Archival Report

Biological Psychiatry: CNNI

A Multivoxel Pattern Analysis of Anhedonia During Fear Extinction: Implications for Safety Learning

Benjamin M. Rosenberg, Vincent Taschereau-Dumouchel, Hakwan Lau, Katherine S. Young, Robin Nusslock, Richard E. Zinbarg, and Michelle G. Craske

ABSTRACT

BACKGROUND: Pavlovian learning processes are central to the etiology and treatment of anxiety disorders. Anhedonia and related perturbations in reward processes have been implicated in Pavlovian learning. Associations between anhedonia symptoms and neural indices of Pavlovian learning can inform transdiagnostic associations among depressive and anxiety disorders.

METHODS: Participants ages 18 to 19 years (67% female) completed a fear extinction (n = 254) and recall (n = 249) paradigm during functional magnetic resonance imaging. Symptom dimensions of general distress (common to anxiety and depression), fears (more specific to anxiety), and anhedonia-apprehension (more specific to depression) were evaluated. We trained whole-brain multivoxel pattern decoders for anhedonia-apprehension during extinction and extinction recall and tested the decoders' ability to predict anhedonia-apprehension in an external validation sample. Specificity analyses examined effects covarying for general distress and fears. Decoding was repeated within canonical brain networks to highlight candidate neurocircuitry underlying whole-brain effects.

RESULTS: Whole-brain decoder training succeeded during both tasks. Prediction of anhedonia-apprehension in the external validation sample was successful for extinction (R^2 = 0.047; r = 0.276, p = .002) but not extinction recall ($R^2 <$ 0.001, $r = -0.063$, p = .492). The extinction decoder remained significantly associated with anhedonia-apprehension covarying for fears and general distress (t_{121} = 3.209, p = .002). Exploratory results highlighted activity in the cognitive control, default mode, limbic, salience, and visual networks related to these effects.

CONCLUSIONS: Results suggest that patterns of brain activity during extinction, particularly in the cognitive control, default mode, limbic, salience, and visual networks, can be predictive of anhedonia symptoms. Future research should examine associations between anhedonia and extinction, including studies of exposure therapy or positive affect treatments among anhedonic individuals.

https://doi.org/10.1016/j.bpsc.2021.12.008

Anhedonia, the loss of interest or pleasure in activities, is a symptom dimension commonly associated with major depression but also relevant to anxiety disorders. Extant research has focused largely on reward-related processes in relation to anhedonia, such as reductions in sensitivity to reward (i.e., reward consumption, or liking) (1–4), motivation to pursue rewards (i.e., reward anticipation, or wanting), and dopaminergic prediction error signaling associated with impairments in capacity to update behavior after reinforcement learning (5–8). This study extends beyond reward processes to address threat-related processes in relation to anhedonia.

A common paradigm for measuring threat-related processes is Pavlovian fear learning. Activation of the insular cortex, dorsal anterior cingulate cortex (dACC), amygdala (9), and other regions such as the ventromedial prefrontal cortex (PFC) has been consistently highlighted in neuroimaging studies of fear learning (10–15), although there are inconsistencies regarding the precise role of the amygdala in human studies (9,13). Behavioral and neural aberrations in Pavlovian fear acquisition and particularly fear extinction have been observed in individuals at risk for and with anxiety disorders, including perturbations in the insular cortex, dACC, amygdala, and ventromedial PFC (16–18). Beyond the typical threat neurocircuitry, fear extinction has been shown to rely on dopaminergic reward pathways in 1) signaling the unexpected omission of an aversive unconditional stimulus (US) (19,20), or relief, which may itself be considered a type of reward (21), and 2) supporting the long-term consolidation of extinction memories (22). Because anhedonia has been associated with reductions in dopaminergic prediction error signaling, classically evaluated within reward learning paradigms (3,23), there is reason to hypothesize that anhedonia influences reward pathways involved in fear extinction. In partial support, behavioral studies show an association between low positive affect (a central feature of anhedonia) and less stable long-term fear extinction, as measured by stronger reacquisition (24) and reinstatement (25) of conditioned fear.

© 2021 Society of Biological Psychiatry. Published by Elsevier Inc. All rights reserved.

ISSN: 2451-9022

We previously found more direct support for the role of anhedonia in neural responses during fear extinction (9). Specifically, we used a dimensional model of symptoms of anxiety and depression (trilevel model) within a regions-of-interest analytic framework and found that the dimension of anhedonia-apprehension, but not dimensions of general distress or fears, was associated with increased activation of several brain regions during extinction learning, including the insular cortex, dACC, and amygdala (9). Notably, these regions overlap with the salience network, where aberrations as a function of anhedonia have been found in studies of reward consumption, anticipation, and decision making (2,26).

Neural processes associated with anhedonia extend beyond the salience network to regions of the limbic (e.g., ventral striatum and hippocampus) and cognitive control (e.g., orbitofrontal cortex and dorsolateral PFC) networks (26,27). Regions of the default mode network (DMN) (e.g., medial PFC and posterior cingulate cortex) are thought to play a central role in the self-referential processes characteristic of depressive disorders (28–30) and may relate to anhedonia symptoms as well. Given the wide range of brain systems associated with anhedonia, it is conceivable that the influence of anhedonia on fear extinction extend beyond regions of the traditional fear network. This study built on the prior study (9) by 1) analyzing patterns of brain activity during fear extinction and recall to predict individual differences in anhedonia and 2) addressing the breadth of brain activity associated with anhedonia during fear learning, and specifically fear extinction, beyond the fear network.

One approach to neuroimaging data, multivoxel pattern analysis (MVPA), is particularly well suited to research in novel areas and may aid efforts to uncover associations between anhedonia and fear extinction. MVPA uses machine learning to decode patterns of brain activity that are consistently associated with a specific psychological construct. Unique strengths of this approach include its 1) emphasis on distributed patterns of brain activity rather than evaluating individual brain areas separately, 2) ability to directly test these patterns by predicting symptoms in an external validation sample, and 3) flexibility to detect unexpected associations by combining the predictive strengths of different features (which individually may not be strong enough to reach significance). MVPA has been widely used in clinical neuroscience (31,32), including studies identifying patterns of brain activity associated with disruptions in Pavlovian fear learning (33), indices of subjective fear and physiological arousal (34), and anxious compared with nonanxious subjects during Pavlovian fear learning (35).

This study aimed to uncover patterns of brain activity associated with anhedonia by decoding anhedonia symptoms using extinction (n = 254) and extinction recall (n = 249) functional magnetic resonance imaging (fMRI) data collected across two study sites. These tasks were selected due to prior evidence of neural associations with anhedonia within this dataset (9) and known associations between anhedonia and prediction error signaling, a process central to the extinction of learned fear. We hypothesized that the decoders would train successfully during both task phases and that the decoded patterns of brain activity would generalize to an external validation sample (i.e., data that were not included in decoder training). We further hypothesized that successful decoders would be specific to anhedonia-apprehension, over and above other trilevel transdiagnostic symptom factors (i.e., general distress or fears). Exploratory analyses repeated the decoding approach by training and validating the decoder 1) between study sites (i.e., training within data from one site and generalizing to the other) and 2) within individual brain networks, highlighting candidate brain circuits that may be central to decoder results and warrant further research.

METHODS AND MATERIALS

Participants

As described previously (9), participants were recruited for the Brain, Motivation and Personality Development (BrainMAPD) study at the University of California Los Angeles (UCLA) and Northwestern University (NU), which investigated depression and anxiety in late adolescence and early adulthood. Participants were 272 individuals aged 18 to 19 years (182 female; mean age = 19.16 years, SD = 0.52). Recruitment was based on self-reported scores of trait neuroticism (36) and reward sensitivity (37). Oversampling on these dimensions was used to ensure that the sample included individuals at risk for the onset of depression and anxiety (see the Supplement). Exclusion criteria were lack of right-handed dominance, not fluent in English, traumatic brain injury, MRI contraindications, pregnancy, color blindness, lifetime psychotic symptoms, bipolar I disorder, clinically significant substance use disorder in the past 6 months, and antipsychotic medication usage.

Of this group, 254 (UCLA: n = 116, NU: n = 138) were included for fear extinction and 249 (UCLA: n = 116, NU: n = 133) were included for extinction recall (see the Supplement for exclusion details). Of the 254 participants, 250 participants completed the Structured Clinical Interview for DSM-5, of whom 79 participants (31.60%) met criteria for a current anxiety disorder but no depressive disorder, 19 (7.60%) met criteria for current anxiety and depressive disorders, and 3 (1.20%) met criteria for a depressive disorder but no anxiety disorder. Overall, 20 participants (8.00%) reported current use of at least one psychotropic medication (see the Supplement for details). All participants provided written, informed consent. Participant demographics and trilevel symptoms are summarized in Table 1.

Trilevel Measures of General Distress, Fear, and Anhedonia-Apprehension

Immediately before MRI scans, participants completed questionnaire measures of anxiety and depression to generate hierarchical trilevel model factor scores for general distress, fear, and anhedonia-apprehension (see the Supplement for details).

Fear Acquisition, Extinction, and Extinction Recall

The 2-day procedure for fear acquisition, extinction, and extinction recall was based on the validated paradigm developed by Milad *et al.* (38,39). As described previously (9), this slow event-related fMRI paradigm consisted of four phases: habituation, acquisition, extinction (all conducted on day 1), and extinction recall (conducted on day 2, 1–7 days later) (see the Supplement for details). Images were offices or conference rooms (context) with different colored lights (red/yellow/blue)

Table 1. Demographics and Symptom Dimensions

Characteristics	UCLA (n = 116)	Northwestern (n = 138)	Statistic	p Value
Gender, n (%)			$\chi^2_2 = 0.84$	.656
Female, cisgender	78 (67.24%)	92 (66.67%)		
Male, cisgender	38 (32.76%)	45 (32.61%)		
Male, transgender	0 (0.0%)	1 (0.72%)		
Age, Years, Mean (SD)	19.09 (0.52)	19.25 (0.52)	$t_{253} = 2.46$[a]	.014
Ethnicity, n (%)			$\chi^2_1 = 1.23$	.268
Not Hispanic/Latino	82 (70.69%)	106 (76.81%)		
Hispanic/Latino	34 (29.31%)	32 (23.19%)		
Race, n (%)			$\chi^2_5 = 20.67$[a]	.001
Asian	45 (38.79%)	27 (19.57%)		
Black	5 (4.31%)	14 (10.14%)		
Multiracial	3 (2.59%)	17 (12.32%)		
Native American	1 (0.86%)	3 (2.17%)		
White	61 (52.59%)	77 (55.80%)		
Declined to report	1 (0.86%)	0 (0.0%)		
Current Psychotropic Medication Use, n (%)	2 (1.72%)	18 (13.04%)	$\chi^2_1 = 11.13$[a]	.001
Symptom Dimension Scores, Mean (SD)				
General distress	−0.032 (0.94)	0.128 (0.89)	$t_{253} = 1.39$	.17
Fears	0.080 (0.93)	−0.115 (0.79)	$t_{253} = 1.81$	.07
Anhedonia-apprehension	0.112 (0.84)	−0.094 (0.94)	$t_{253} = 1.82$	.07

Demographic factors and symptom dimension scores of participants compared across scanning site. The racial identity of individuals across sites was significantly different, with a higher proportion of Asian participants at UCLA and a higher proportion of Black and Multiracial participants at Northwestern University.

UCLA, University of California Los Angeles.

[a]Denotes statistical significance ($p < .05$).

as conditional stimuli (CS) (color order and context images were counterbalanced across participants). During all task phases, intertrial intervals varied from 12 to 18 s (mean = 15 s) and included a jitter of 125 ms per trial to reduce slice timing bias. The task was programed in E-Prime (version 2.0 SP1) and presented to participants using a mirror and projector system.

fMRI Acquisition and Analysis

We used identical Siemens Prisma 3T MRI scanners at the UCLA Ahmanson-Lovelace Brain Mapping Center and the NU Center for Translational Imaging. High-resolution structural (T1-weighted) images and blood oxygenation level–dependent (T2*-weighted) functional images were acquired and preprocessing procedures applied (see the Supplement for details).

As has been done in prior fMRI studies of fear extinction and extinction recall (35,40,41), analyses specifically focused on the end of fear extinction (i.e., the final four trials for extinguished CS+ minus the final four trials for CS−) and the beginning of extinction recall (i.e., the first four trials for extinguished CS+ minus the first four trials for CS−). Functional images were masked using a standard Montreal Neurological Institute template (42). MVPA was implemented in the scikit-learn toolbox (43) using the ElasticNetCV function (see the Supplement and Table S1 for parameters evaluated during the training stage). Subjects were randomized into training and testing datasets, yielding a training sample of 127 subjects (UCLA: n = 58, NU: n = 69) and a testing sample of 127 subjects (UCLA: n = 58, NU: n = 69) for the extinction task. A total of 5 subjects were not included in the extinction recall analysis, yielding a training sample of 127 subjects (UCLA: n = 58, NU: n = 69) and a testing sample of 122 subjects (UCLA: n = 58, NU: n = 64). Within the testing sample, whole-brain decoders yielded brain-predicted anhedonia-apprehension values for each participant. The coefficient of determination (R^2) and correlation coefficient (r) were calculated between anhedonia-apprehension and brain-predicted anhedonia-apprehension values to determine successful prediction of scores in the external validation sample. To determine the R^2 cutoff score corresponding with statistical significance ($p < .05$), anhedonia-apprehension scores were permuted 10,000 times, and R^2 was computed for each permutation.

Motion Outliers

To account for confounds due to motion, analyses tested the association between the percent of fMRI volumes censored due to motion (see the Supplement) and anhedonia-apprehension and brain-predicted anhedonia-apprehension values in the training sample (covarying for fears, general distress, and site). The percent of volumes censored due to motion was also included as a covariate in specificity analyses (see below).

Specificity Analysis: Associations Over and Above General Distress, Fears, Site, and Motion

Within the external validation sample, the correlation coefficient r was first computed separately between brain-predicted anhedonia-apprehension and each of the trilevel

factors (anhedonia-apprehension, fears, and general distress), covarying for site and motion. The correlation coefficient was then calculated between brain-predicted anhedonia-apprehension and anhedonia-apprehension, covarying for fears, general distress, site, and motion.

Exploratory Analyses: Network-by-Network Effects

To explore localization of the decoder effects, we used the brain atlas developed by Schaefer *et al.* to divide the brain into 100 parcels, grouped into seven functional brain networks (cognitive control, dorsal attention, default mode, limbic, salience, somatomotor, and visual) (44,45). We then reran the decoding procedure seven times, masking within each network. Among significant networks, we additionally reran the decoding procedure masking within each individual region of the network.

Exploratory Analyses: Testing Between-Site Decoding of Anhedonia-Apprehension

To explore the robustness of decoder effects, we reran significant decoders using a between-site external validation approach (see the Supplement). Decoder training was completed within the NU cohort (n = 138), and validation was completed within the UCLA cohort (n = 116). R^2 and r were calculated between anhedonia-apprehension and brain-predicted anhedonia-apprehension values to determine successful prediction of data within the external validation sample. We then reran the decoding procedure seven times, masking within each network, as described above.

RESULTS

Whole-Brain Decoder Effects

Permutation testing yielded a significance cutoff of R^2 = 0.0186 for external validation (corresponding with two-tailed $p < .05$). Initial training of the whole-brain decoder during fear extinction was successful, such that the decoder predicted anhedonia-apprehension values (R^2 = 0.168). The extinction decoder significantly predicted anhedonia-apprehension values in the external validation sample (R^2 = 0.047; r = 0.276, p = .002) (Figure 1).

Initial training of the whole-brain decoder during extinction recall was successful, such that the decoder predicted anhedonia-apprehension values (R^2 = 0.336). However, the extinction recall decoder did not significantly predict anhedonia-apprehension values in the external validation sample ($R^2 < 0.001$, $r = -0.063$, p = .492). Therefore, the extinction recall decoder was not evaluated in subsequent analyses.

Motion Outliers

Covarying for fears, general distress, and site, there was a significant association between percentage of volumes censored and anhedonia-apprehension ($t_{249} = -2.059$, p = .041, r = 0.130), such that individuals with greater anhedonia tended to exhibit less movement in the MRI scanner. Covarying for fears, general distress, and site, there was no association between percentage of volumes censored and brain-predicted anhedonia-apprehension values within the training sample ($t_{122} = -1.277$, p = .204, r = 0.115) or the testing sample (t_{122} = 0.016, p = .987, r = 0.002).

Specificity Analysis: Associations Over and Above General Distress, Fears, Site, and Motion

Covarying for site and motion, brain-predicted anhedonia-apprehension was significantly associated with anhedonia-apprehension (t_{123} = 3.192, p = .002, r = 0.274) but not fears (t_{123} = 0.660, p = .511, r = 0.060) or general distress (t_{123} = 0.338, p = .736, r = 0.031). Covarying for fears, general distress, site, and motion, brain-predicted anhedonia-apprehension was significantly associated with anhedonia-apprehension (t_{121} = 3.209, p = .002, r = 0.277).

Exploratory Analyses: Network-by-Network Effects

Exploratory analyses demonstrated that the following network-masked decoders significantly predicted anhedonia-apprehension within the external validation sample: cognitive control (R^2 = 0.020, r = 0.245, p = .006), default mode (R^2 = 0.040, r = 0.263, p = .003), limbic (R^2 = 0.029, r = 0.217, p = .014), and visual (R^2 = 0.022, r = 0.181, p = .042) (Figure 2, Figure S1, and Table 2 for details). The salience decoder met criteria for significance using r, but not R^2, as the metric of external validation (R^2 = 0.004, r = 0.229, p = .010). Exploratory analyses within these networks further demonstrated that several region-specific decoders could significantly predict anhedonia-apprehension (Table 3 for details).

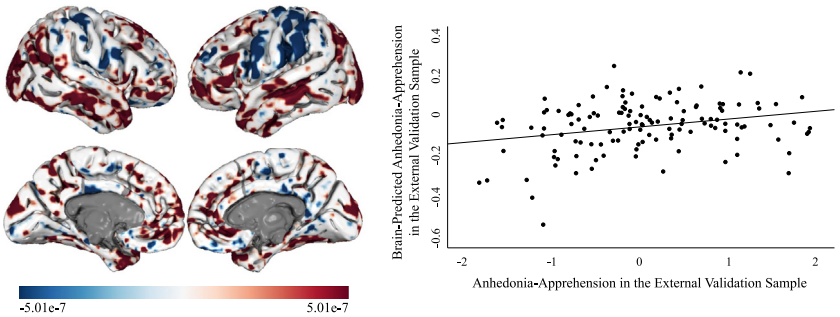

Figure 1. Whole-brain decoder results. Weights of the whole-brain decoder are presented for illustration purposes only. They should not be interpreted as indicating involvement of a specific brain region (as is the case in mass univariate analyses) (1). Predicted anhedonia apprehension values were significantly associated with anhedonia apprehension in the external validation sample.

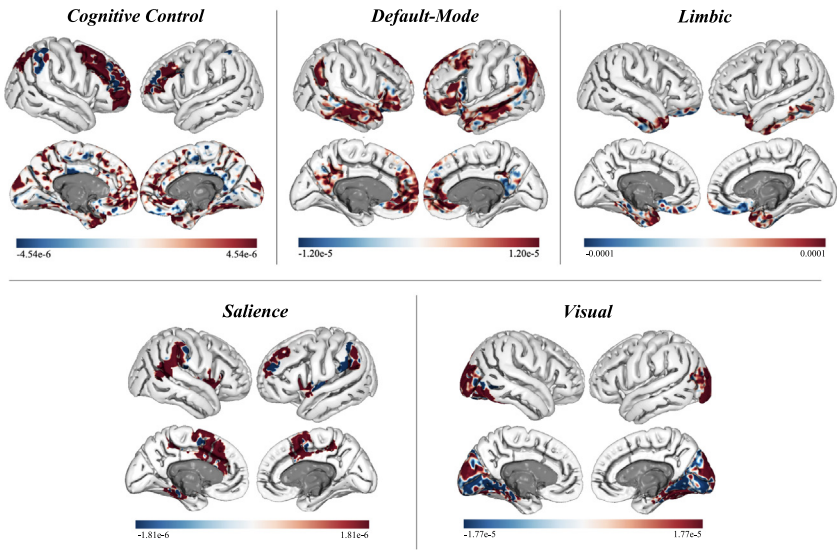

Figure 2. Network-by-network decoder results. Decoder plots for individual network decoders that were significantly associated with anhedonia-apprehension in the external validation sample.

Exploratory Analyses: Decoding of Anhedonia-Apprehension With a Between-Sites Approach

Initial training of the whole-brain decoder during fear extinction was successful (R^2 = 0.420) (see the Supplement for details). The between-sites decoder met criteria for significance using r, but not R^2, as the metric of external validation (R^2 = −0.024, r = 0.190, p = .040). Specificity analyses revealed that using the between-sites approach, the dorsal attention (R^2 = −0.025, r = 0.186, p = .045) and visual (R^2 = −0.015, r = 0.236, p = .011) network decoders met criteria for significance using r, but not R^2, as the metric of external validation (see the Supplement for details).

DISCUSSION

This study used MVPA to characterize unique patterns of functional brain activity during fear extinction and extinction recall associated with anhedonia symptoms. We found anhedonia-specific whole-brain patterns of functional activity during fear extinction that generalized to an external validation sample. These patterns were significantly associated with the dimension of anhedonia-apprehension over and above other symptom dimensions of general distress and fears.

Within individual networks and regions, the patterns of activity appeared complex. Although plotting the decoder weights can aid in the interpretation of which regions and networks are implicated in the whole-brain decoder, these results should be interpreted with caution. For example, high beta weights could indicate voxels that cancel out noise rather than increased activation. Similarly, if two voxels provide an equivalent amount of information, the decoder may arbitrarily select one voxel and omit the other [for additional information on interpreting decoder results, see (46)]. Hence, exploration of specific brain networks implicated in the anhedonia decoder is highly tentative. With that caveat in mind, we identified activity within the cognitive control, default mode, limbic, salience, and visual networks that generalized across the training and external validation samples.

The anhedonia-apprehension decoder appeared to involve predominantly positive beta weights among regions of the salience network. Regions of the salience network, such as the insular cortex, dACC, and amygdala, overlap with the fear

Table 2. Network-by-Network Decoder Results

	Decoder Training	Testing in External Validation Sample		
Network Mask	R^2	R^2	r	p
Cognitive Control	0.045	0.020[a]	0.245[a]	.006
Default Mode	0.088	0.040[a]	0.263[a]	.003
Dorsal Attention	0.271	0.008	0.140	.116
Limbic	0.183	0.029[a]	0.217[a]	.014
Salience	0.020	0.004	0.229[a]	.010
Somatomotor	Failed training	N/A	N/A	N/A
Visual	0.155	0.022[a]	0.181[a]	.042

Exploratory network-by-network decoder results implicated in the whole-brain decoder.
N/A, not applicable.
[a]Denotes statistical significance ($R^2 > 0.0186$ or $p < .05$).

Table 3. Region-by-Region Decoder Results

Regional Masks	Schaefer Atlas ROI Index	Testing in External Validation Sample		
		R^2	r	p
Cognitive Control Network				
L dorsolateral PFC, lateral	35	0.016	0.267[a]	.002
R dorsolateral PFC, anterior	83	0.016	0.215[a]	.015
R dorsolateral PFC, lateral	84	0.017	0.249[a]	.005
R dorsolateral PFC, dorsal	85	0.026[a]	0.227[a]	.010
R frontal eye field	86	0.034[a]	0.250[a]	.005
R medial posterior PFC/frontal eye field	88	0.018	0.222[a]	.012
Default Mode Network				
L medial temporal gyrus	38	0.013	0.177[a]	.046
L medial temporal gyrus	39	0.057[a]	0.299[a]	.001
L angular gyrus	41	0.028[a]	0.210[a]	.018
L pars orbitalis	42	0.014	0.290[a]	.001
L pars orbitalis	43	0.044[a]	0.259[a]	.003
L dorsal anterior cingulate cortex	44	0.028[a]	0.215[a]	.015
L anterior PFC	45	0.028[a]	0.293[a]	.001
L dorsolateral PFC, dorsal	46	0.025[a]	0.200[a]	.024
L premotor/supplementary motor area	47	0.023[a]	0.224[a]	.011
L frontal eye field	48	0.039[a]	0.295[a]	.001
L ventral posterior cingulate cortex	49	-0.007	0.225[a]	.011
L ventral posterior cingulate cortex	50	0.021[a]	0.229[a]	.010
R superior temporal gyrus	93	0.011	0.237[a]	.007
R pars orbitalis	94	0.026[a]	0.285[a]	.001
R Broca's triangle	95	0.036[a]	0.274[a]	.002
R anterior PFC	96	0.032[a]	0.222[a]	.012
R dorsolateral PFC, dorsal	97	0.020[a]	0.178[a]	.045
R frontal eye field	98	0.047[a]	0.295[a]	.001
R ventral posterior cingulate cortex	100	0.001	0.194[a]	.029
Limbic Network				
L orbitofrontal cortex	31	0.014	0.217[a]	.014
L temporal pole	32	0.037[a]	0.256[a]	.004
R orbitofrontal cortex	79	0.010	0.211[a]	.018
Salience Network				
L insula/frontal operculum	25	-0.001	0.288[a]	.001
L insula/frontal operculum	26	0.001	0.225[a]	.011
L anterior lateral PFC	27	0.031[a]	0.269[a]	.002
L dorsal anterior cingulate cortex	28	0.024[a]	0.221[a]	.013
L premotor/supplementary motor area	30	0.007	0.191[a]	.032
R insula/frontal operculum	76	0.009	0.255[a]	.004
Visual Network				
L visual association area	2	0.018	0.175[a]	.049
L visual association area	3	0.021[a]	0.272[a]	.002
L visual association area	7	0.013	0.214[a]	.016
L visual association area	8	0.032[a]	0.209[a]	.018
R fusiform gyrus	51	0.022[a]	0.227[a]	.010
R fusiform gyrus	52	0.023[a]	0.194[a]	.029
R primary visual cortex	55	0.016	0.194[a]	.029
R visual association area	57	0.028[a]	0.197[a]	.026

Significant exploratory ROI-by-ROI decoder results within networks implicated in the whole-brain decoder.
L, left hemisphere, PFC, prefrontal cortex; R, right hemisphere; ROI, region of interest.
[a]Denotes statistical significance ($R^2 > 0.0186$ or $p < .05$).

network (9,13). One interpretation of heightened activation in the salience network is persistent attentional salience of extinguished stimuli, perhaps representing strength of CS-US associations (i.e., weakened extinction). Additional research is needed to explore this possibility particularly considering the limited interpretability of directional findings in MVPA analyses. Likewise, it has been suggested that the salience network integrates information from both the default mode and cognitive control networks in directing external and internal attention (47–49) and that this process is altered in major depression (50,51). Additional research may explore the extent to which persistent activation among regions of the cognitive control network or DMN, in coordination with the salience network, relates to deficits in extinction associated with anhedonia.

Another potential pattern was for the anhedonia-apprehension decoder to involve predominantly heightened activity within the cognitive control network. Prior studies of major depression have highlighted aberrant activity in this network, particularly the dorsolateral PFC, in studies of attentional bias and emotion regulation (52–54). Heightened dorsolateral PFC activation has been implicated as compensation for diminished reward processing as a function of depression (26) and transdiagnostically (55) and thus offers another pathway for the relationship between anhedonia and extinction.

Furthermore, the anhedonia-apprehension decoder appeared to involve predominantly positive beta weights among regions of the DMN. Hyperactivation and hyperconnectivity of the DMN have been implicated in studies of depression, particularly during unconstrained rest or during tasks involving internally directed attention, such as autobiographical memory and rumination (28–30,51). These results support the potential applicability of the DMN within studies of anhedonia or extinction learning, although additional research is needed to elucidate these associations more precisely.

These findings highlight the role of anhedonia in relation to fear learning constructs that have been traditionally considered primarily within the context of anxiety disorders. Although anhedonia has been considered mostly within the context of depression, it is transdiagnostic and associated with several anxiety disorders, including social anxiety (56), obsessive-compulsive disorder (57), and posttraumatic stress (58). Greater recognition of the role of anhedonia in anxiety disorders and fear learning processes is consistent with dimensional models of psychopathology that cut across conventional diagnostic categories (59–66) and have direct implications for optimal care (67,68). Pending replication of these findings, the role of anhedonia in fear learning could be leveraged in the development of personalized, process-targeted treatments. For example, studies of fear extinction have provided a foundation for contemporary models of exposure therapy (69–71), which emphasize prediction error (and other features, such as contextual modulation) for optimizing exposure therapy effectiveness (18,72–74). Given the potential interference with prediction error posed by anhedonia, novel exposure protocols may incorporate interventions to increase positive affect already shown to augment extinction (75,76) for anxious individuals with anhedonic symptoms.

In addition, neuromodulation targeting the control, default mode, or salience networks may augment exposure therapy for individuals with anhedonia. For example, preliminary studies of transcranial magnetic stimulation have targeted the dorsolateral PFC to augment the effects of exposure therapy for posttraumatic stress disorder (77–79). Combining brain stimulation and exposure therapy may prove particularly useful for patients with anxiety who present with elevated anhedonia or a comorbid depressive disorder.

Decoder cross-validation was associated with relatively small coefficient of determination (R^2) values in this study. Despite the potential advantages of R^2 in prediction studies (80), the application of R^2 in MVPA studies may also be limited due to scaling issues. For example, this study collected data at two different fMRI scanners, which could affect the R^2 metric. For this reason, we have also reported correlation coefficients (r), which are relatively independent of the scale used and tended to indicate stronger associations between anhedonia and brain activation.

This study involves several strengths, namely the 1) comparatively large sample size of both the training and external validation datasets, 2) emphasis on dimensional psychopathology, 3) test of effects in an external validation sample, and 4) exploration of effects using a between-sites approach. The narrow age range of participants could also be considered a strength of this study because the reported effects are unlikely to be explained by variations in the age of participants. However, the narrow age range may also reduce generalizability of these results to other developmental stages. Additional limitations include the 1) small number of experimental trials analyzed in the extinction and extinction recall tasks; 2) exploratory nature of network- and region-specific analyses, including some cases in which r but not R^2 met criteria for statistical significance; 3) comparatively small size of decoder prediction values; and 4) tentative interpretability of directional results. Future research is needed to replicate the findings of this study, to explore avenues for strengthening decoder predictions, and to evaluate the directionality of network- and region-specific results.

In sum, this study suggests that patterns of brain activity during extinction learning are predictive of anhedonia symptoms. Extinction is a fear learning process traditionally considered in relation to anxiety symptoms but rarely in relation to transdiagnostic symptom dimensions, such as anhedonia. The patterns of brain activity identified in this study may be characteristic of anhedonia-specific deficits during fear learning and warrant additional research.

ACKNOWLEDGMENTS AND DISCLOSURES

This research was supported by the National Institute of Mental Health (NIMH) (Grant No. R01 MH100117-01 [to MGC, RN, and REZ]). BMR was supported by the National Science Foundation Graduate Research Fellowship (Grant No. DGE-1650604) during the completion of this work.

The content is the responsibility of the authors and does not necessarily represent the official views of the NIMH.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Department of Psychology (BMR, MGC), College of Life Sciences, and Department of Psychiatry and Biobehavioral Sciences (MGC), University of California, Los Angeles, Los Angeles, California; Department of Psychology (RN, REZ), Institute for Policy Research (RN), and Family Institute at Northwestern University (REZ), Northwestern University, Evanston, Illinois; Department of Psychiatry and Addictology (VT-D), University of Montréal, Montreal; Centre de Recherche de l'Institut Universitaire en Santé Mentale de Montréal (VT-D), Montreal, Quebec, Canada; Social, Genetic and Development Psychiatry Centre (KSY), Institute of Psychology, Psychiatry and Neuroscience, and National Institute for Health Research Maudsley Biomedical Research Centre (KSY), King's College London, London, United Kingdom; and RIKEN Center for Brain Science (HL), Saitama, Japan.

Address correspondence to Benjamin M. Rosenberg, M.A., C.Phil., at ben.rosenberg@g.ucla.edu.

Received Sep 8, 2021; revised Dec 1, 2021; accepted Dec 9, 2021.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.bpsc.2021.12.008.

REFERENCES

1. Berridge KC, Robinson TE (2003): Parsing reward [published correction appears in Trends Neurosci 2003; 26:581]. Trends Neurosci 26:507–513.
2. Treadway MT, Zald DH (2011): Reconsidering anhedonia in depression: Lessons from translational neuroscience. Neurosci Biobehav Rev 35:537–555.
3. Huys QJ, Pizzagalli DA, Bogdan R, Dayan P (2013): Mapping anhedonia onto reinforcement learning: A behavioural meta-analysis. Biol Mood Anxiety Disord 3:12.
4. Thomsen KR (2015): Measuring anhedonia: Impaired ability to pursue, experience, and learn about reward. Front Psychol 6:1409.
5. Kumar P, Waiter G, Ahearn T, Milders M, Reid I, Steele JD (2008): Abnormal temporal difference reward-learning signals in major depression. Brain 131:2084–2093.
6. Pizzagalli DA, Iosifescu D, Hallett LA, Ratner KG, Fava M (2008): Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. J Psychiatr Res 43:76–87.
7. Gradin VB, Kumar P, Waiter G, Ahearn T, Stickle C, Milders M, *et al.* (2011): Expected value and prediction error abnormalities in depression and schizophrenia. Brain 134:1751–1764.
8. Pizzagalli DA (2014): Depression, stress, and anhedonia: Toward a synthesis and integrated model. Annu Rev Clin Psychol 10:393–423.
9. Young KS, Bookheimer SY, Nusslock R, Zinbarg RE, Damme KSF, Chat IK, *et al.* (2021): Dysregulation of threat neurocircuitry during fear extinction: The role of anhedonia. Neuropsychopharmacology 46:1650–1657.
10. Etkin A, Wager TD (2007): Functional neuroimaging of anxiety: A meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry 164:1476–1488.
11. Sehlmeyer C, Schöning S, Zwitserlood P, Pfleiderer B, Kircher T, Arolt V, Konrad C (2009): Human fear conditioning and extinction in neuroimaging: A systematic review. PLoS One 4:e5865.
12. Milad MR, Rosenbaum BL, Simon NM (2014): Neuroscience of fear extinction: Implications for assessment and treatment of fear-based and anxiety related disorders. Behav Res Ther 62:17–23.
13. Fullana MA, Harrison BJ, Soriano-Mas C, Vervliet B, Cardoner N, Àvila-Parcet A, Radua J (2016): Neural signatures of human fear conditioning: An updated and extended meta-analysis of fMRI studies. Mol Psychiatry 21:500–508.
14. Greco JA, Liberzon I (2016): Neuroimaging of fear-associated learning. Neuropsychopharmacology 41:320–334.
15. Battaglia S, Garofalo S, di Pellegrino G, Starita F (2020): Revaluing the role of vmPFC in the acquisition of Pavlovian threat conditioning in humans. J Neurosci 40:8491–8500.
16. Graham BM, Milad MR (2011): The study of fear extinction: Implications for anxiety disorders. Am J Psychiatry 168:1255–1265.
17. Craske MG, Stein MB, Eley TC, Milad MR, Holmes A, Rapee RM, Wittchen HU (2017): Anxiety disorders [published correction appears in Nat Rev Dis Primers 2017; 3:17100]. Nat Rev Dis Primers 3:17024.
18. Craske MG, Hermans D, Vervliet B (2018): State-of-the-art and future directions for extinction as a translational model for fear and anxiety [published correction appears in Philos Trans R Soc Lond B Biol Sci 2018; 373:20180432]. Philos Trans R Soc Lond B Biol Sci 373:20170025.
19. Salinas-Hernández XI, Vogel P, Betz S, Kalisch R, Sigurdsson T, Duvarci S (2018): Dopamine neurons drive fear extinction learning by signaling the omission of expected aversive outcomes. Elife 7:e38818.
20. Papalini S, Beckers T, Vervliet B (2020): Dopamine: From prediction error to psychotherapy. Transl Psychiatry 10:164.
21. Carver CS (2009): Threat sensitivity, incentive sensitivity, and the experience of relief. J Pers 77:125–138.
22. Kalisch R, Gerlicher AMV, Duvarci S (2019): A dopaminergic basis for fear extinction. Trends Cogn Sci 23:274–277.
23. Rizvi SJ, Pizzagalli DA, Sproule BA, Kennedy SH (2016): Assessing anhedonia in depression: Potentials and pitfalls. Neurosci Biobehav Rev 65:21–35.
24. Zbozinek TD, Craske MG (2017): Positive affect predicts less reacquisition of fear: Relevance for long-term outcomes of exposure therapy. Cogn Emot 31:712–725.
25. Zbozinek TD, Holmes EA, Craske MG (2015): The effect of positive mood induction on reducing reinstatement fear: Relevance for long term outcomes of exposure therapy. Behav Res Ther 71:65–75.
26. Zhang WN, Chang SH, Guo LY, Zhang KL, Wang J (2013): The neural correlates of reward-related processing in major depressive disorder: A meta-analysis of functional magnetic resonance imaging studies. J Affect Disord 151:531–539.
27. Drevets WC (2007): Orbitofrontal cortex function and structure in depression. Ann N Y Acad Sci 1121:499–527.
28. Sheline YI, Barch DM, Price JL, Rundle MM, Vaishnavi SN, Snyder AZ, *et al.* (2009): The default mode network and self-referential processes in depression. Proc Natl Acad Sci U S A 106:1942–1947.
29. Hamilton JP, Farmer M, Fogelman P, Gotlib IH (2015): Depressive rumination, the default-mode network, and the dark matter of clinical neuroscience. Biol Psychiatry 78:224–230.
30. Zhou HX, Chen X, Shen YQ, Li L, Chen NX, Zhu ZC, *et al.* (2020): Rumination and the default mode network: Meta-analysis of brain imaging studies and implications for depression. Neuroimage 206:116287.
31. Dwyer DB, Falkai P, Koutsouleris N (2018): Machine learning approaches for clinical psychology and psychiatry. Annu Rev Clin Psychol 14:91–118.
32. Zhang X, Braun U, Tost H, Bassett DS (2020): Data-driven approaches to neuroimaging analysis to enhance psychiatric diagnosis and therapy. Biol Psychiatry Cogn Neurosci Neuroimaging 5:780–790.
33. Hennings AC, McClay M, Lewis-Peacock JA, Dunsmoor JE (2020): Contextual reinstatement promotes extinction generalization in healthy adults but not PTSD. Neuropsychologia 147:107573.
34. Taschereau-Dumouchel V, Kawato M, Lau H (2020): Multivoxel pattern analysis reveals dissociations between subjective fear and its physiological correlates. Mol Psychiatry 25:2342–2354.
35. Wen Z, Marin MF, Blackford JU, Chen ZS, Milad MR (2021): Fear-induced brain activations distinguish anxious and trauma-exposed brains. Transl Psychiatry 11:46.
36. Eysenck HJ, Eysenck SBG (1975): Manual of the Eysenck Personality Questionnaire (Junior and Adult). London: Hodder & Stoughton.
37. Carver CS, White TL (1994): Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. J Pers Soc Psychol 67:319–333.
38. Milad MR, Quirk GJ, Pitman RK, Orr SP, Fischl B, Rauch SL (2007): A role for the human dorsal anterior cingulate cortex in fear expression. Biol Psychiatry 62:1191–1194.
39. Milad MR, Wright CI, Orr SP, Pitman RK, Quirk GJ, Rauch SL (2007): Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. Biol Psychiatry 62:446–454.
40. Marin MF, Zsido RG, Song H, Lasko NB, Killgore WDS, Rauch SL, *et al.* (2017): Skin conductance responses and neural activations

during fear conditioning and extinction recall across anxiety disorders. JAMA Psychiatry 74:622–631.
41. Marin MF, Hammoud MZ, Klumpp H, Simon NM, Milad MR (2020): Multimodal categorical and dimensional approaches to understanding threat conditioning and its extinction in individuals with anxiety disorders. JAMA Psychiatry 77:618–627.
42. Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, *et al.* (2001): A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos Trans R Soc Lond B Biol Sci 356:1293–1322.
43. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, *et al.* (2011): Scikit-learn: Machine learning in Python. J Mach Learn Res 12:2825–2830.
44. Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, *et al.* (2011): The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165.
45. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, *et al.* (2018): Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb Cortex 28:3095–3114.
46. Kriegeskorte N, Douglas PK (2019): Interpreting encoding and decoding models. Curr Opin Neurobiol 55:167–179.
47. Sridharan D, Levitin DJ, Menon V (2008): A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proc Natl Acad Sci U S A 105:12569–12574.
48. Menon V, Uddin LQ (2010): Saliency, switching, attention and control: A network model of insula function. Brain Struct Funct 214:655–667.
49. Goulden N, Khusnulina A, Davis NJ, Bracewell RM, Bokde AL, McNulty JP, Mullins PG (2014): The salience network is responsible for switching between the default mode network and the central executive network: Replication from DCM. Neuroimage 99:180–190.
50. Hamilton JP, Furman DJ, Chang C, Thomason ME, Dennis E, Gotlib IH (2011): Default-mode and task-positive network activity in major depressive disorder: Implications for adaptive and maladaptive rumination. Biol Psychiatry 70:327–333.
51. Kaiser RH, Andrews-Hanna JR, Wager TD, Pizzagalli DA (2015): Large-scale network dysfunction in major depressive disorder: A meta-analysis of resting-state functional connectivity. JAMA Psychiatry 72:603–611.
52. Gotlib IH, Hamilton JP (2008): Neuroimaging and depression: Current status and unresolved issues. Curr Dir Psychol Sci 17:159–163.
53. Koenigs M, Grafman J (2009): The functional neuroanatomy of depression: Distinct roles for ventromedial and dorsolateral prefrontal cortex. Behav Brain Res 201:239–243.
54. Disner SG, Beevers CG, Haigh EAP, Beck AT (2011): Neural mechanisms of the cognitive model of depression. Nat Rev Neurosci 12:467–477.
55. Brolsma SCA, Vassena E, Vrijsen JN, Sescousse G, Collard RM, van Eijndhoven PF, *et al.* (2021): Negative learning bias in depression revisited: Enhanced neural response to surprising reward across psychiatric disorders. Biol Psychiatry Cogn Neurosci Neuroimaging 6:280–289.
56. Kashdan TB (2007): Social anxiety spectrum and diminished positive experiences: Theoretical synthesis and meta-analysis. Clin Psychol Rev 27:348–365.
57. Abramovitch A, Pizzagalli DA, Reuman L, Wilhelm S (2014): Anhedonia in obsessive-compulsive disorder: Beyond comorbid depression. Psychiatry Res 216:223–229.
58. Nawijn L, van Zuiden M, Frijling JL, Koch SBJ, Veltman DJ, Olff M (2015): Reward functioning in PTSD: A systematic review exploring the mechanisms underlying anhedonia. Neurosci Biobehav Rev 51:189–204.
59. Krueger RF, Kotov R, Watson D, Forbes MK, Eaton NR, Ruggero CJ, *et al.* (2018): Progress in achieving quantitative classification of psychopathology. World Psychiatry 17:282–293.
60. Naragon-Gainey K, Prenoveau JM, Brown TA, Zinbarg RE (2016): A comparison and integration of structural models of depression and anxiety in a clinical sample: Support for and validation of the tri-level model. J Abnorm Psychol 125:853–867.
61. Prenoveau JM, Zinbarg RE, Craske MG, Mineka S, Griffith JW, Epstein AM (2010): Testing a hierarchical model of anxiety and depression in adolescents: A tri-level model. J Anxiety Disord 24:334–344.
62. Brown TA, Barlow DH (2009): A proposal for a dimensional classification system based on the shared features of the DSM-IV anxiety and mood disorders: Implications for assessment and treatment. Psychol Assess 21:256–271.
63. Insel TR (2014): The NIMH Research Domain Criteria (RDoC) project: Precision medicine for psychiatry. Am J Psychiatry 171:395–397.
64. Thompson-Hollands J, Sauer-Zavala S, Barlow DH (2014): CBT and the future of personalized treatment: A proposal. Depress Anxiety 31:909–911.
65. Kotov R, Krueger RF, Watson D, Achenbach TM, Althoff RR, Bagby RM, *et al.* (2017): The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. J Abnorm Psychol 126:454–477.
66. Kotov R, Krueger RF, Watson D (2018): A paradigm shift in psychiatric classification: The Hierarchical Taxonomy of Psychopathology (HiTOP). World Psychiatry 17:24–25.
67. Ruggero CJ, Kotov R, Hopwood CJ, First M, Clark LA, Skodol AE, *et al.* (2019): Integrating the Hierarchical Taxonomy of Psychopathology (HiTOP) into clinical practice. J Consult Clin Psychol 87:1069–1084.
68. Hopwood CJ, Bagby RM, Gralnick T, Ro E, Ruggero C, Mullins-Sweatt S, *et al.* (2020): Integrating psychotherapy with the hierarchical taxonomy of psychopathology (HiTOP). J Psychother Integr 30:477–497.
69. Pittig A, van den Berg L, Vervliet B (2016): The key role of extinction learning in anxiety disorders: Behavioral strategies to enhance exposure-based treatments. Curr Opin Psychiatry 29:39–47.
70. Scheveneels S, Boddez Y, Vervliet B, Hermans D (2016): The validity of laboratory-based treatment research: Bridging the gap between fear extinction and exposure treatment. Behav Res Ther 86:87–94.
71. Fullana MA, Dunsmoor JE, Schruers KRJ, Savage HS, Bach DR, Harrison BJ (2020): Human fear conditioning: From neuroscience to the clinic. Behav Res Ther 124:103528.
72. Craske MG, Kircanski K, Zelikowsky M, Mystkowski J, Chowdhury N, Baker A (2008): Optimizing inhibitory learning during exposure therapy. Behav Res Ther 46:5–27.
73. Craske MG, Liao B, Brown L, Vervliet B (2012): Role of inhibition in exposure therapy. J Exp Psychopathol 3:322–345.
74. Craske MG, Treanor M, Conway CC, Zbozinek T, Vervliet B (2014): Maximizing exposure therapy: An inhibitory learning approach. Behav Res Ther 58:10–23.
75. Craske MG, Meuret AE, Ritz T, Treanor M, Dour HJ (2016): Treatment for anhedonia: A neuroscience driven approach. Depress Anxiety 33:927–938.
76. Craske MG, Meuret AE, Ritz T, Treanor M, Dour H, Rosenfield D (2019): Positive affect treatment for depression and anxiety: A randomized clinical trial for a core feature of anhedonia. J Consult Clin Psychol 87:457–471.
77. Osuch EA, Benson BE, Luckenbaugh DA, Geraci M, Post RM, McCann U (2009): Repetitive TMS combined with exposure therapy for PTSD: A preliminary study. J Anxiety Disord 23:54–59.
78. Karsen EF, Watts BV, Holtzheimer PE (2014): Review of the effectiveness of transcranial magnetic stimulation for post-traumatic stress disorder. Brain Stimul 7:151–157.
79. Fryml LD, Pelic CG, Acierno R, Tuerk P, Yoder M, Borckardt JJ, *et al.* (2019): Exposure therapy and simultaneous repetitive transcranial magnetic stimulation: A controlled pilot trial for the treatment of posttraumatic stress disorder. J ECT 35:53–60.
80. Poldrack RA, Huckins G, Varoquaux G (2020): Establishment of best practices for evidence for prediction: A review. JAMA Psychiatry 77:534–540.