to the first four CS- trials.

For imaging, we created masks for the following ROIs: amygdala, hippocampus, insular cortex, dorsal anterior cingulate cortex, and ventromedial prefrontal cortex. For amygdala, hippocampus and insular cortex, masks were created using the Automated Anatomical Labeling (AAL). For dACC and vmPFC, masks were created with the use of Neurosynth (neurosynth.org), which is a searchable online automated synthesis of fMRI data. By using the keywords 'fear conditioning', we identified clusters within 248 studies with estimated peaks at the following coordinates: dACC (MNI_{xyz} = 0, 14, 28) and vmPFC (MNI_{xyz} = 6, 40, -20) and created a mask for each region. Activations detected in one of these five ROIs with an uncorrected threshold of $P < 0.005$ were tested for small volume corrections using the corresponding mask. Only clusters that survived the small volume correction (family-wise error; pfwe<0.05) are reported. For clusters reaching statistical significance, beta-weights were extracted for visual display of the data and to establish the between-group differences. Note that analyses that included covariates graphed the estimated marginal means (which take into account the covariates).

Composite score of Anxiety, Mood, and Personality traits (C.A.M.P.): In order to reduce the number of comparisons and to compute one global index of self-report measures, all raw scores obtained ASI, BAI, BDI, STAI-T, and the five subscales of the NEO-PI-5 were transformed to a z score. Z scores were inversed for extraversion, agreeableness, conscientiousness, and openness. An average z score was then computed for each participant to create a composite score of anxiety, mood, and personality traits (C.A.M.P.).

Analytic Approach

RDoC-based analytic approach: Given that recent studies have highlighted arousal as a key metric for predicting PTSD symptoms months after trauma, the high relevance of arousal to anxiety disorders, and that arousal is one key domain within RDoC, we calculated the threat-induced arousal for each individual. For each participant suffering from an anxiety disorder, the average UCR to the 10 CS+ trials for which a shock was delivered was computed. UCR averages were then Z score transformed and this was used to rank all anxious participants (from lowest arousal to highest arousal). Irrespective of DSM diagnosis, we artificially created four groups based on arousal, with group 1 exhibiting the lowest level of arousal in response to the shock and group 4 exhibiting the highest level of arousal in response to the shock.

In order to validate this new metric at the neural level, a whole-brain analysis was performed with the z score obtained from the arousal metric as the regressor during the Shock vs. No Shock contrast. For this analysis, we investigated significant relationships within the threat and extinction network, without using predefined masking (whole-brain approach). We used a threshold of $P < 0.005$ and 10 contiguous voxels. Only clusters that survived the small volume correction (using a 8-mm sphere around the peak, family-wise error; pfwe < 0.05) are reported. For clusters reaching statistical significance, beta-weights were extracted for visual display of the data and a Pearson correlation coefficient (r) was calculated.

For both approaches (DSM and RDoC), between-group SCR differences were tested during fear conditioning (CS+ vs. CS- and Shock vs. No Shock) using Stimuli X Group ANCOVAs. Extinction learning for SCR was examined through a Time X Group ANCOVA (using 8 bins of 2 CS+E trials each). Finally, SCR during early extinction recall was tested using a Stimuli X Group ANCOVA. For the fMRI data, the following contrasts were tested with an ANCOVA: during threat conditioning (CS+ vs. CS- and Shock vs. No Shock); during extinction learning (early CS+E vs. early CS- and late CS+E vs. CS-); during extinction recall (early CS+E vs. early CS-).

***DSM-IV* vs. RDoC**

For each cluster within the ROIs for which a significant result was obtained and computed an ANOVA for the alternative approach (for a significant cluster obtained with the DSM approach, we compared the activations for that same cluster across the four groups formed on arousal and the reverse for the clusters identified with the dimensional approach). These analyses were performed in anxious individuals only. To keep consistency with the analytic approach described above, the DSM grouping included age and education as covariates whereas the RDoC grouping included age as a covariate.

Exploratory analyses for threat-induced arousal and self-reported metric

We conducted two sets of exploratory analyses for the different phases of the paradigm using 1) threat-induced arousal z scores as a regressor and then 2) C.A.M.P. z scores as a regressor. For these exploratory analyses, we investigated significant relationships within the threat and extinction network, without using predefined masking (whole-brain approach). We used a threshold of $p<0.005$ and 10 contiguous voxels.

eResults. Sample Characterization for *DSM-IV* Approach, Sample Characterization for RDoC Approach, and Whole-Brain Analysis

Sample characterization for *DSM-IV* approach

When comparing the four anxiety disorders (SP (n=20), SAD (n=28), GAD (n=27), and PD (n=18)) and the healthy controls (n=21) for their demographic characteristics, a main effect of Group was found for age ($F_{4,109}$=5.13, p=0.001), and for education ($F_{4,109}$=3.29, p=0.01). Tukey post-hoc tests revealed that the SP group was significantly older than all other groups (all ps≤0.05, **eFigure 2a**), whereas the HC group had significantly more years of education than the SAD group (p=0.01, **eFigure 2b**). All five groups selected similar shock level, ($F_{4,99}$=0.69, p=0.6, **eFigure 2c**). The one-way ANCOVA ($F_{4,108}$=12.59, p<0.001) revealed that both HC and SP groups had lower C.A.M.P. z scores relative to SAD, GAD, and PD groups (all ps≤0.002, **eFigure 2d**). The main DSM diagnosis had an impact on whether participants suffered from one or multiple disorders, $\chi^2(3)$=10.50, p=0.02 (**eFigure 2e**).

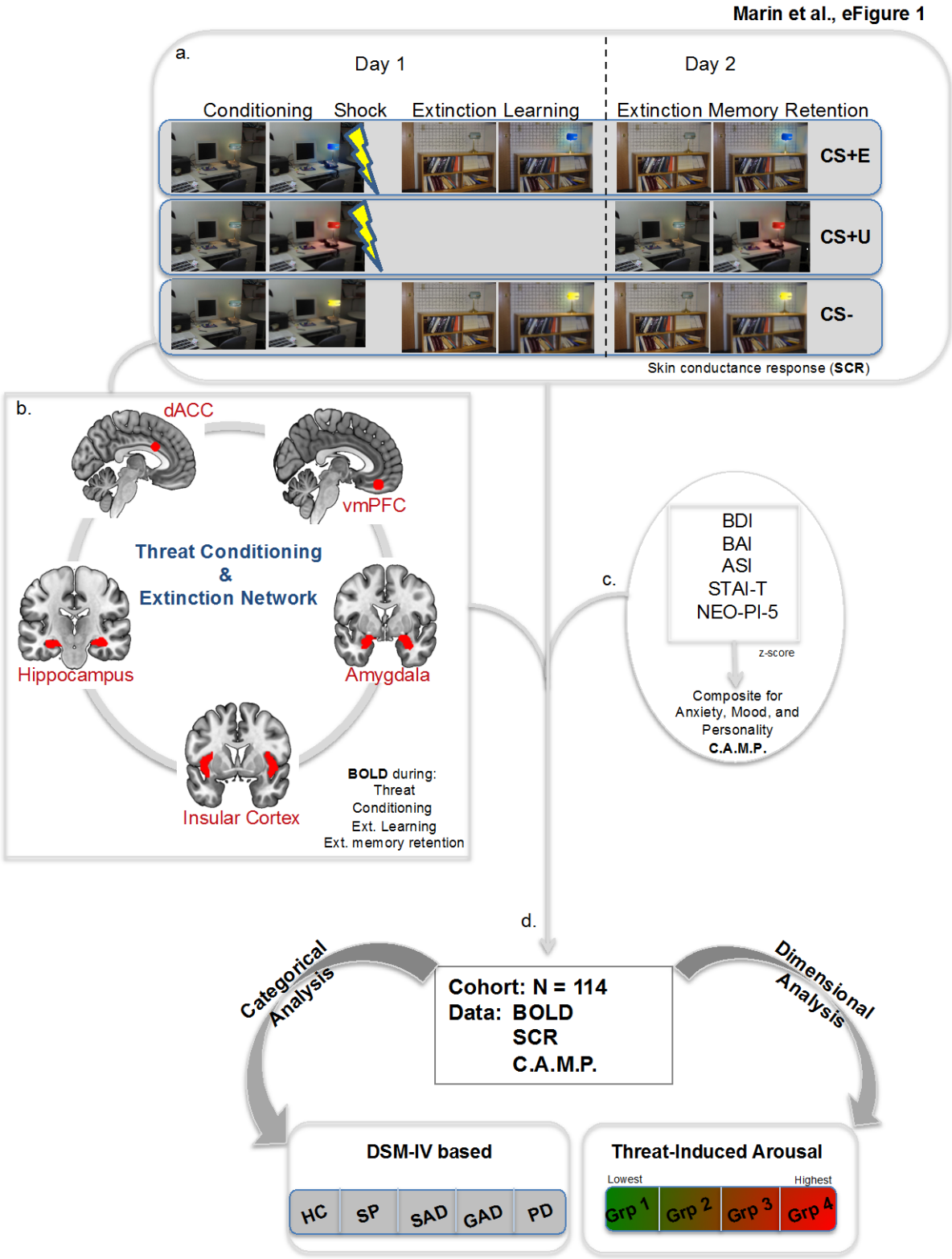
Sample characterization for RDoC approach

With regards to the demographic characterization of the new grouping based on the threat-induced arousal (TIA), a main effect of Group was found for age, ($F_{3,83}$=6.15, p=0.001), with groups 1 and 2 being older than group 4 (p≤0.04**,** **eFigure 3a**). No significant effect of Group was found for education ($F_{3,83}$= 1.13, p=0.34, **eFigure 3b**), and shock level ($F_{3,73}$= 1.65, p=0.19, **eFigure 3d**). A group effect was found for C.A.M.P. scores ($F_{3,82}$=3.81, p=0.01), where group 1 and 3 had lower scores relative to group 4, ps≤0.01 (**eFigure 3d**). Groups did not differ in having a single or multiple anxiety disorders diagnosis, $\chi^2(3)$=5.88, p=0.12 (**eFigure 3e**).

Whole-brain analyses: TIA and C.A.M.P. as regressors

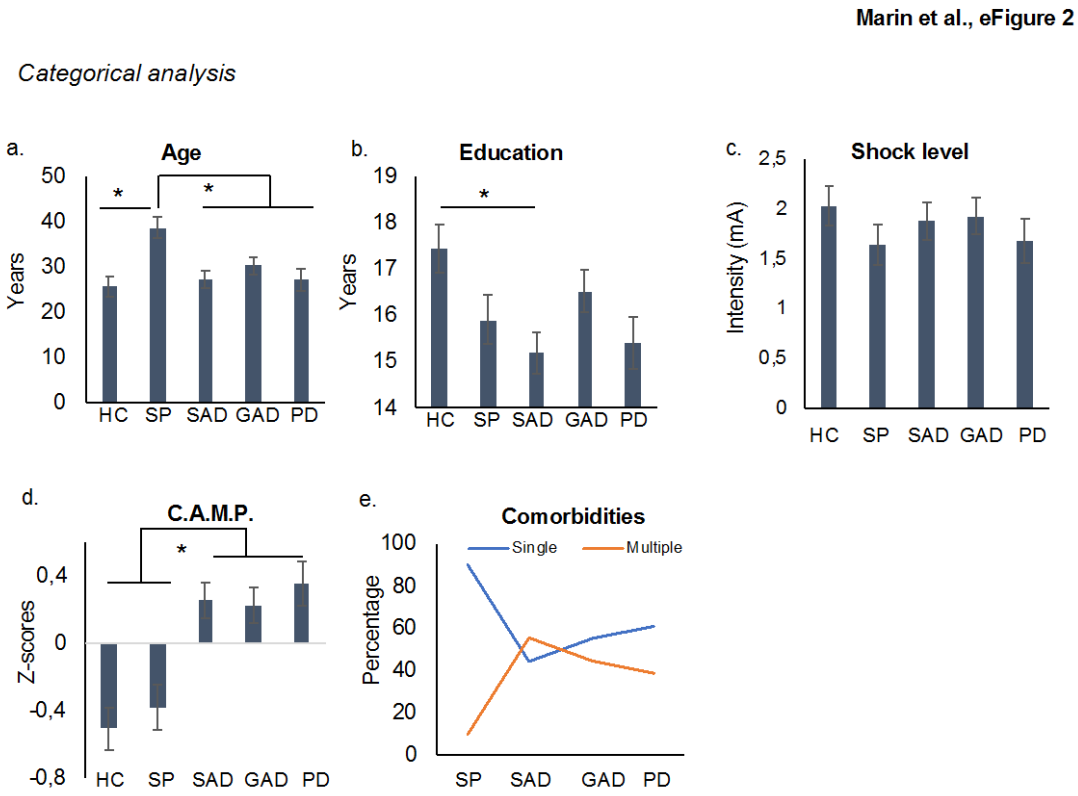
Whole-brain exploratory analyses revealed that threat-induced arousal scores obtained during shock delivery were negatively associated with vmPFC activation during threat conditioning (CS+ vs. CS- contrast) (MNI_{xyz}= 8, 54, -16; cluster size = 171, t_{102}=3.38, p_{FWE}=0.02) and with rostral anterior cingulate cortex – rACC during extinction memory retention (MNI_{xyz}= -6, 44, 14; cluster size = 20, t_{95}=3.01, p_{FWE}=0.05) (**eFigure 4a**). The z-score composite for anxiety, mood, and personality (C.A.M.P.) was negatively associated with vmPFC activation during both threat extinction learning (MNI_{xyz}= -6, 34, -12; cluster size =62, t_{106}=3.38, p_{FWE}=0.02) and extinction memory retention (MNI_{xyz}= 16, 54, -10; cluster size = 73, t_{100}=3.35, p_{FWE}=0.02), whereas it positively correlated with posterior hippocampal activation (MNI_{xyz}=32, -38, -4; cluster size =25, t_{100}=3.35, p_{FWE}=0.02) during extinction memory retention (**eFigure 4b**).

eFigure 1. Schematic Representation of the Experimental Protocol, the Measures, and the 2 Analytic Approaches



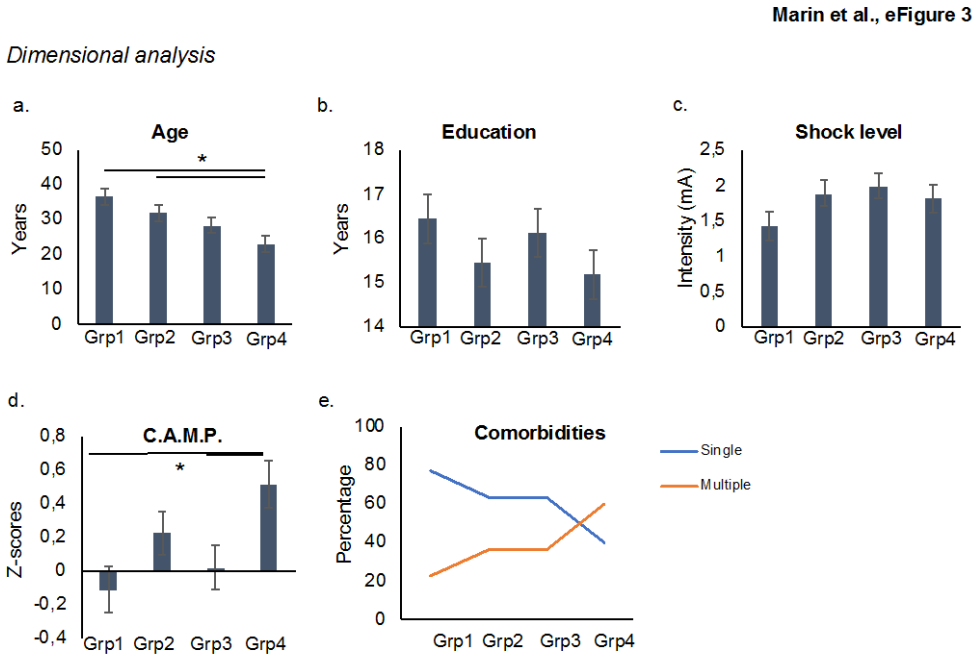
eFigure 1. A) Two-day threat-conditioning and extinction protocol. Skin conductance responses (SCR) were measured at each phase of the protocol to assess conditioned responses. B) Main brain regions involved in threat conditioning and extinction network that were used as regions of interest (ROIs). Blood oxygenated level dependent (BOLD) signal was measured during all three phases of the protocol in those ROIs. dACC = dorsal anterior cingulate cortex; vmPFC = ventromedial prefrontal cortex. C) Participants filled out various questionnaires assessing personality traits as well as mood and anxiety-related symptoms. BDI = Beck Depression Inventory; BAI = Beck Anxiety Inventory; ASI = Anxiety Sensitivity Inventory; STAI-T = State and Trait Anxiety Inventory – Trait Form; NEO-PI-5 = NEO Personality Inventory – 5. Each measure was Z-score transformed and an average z score was computed, which results in a composite score for anxiety, mood, and personality (C.A.M.P.) D) Analytic approach. BOLD and SCR data during the 2-day protocol as well as C.A.M.P. scores were obtained in a total of 114 individuals. Categorial analysis consisted of comparing groups based on their main diagnosis as defined by DSM-IV. HC = healthy controls; SP = specific phobia; SAD = social anxiety disorder; GAD = generalized anxiety disorder; PD = panic disorder. For dimensional analysis, all individuals diagnosed with an anxiety disorder were grouped anew (in 4 groups) based on their SCR to the shock (threat-induced arousal index).

eFigure 2. Characterization of the Groups Based on the Categorical Approach



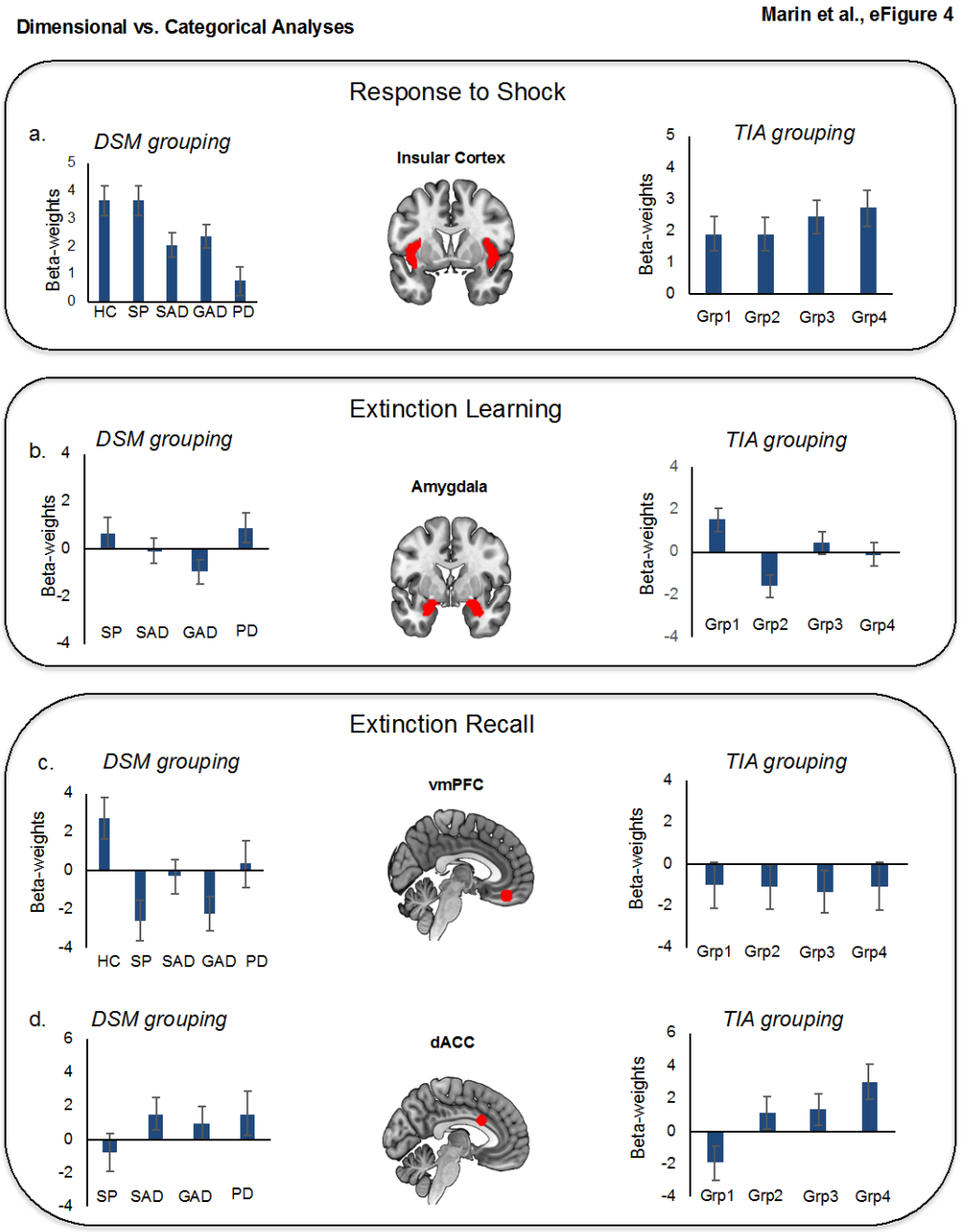
eFigure 2. Upper panel = Categorical Analysis A) Mean age (years) of participants based on their main DSM diagnosis. B) Mean education level (years) of participants based on their main DSM diagnosis. C) Mean selected shock levels (mA) of participants based on their DSM diagnosis. D) Mean z score of Composite of Anxiety, Mood, and Personality (C.A.M.P.) as a function of the DSM groups. E) Percentage of individuals who have comorbid disorders based on their main DSM diagnosis. HC = healthy controls; SP = specific phobia; SAD = social anxiety disorder; GAD = generalized anxiety disorder; PD = panic disorder. Blue line 'single' indicates those who have a single diagnosis (no comorbidities) and the red line 'multiple' indicates those who have at least one comorbid diagnosis. Error bars are S.E.M. Asterisk indicates $p<0.05$.

eFigure 3. Characterization of the Groups Based on the Dimensional Approach



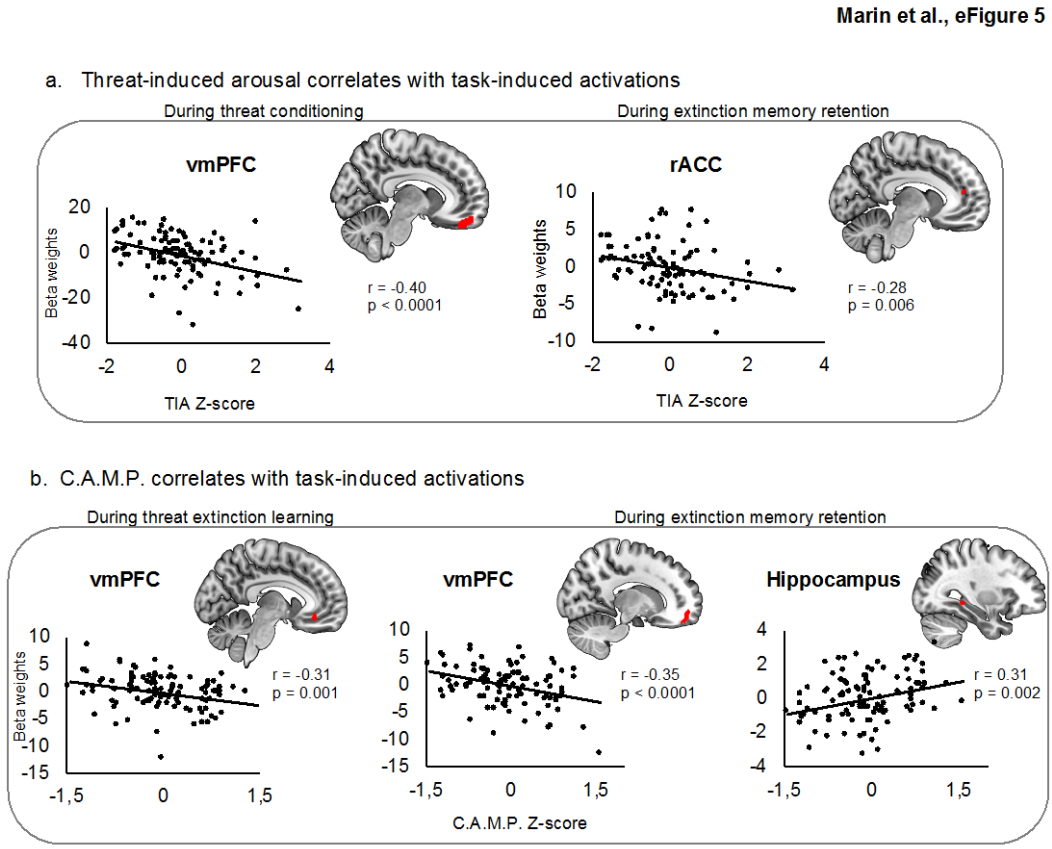
eFigure 3. Dimensional Analysis (only including individuals diagnosed with an anxiety disorder). A) Mean age (years) of participants when divided as a function of their threat-induced arousal (TIA) B) Mean education level (years) of participants when divided as a function of their threat-induced arousal. C) Mean shock level (mA) of participants when divided as a function of their threat-induced arousal. D) Mean z score of Composite of Anxiety, Mood, and Personality (C.A.M.P.) of participants when divided as a function of their threat-induced arousal. E) Percentage of individuals who have comorbid disorders based on their group as defined by threat-induced arousal. Blue line 'single' indicates those who have a single diagnosis (no comorbidities) and the red line 'multiple' indicates those who have at least one comorbid diagnosis. Group 1 (Grp1) to Group 4 (Grp4) are based on the skin conductance levels obtained in response to shock (threat-induced arousal); Group 1 has the lowest threat-induced arousal and Group 4 has the highest threat induced arousal. Error bars are S.E.M. Asterisk indicates $p<0.05$.

eFigure 4. Comparison of Categorial vs Dimensional Analytic Approach



eFigure 4. A) During shock delivery, a significant difference was found in the IC using the DSM approach (left panel). The right panel indicates activations in that same cluster as a function of RDoC grouping. B) During extinction learning, a significant difference was observed in the amygdala using the RDoC approach (right panel). The left panel indicates activations in that same cluster as a function of DSM grouping. C) During extinction recall, a significant difference was observed in the vmPFC using the DSM approach (left panel). The right panel indicates activations in that same cluster as a function of RDoC grouping. A significant difference was found in the dACC using the RDoC approach (right panel). The left panel indicates activations in that same cluster as a function of DSM grouping. The masks used for the different ROIs are displayed on the figure. This analysis was performed in anxious individuals. Therefore, panel A (left) and C (left) have healthy controls in their figure only because the effect was detected in the DSM-grouping (which included healthy controls).

eFigure 5. Voxelwise Analyses of TIA and CAMP Scores Across the Different Phases of the Threat and Extinction Protocol



eFigure 5. A. Upper panel: Threat-induced arousal (TIA) (measured with skin conductance responses to shock delivery) z scores as a predictor of brain activations during threat conditioning and extinction memory retention. TIA correlated negatively with vmPFC activation during threat conditioning (CS+ vs. CS- contrast) and with rostral anterior cingulate cortex (rACC) during extinction memory retention. (B) Composite for Anxiety, Mood, and Personality (C.A.M.P.) z-scores as a predictor of brain activations during threat extinction learning and extinction memory retention. C.A.M.P. scores correlated negatively with vmPFC activation during both threat extinction learning and extinction memory recall, but it predicted higher hippocampal activation during extinction memory retention. Correlations coefficients are indicated for each graph (r). vmPFC = ventromedial prefrontal cortex; rACC = rostral anterior cingulate cortex.