Learning about location-dependent threat: neural abnormalities in clinical anxiety

Benjamin Suarez-Jimenez (benjamin.suarez@nyspi.columbia.edu)
Columbia University Medical Center https://orcid.org/0000-0002-4765-2458
Nicholas Balderston
University of Pennsylvania https://orcid.org/0000-0002-8565-1544
James Bisby
University College London https://orcid.org/0000-0002-5334-3969
Joseph Leshin
The University of North Carolina
Abigail Hsiung
Duke University
John King
University College London
Daniel (S) Pine
NIMH, Bethesda, MD
Neil Burgess
University College London https://orcid.org/0000-0003-0646-6584
Christian Grillon
NIMH
Monique Ernst
National Institute of Mental Health

Article

Keywords:

Posted Date: August 31st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-61318/v1

Version of Record: A version of this preprint was published at Communications Biology on November 4th, 2021. See the published version at https://doi.org/10.1038/s42003-021-02775-x.

Abstract

Anxiety disorders are characterized by maladaptive defensive responses to distal or uncertain threats. Elucidating neural mechanisms of anxiety is essential to formulate new treatment strategies tartgeting these circuits. In an fMRI scanner, patients with pathological anxiety (ANX, n=23) and healthy comparisons (HC, n=28) completed a contextual threat learning paradigm, in which they picked flowers in a virtual environment comprising a danger zone in which flowers were paired with shock and a safe zone (no shock). ANX compared with HC showed 1) global decreased ventromedial prefrontal cortex and anterior hippocampus activation during the task, 2) increased insula and dorsomedial prefrontal cortex activation in the danger zone during the task; and 3) increased amygdala and midbrain/periaqueductal gray activation in the danger zone prior to potential shock delivery. Our findings suggest that ANX exhibit a heightened reactivity to threat and show decreased activation in modulatory areas responsible for regulating context-appropriate emotional responses.

Introduction

When exploring our environment, we might encounter items that require us to learn about their threat value. Learning about potential threatening environments may induce anxiety, an anticipatory response to potential threats, and a fearful response evoked by an imminent acute threat.[1–5] Although anxious and fearful states are normal responses to threats chronic manifestations of these states can be highly debilitating.[6,7] Research shows that patients with chronic anxiety lack the ability to integrate contextual cues to guide learning of threat and safety.[9–11] A previous investigation has delineated the neural mechanisms underlying learning and discriminating threats within specific spatial locations in healthy adults.[8] However, very little is known about how patients with pathological anxiety learn about threat within an environment. Understanding how patients with pathological anxiety learn about threat within specific spatial location is essential to better understand the disorder and create novel treatments.

Traditional context conditioning paradigms have shown that a defensive response can be triggered not only by the presentation of an aversive stimulus but also by the presentation of the context where an aversive stimulus was previously encountered. Typically, healthy individuals can learn to distinguish between a safe and dangerous context. That is, when healthy individuals associate a cue with an aversive stimulus (conditioned stimulus; CS) in context A (CS+A), they display a threat response that is dependent on the context in which the association was made. However, the same cue (CS) in a different context B (CS-B) elicits a weaker defensive response. Context conditioning requires the use of spatial processing strategies[9–11] (i.e., attend to the environment and surrounding landmarks to create a spatial representation of where a threat was encountered). These strategies have been mapped to neural systems that regulate emotion and memory, such as the hippocampus, amygdala, and prefrontal cortex (PFC).[11-21]

We developed a virtual-environment paradigm to probe how brain regions interact to shape behavior (i.e. threat learning and discrimination) over time[8] to quantify threat learning relevant to context-specific

threat. The virtual environment depicts a grass field surrounded by mountains, divided equally into two zones, a safe and a danger zone. In both zones, flowers appear and need to be “picked” up. Picking flowers in the danger zone potentially causes an electric shock to the wrist (or “bee-sting”), while flowers in the safe zone are never associated with shock. To learn threat contingencies, participants must rely on distal environmental cues (e.g., shape of the mountains and clouds, which differ in both zones, and beehives) to locate themselves and learn ‘where’ they are in the environment and not on the physical properties of the stimuli (i.e., the flowers), which are all identical in both zones.[8]

Previous findings with this paradigm in healthy volunteers informed three processes: threat learning, threat appraisal (anxiety-state), and threat anticipation (fear-state). (1) Threat learning: Healthy adults demonstrated behavioral/physiological learning: as the task progressed, shock expectancy ratings and skin conductance for flowers increased in the danger but not the safe zone. The neural substrates associated with learning about environmental threats during flower approach, in either zone, engaged key nodes of the learning circuit, including the anterior hippocampus, amygdala, ventromedial prefrontal cortex (vmPFC), and vmPFC-hippocampal functional connectivity. (2) Threat appraisal: approach of the flower, in the danger zone compared to the safe zone, recruited sensory and control-related regions, i.e., the insula, dorsal anterior cingulate cortex (dACC), extending to the dorsomedial prefrontal cortex (dmPFC), and insula-hippocampal functional connectivity. (3) Threat anticipation: During imminent threat of potential shock, upon picking a flower, a progressively increasing response of the periaqueductal gray matter (PAG) and posterior hippocampus, and insula-dACC coupling were observed.[8]

Research has shown that individuals with anxiety disorders display a higher defensive response to a CS+, compared to healthy individuals, and they often generalize the threat response to safe cues (CS-),[6,7,12,14] but this not clear if this extends to context. Still, hippocampal dysfunction and decreased hippocampal volume have been associated with anxiety disorders.[12,13,15] For example, studies in both humans and rats suggest that impairment in hippocampal function leads to compensatory learning strategies that do not involve the hippocampus. These abnormal modulation of attention, linked to attention shifts to the cue and not the context, use compensatory neural mechanisms that leads to generalization of threat.[6,7,10–12,14,16–20,20,21]

Using the virtual environment paradigm described above, three main hypotheses are tested. Individuals with an anxiety disorder (ANX; generalized anxiety disorder, social anxiety disorder) compared with a sample of healthy comparisons (HC) will show: (1) Threat learning: poor learning and discrimination associated with compensatory learning strategies that do not involve the hippocampus leading to generalization of threat, (2) Threat appraisal: stronger engagement of the dACC, dmPFC, amygdala, and insula in danger zones when approaching the flowers (higher anxiety-state), and (3) Threat anticipation: stronger engagement of the periaqueductal gray matter (PAG) activation in danger zones in anticipation of potential shock delivery (higher fear-state).

Results

Physiological and Behavioral Measures of Threat Learning

Participants explored a virtual environment (**Fig. 1A and B**; see Methods for details) that consisted of a mountain landscape and defined two half-zones recognizable by the unique shape of the mountains in the horizon. For each trial, participants freely explored the arena and were instructed to pick up flowers that appeared one at a time in random locations across the environment (approach period). When a participant picked a flower, their position was held stationary for a variable duration (2-8 seconds; stationary period), during which the participant rated the expectancy of receiving a shock (rating of 0-9). Flowers located in one-half of the environment were paired with a shock delivered at the end of the stationary period on 50% of the trials within the danger zone. Flowers in the other half of the environment were never paired with a shock (safe zone). Since all flowers were identical, predictive value (danger or safety) could not be attributed to their physical characteristics.

For analysis, the data were divided by zones (safe, danger) and segregated into 4 learning blocks (10 trials in each). All these physiological and behavioral analyses can be found in the supplementary material.

Skin conductance. We measured participants' skin conductance level (SCL), as they navigated towards the flower (approach period), and skin conductance responses (SCR), during the stationary period. A 2x4x2 ANOVA (zone x block x group) of both SCL and SCR revealed a significant main effect of zone, due to higher responses in the danger zone as compared to the safety zone (SCL, $F(1,49)=16.24$, $p<0.01$; SCR, $F(1,49)=4.13$, $p<0.05$; For a full description of skin conductance results, see supplementary material and **Figure S1**).

Shock expectancy. A 2x4x2 ANOVA (zone x block x group) assessing shock expectancy ratings after picking the flower revealed a significant zone x block interaction ($F(3,147)=31.07$, $p<0.001$). This interaction stemmed from consistently (across the 4 blocks) high expectancy ratings of the shock in the danger zone (danger vs safe: $t(50)=0.06$, $p>0.05$), while ratings significantly decreased between block 1 and 4 in the safe zone suggesting that there was discrimination learning over blocks for all subjects (block 1 vs 4: $t(50)=8.83$, $p<0.001$; For a full description of the results see supplementary material and **Figure S1**).

fMRI Measures of Threat Learning

Threat appraisal (Anxiety-state): Approaching flowers (approach period) in the danger vs. the safety zone

Each approach period (approaching a flower) began at trial onset, when the flower appeared in the environment and ended when the flower was "collected." For analyses we excluded the initial orienting period (looking for the flower) of the approach period, only including the last 75% of the approach (active navigation towards the flower).

We compared brain activation (i.e., presumed metabolic activity) between diagnostic groups (ANX vs. HC) as individuals approached flowers located in the danger/safe zones of the environment (see Table S1 for

full results from this analysis). To assess discrimination learning, and increase signal-to-noise ratio, trials were divided into two blocks comprising the first- (early) and second-half (late) of the experiment. To directly examine group differences, the first-level analysis contrasted factors of zone (safe vs. danger) and block (early vs. late), whereas the second-level analysis directly compared groups (group: ANX vs HC). Significant peak activation was extracted and analyzed to disentangle the directionality of the results.

An overall group contrast between ANX and HC of approach periods (assessing changes across both zone and block) identified two opposing patterns of activity changes in a range of areas comprising posterior cingulate cortex (PCC; $p<0.05$ FWE), vmPFC, orbitofrontal cortex (OFC), and bilateral anterior hippocampus ($p<0.05$ FWE small volume corrected; SVC, **Fig. 2A**). Compared to the HC group, the ANX group demonstrated a greater increase in activity in these areas from early to late blocks (late > early) of the safe zone, and a decrease in activity from early to late blacks of the danger zone. To further understand these distinct patterns of activity, we next performed direct group comparisons on separate components of the task.

When approaching flowers in the danger zone (danger > safe), the ANX group (ANX > HC) showed greater activation of bilateral insula ($p<0.05$ FWE) and dmPFC ($p<0.05$ FWE SVC; **Fig. 2B**) compared to the HC group. That is, these areas were more responsive in the ANX group when approaching flowers in the danger compared to safe zone. No group differences were observed when looking for areas that showed greater activity when approaching the safe compared to danger zone.

For the HC group, compared to ANX (HC > ANX), approaching flowers in the second half of the experiment, compared with the first half (late > early) showed that, regardless of zone, there was increased activation from early to late blocks in PCC, vmPFC, OFC, and anterior hippocampus ($p<0.05$ FWE SVC). No other significant results were found ($p>0.001$).

Next, psychophysiological interactions (PPI) analyses were performed for each participant group separately to identify brain regions in which connectivity changed during the danger vs. safe contrast. PPI examined brain connectivity of each significant cluster (i.e., seed ROI) from the approaching flowers period. In the HC group, a positive association between the dmPFC seed (ROI, MNI coordinated: 9, 26, 45) and bilateral insula ($p<0.001$ Bonferroni corrected) was found in the contrast (dange > safe On the other hand, in ANX, a negative correlation between the dmPFC-OFC and dmPFC-vmPFC ($p<0.001$ Bonferroni corrected) was found in the same contrast (danger > safe). No other PPI analysis revealed significant results.

Overall, these results suggest that when approaching flowers in the dangerous zone, ANX showed reduced activation in vmPFC, PCC, and anterior hippocampus while showing greater activation in the insula and dmPFC. These findings are further highlighted by negative connectivity between the dmPFC and vmPFC areas. On the other hand, HC displayed greater activation as a function of time in the vmPFC, OFC, PCC, and anterior hippocampus, regardless of the zone suggesting appropriate contextual learning.

Threat anticipation (Fear-state): Held stationary in the danger vs. the safety zone

We next examined changes in brain activation when participants' positions were held stationary after picking flowers and anticipating a potential shock. We again examined the effects of zone (danger vs. safe zone) and block (early vs. late blocks) in the first level analysis. The second level analysis consisted of the group comparison (ANX vs. HC; see Table S2 for full results). Analyses followed the same model as for the approach period.

An overall group contrast between ANX and HC during stationary periods (assessing changes across both zone and block) identified two opposing patterns of activity changes in a range of areas comprising PCC ($p<0.05$ FWE), vmPFC, and OFC ($p<0.05$ FWE SVC). Compared to the HC group, the ANX group demonstrated a greater increase in activity in these areas from early to late blocks (late > early) of the safe zone, and a decrease in activity from early to late blacks of the danger zone. To further understand these distinct patterns of activity, we next performed direct group comparisons on separate components of the task.

When held stationary after picking a flower located in a zone of the environment associated with danger (danger > safe), the ANX group compared to HC (ANX > HC) showed greater activation in dmPFC, dACC, bilateral insula, caudate, thalamus, amygdala, and midbrain areas, including the periaqueductal gray (PAG; $p<0.05$ FWE SVC; **Fig. 3A**). A group contrast (ANX > HC) of stationary periods (irrespective of danger or safety) showed increased dmPFC activation from early to late block (late > early) in ANX compared with HC ($p<0.05$ FWE SVC).

For the HC group, compared to ANX (HC > ANX), flowers located in a zone of the environment associated with safety (safe > danger) generated greater activation in the PCC ($p<0.05$ FWE), vmPFC, OFC, and anterior hippocampus ($p<0.05$ FWE SVC; **Fig. 3B**). For HC, compared to ANX (HC > ANX), we also found increased activation during the last half of learning (late > early) regardless of zone in vmPFC and OFC ($p<0.05$ FWE SVC). No other significant results were found ($p>0.001$).

Given the reported group differences during stationary periods, PPI analyses were next performed for each group separately to identify brain regions in which connectivity changed during danger vs. safe. PPI examined brain connectivity of each significant cluster (i.e., seed ROI) from the stationary period. PPI analyses used dmPFC (MNI coordinates: 0, -8, 71) and each amygdala side as seed regions (MNI coordinates: (R) 26, -2, -15, (L) -20, 2, -15). ANX showed increased functional connectivity between the dmPFC-bilateral insula, left amygdala-bilateral insula, and right amygdala-bilateral insula in danger compared to safe zones ($p<0.001$ Bonferroni corrected). HC showed increased functional connectivity between the dmPFC-OFC, left amygdala-bilateral insula, right amygdala-bilateral insula, left amygdala-vmPFC, right amygdala-vmPFC, left amygdala-OFC, and right amygdala-OFC in danger compared to safe zone during the stationary period ($p<0.001$ Bonferroni corrected). No other PPI analyses revealed significant results.

In summary, during the stationary period in the danger zone (after collecting a flower) ANX, compared to HC, demonstrated reduced activation in vmPFC and PCC over time, with greater activation in the insula, amygdala, and PAG. This was further highlighted by increased connectivity among the dmPFC, amygdala, and insula while lacking any significant connectivity from vmPFC areas. On the other hand, HC's recruitment of the vmPFC, OFC, and PCC was seen as a function of time, where those areas were recruited during the stationary periods in both the safe and dangerous zones. Furthermore, HC had significant connectivity of the vmPFC and OFC to areas such as the dmPFC and amygdala.

Discussion

The present study examined how patients with pathological anxiety learn to distinguish contextual features that inform the dangerous vs. safe status of stimuli within a virtual environment. This virtual environment task shows how brain regions interact to shape behavior over time. We hypothesized that patients with pathological anxiety would show (1) poor threat learning associated with lower recruitment of the anterior hippocampus; (2) biased threat appraisal with stronger engagement of the dACC, dmPFC, amygdala, and insula (higher anxiety-state); and (3) higher threat anticipation with stronger engagement of the PAG (higher fear-state).

Surprisingly, inconsistent with our hypotheses, patients with pathological anxiety showed no impairment in discriminating between threat and safety within the environment based on their skin conductance and subjective reports. However, consistent with our hypothesis, patients with pathological anxiety, compared to HC, displayed (1) weaker engagement of the anterior hippocampus and vmPFC implicated in memory processing and emotional regulation during both approach and stationary periods (*Threat learning and discrimination*); (2) higher engagement of the dACC, dmPFC, amygdala, and insula involved in negative valence, emotional expression, and conflict processing during both approach and stationary periods (*Threat appraisal*); and (3) higher engagement of midbrain areas, including the PAG, during anticipation of potential shock delivery (*Threat anticipation*; see **Fig. 4**). These results suggest that patients with pathological anxiety recruit compensatory learning strategies that do not involve the anterior hippocampus or vmPFC, which might explain the increased activation in the salience network (e.g., dmPFC, insula, PAG). In patients with pathological anxiety, reintegrating the anterior hippocampus and vmPFC into the circuitry could rescue appropriate learning strategies to discriminate between safety and danger.

Threat Learning and Discrimination

We observed changes in activation in several brain regions as the experiment progressed (first vs. second block) that differed between patients with pathological anxiety and healthy comparisons. Healthy comparisons showed increased activation over time, for regions involved in emotion regulation (ventromedial prefrontal cortex; vmPFC), valuation (orbitofrontal cortex; OFC), spatial memory (hippocampus) and self-referential processing (posterior cingulate cortex; PCC) regardless of safe/danger zone. In patients with pathological anxiety, these brain areas (vmPFC, OFC, hippocampus, and PCC)

differentiated between zones, with increased activation in the safe zone and decreased activation in the danger zone over time. The anterior hippocampus is essential for the integration of spatial information in mediating anxiety-like behaviors.[22] As seen in the current study with healthy comparisons, previous findings show that the hippocampus is involved in learning and discrimination of safe and danger zones, with increased activation with task experience.[8] Therefore, decreased engagement of the anterior hippocampus might reflect an inability of patients with pathological anxiety to appropriately use context-specific cues for emotional regulation. These results are in line with research showing hippocampus and medial prefrontal cortex abnormalities in patients suffering from pathological anxiety[23,24] who report an exaggerated response to threats in contexts predicting safety. Reduced activation of the anterior hippocampus, particularly during threat, could impair emotional regulation abilities when shifting contexts.

Patients with pathological anxiety displayed reduced activation in the vmPFC, OFC, and PCC throughout the flower task; particularly, when interacting with flowers in the danger zone during the late blocks. These vmPFC/OFC areas are considered to have key roles in regulating contextually appropriate emotions.[25–28] Previous studies reported that, in healthy individuals, vmPFC, PCC, and anterior hippocampus activation did not differentiate danger and safety during this flower task, but rather increased with experience (from early to late block of the experiment), suggesting a role in learning and discrimination.[8] In the present study, patients with pathological anxiety, unlike healthy comparisons, demonstrated increased engagement for safe flowers and decreased engagement for dangerous flowers. Given the role of the anterior hippocampus and vmPFC/OFC, the hypoactivation observed in patients with pathological anxiety may suggest that they are unable to use contextual cues to regulate their emotional output, particularly in a potentially dangerous location.

Threat Appraisal (Anxiety-State)

Compared to healthy comparisons, patients with pathological anxiety demonstrated greater activation in brain areas involved in threat appraisal (dmPFC, dACC, amygdala, insula) throughout the whole experiment, particularly in the danger zone. Greater engagement of these brain areas likely reflects higher arousal when interacting with the flowers, particularly in the danger zone. Activation in the dmPFC, along with amygdala and its connectivity, has been associated with threat arousal[29–32] and may play a key role in integrating threat information to coordinate emotional response with other brain regions. Increased dmPFC activation to threat stimuli has been documented in patients with pathological anxiety, as compared to healthy comparisons,[25,26,30,31] suggesting an overactive appraisal of surroundings and cues. Accordingly, dACC and insula activations have been reported during approach of threat in healthy adults using the flower task,[8] likely reflecting the integration of visceral feelings and cognitive appraisals of threat to trigger fear expression.[33] However, in this study, patients with pathological anxiety showed greater dmPFC, dACC, and insula activation throughout the task, possibly reflecting an overactive threat appraisal and detection network.

Threat Anticipation (Fear-state)

During imminent threat, patients with pathological anxiety exhibited greater activation in brain areas related to error detection, conflict resolution, and emotional expression (dACC and amygdala), and regions related to salience and pain judgment (bilateral insula and PAG). Midbrain areas, including the PAG, and amygdala have been reported during imminent threat in healthy adults using the same task,[8] likely reflecting the integration of cognitive appraisals to the emotional anticipation of threat to trigger appropriate fear expression.[33] However, in this study, patients with pathological anxiety, as compared to healthy comparisons, showed greater PAG and amygdala activation during imminent threat, possibly reflecting an overactive threat anticipation. Particularly, increased PAG activation during stationary periods in the danger zone might suggest greater anticipation of the potential shock just prior to its likely delivery.

Brain circuitry connectivity

Unlike healthy comparisons, patients with pathological anxiety failed to exhibit task-related functional connectivity with higher-level cortical areas related to emotional regulation (vmPFC and OFC) throughout the task. During flower-approach, patients with pathological anxiety displayed negative functional connectivity between the dmPFC and both the vmPFC and OFC. During stationary periods in the danger zone, patients with pathological anxiety displayed positive dmPFC-insula and amygdala-insula functional connectivity, while in healthy comparisons these brain areas were functionally connected to the OFC and vmPFC. Coupling between the dmPFC and amygdala has been previously observed with induced threat of shock[25,34,35] and in patients with pathological anxiety[36] during an emotional identification task. In pathological anxiety, this circuit may become hyperactive and unable to “turn off,” contributing to an attentional bias towards threat. In this study, functional connectivity deficits with vmPFC/OFC areas, in patients with pathological anxiety, could reflect a faulty flexible learning spatial strategy. For example, patients with pathological anxiety may fail to integrate the learned contingencies within the environment and surrounding landmarks to create a spatial representation that can modulate emotional output to stimuli while navigating through the safe and danger zones. Therefore, an inflexible or exaggerated evaluation of the cue may hinder the use of spatial strategies to regulate emotional output appropriately. In other words, patients with pathological anxiety might have shock-related worry that gets progressively worse as they approach and pick up the flower, and as the experiment progresses.

In conclusion, patients with pathological anxiety show lower activation in the anterior hippocampus and vmPFC/OFC and lower connectivity to emotion expression brain areas (dmPFC, amygdala, and insula), which could explain poor emotional regulation. This possibility is further supported by higher dmPFC activation and other emotion valuation and expression brain areas. Our findings suggests that patients’ threat-related attention may cause deficits in emotion regulation. The current results suggest that disconnected circuitry in brain areas essential for emotional regulation might lead to disrupted emotional output while exploring environments and learning to discriminate cues within these environments. Finding novel psychological intervention or trainings to reintegrate the vmPFC and hippocampus into the

learning and discrimination circuitry could help patients with pathological anxiety display appropriate emotional responses while navigating novel environments.

Methods

Participants. Seventy participants, aged 18-50 years, were recruited from the Washington D.C. and Maryland areas. Thirty participants were diagnosed with an anxiety disorder (ANX; generalized anxiety disorder and/or social anxiety disorder) using the Structured Clinical Interview for Diagnostic and Statistical Manual of Mental Disorders 4^{th} edition, while the other 40 volunteers were healthy comparisons (HC). Before taking part, all participants provided written informed consent and, after completion, were debriefed and reimbursed for their time. The study was approved by the NIH Institutional Review Board. All participants were right-handed, free from neurological impairment, or any psychological disorders (except for an anxiety disorder in the patient group). Four ANX and 10 HC participants were excluded from analyses because they were unable to explain the shock contingencies between the locations at the end of the task (see procedure below). Three ANX and two HC were omitted due to technical issues during scanning or excessive head motion (>20%; see below). Therefore, the final sample included 23 ANX (mean age=29.6; SD=8.22) and 28 HC (mean age=27.25; SD=8.21). Participants were free of medications and no significant differences in any demographic information between groups was found.

Skin conductance. Skin conductance was measured as an index of anxiety via 8mm Ag/AgCl electrodes attached to the medial phalanges of the index and middle fingers of the participant's left hand. Data were acquired using a Biopac EDA100C MRI system (Biopac Systems, Inc., Goleta, CA, USA) at a sampling rate of 1000Hz.

Shocks. Shocks (20 samples) were applied using a Digitimer DS7A electrical stimulator (Digitimer, Welwyn Garden City, UK) to the left hand with intensity up to 50mA for 2ms duration through a silver chloride electrode. Shock intensity was adjusted individually for each participant before starting the experiment. Individual adjustment procedures delivered a series of shocks to each participant, starting at 12mA. Participants were asked to rate the level of pain with each shock on a 1-10 scale. Shock intensity was increased until the level was irritating, but not painful.

Task. A description of the virtual environment task, developed with Unity Software (Unity Technologies, USA) is available in a prior publication.[8] The virtual environment consisted of a circular grassland with a perimeter boundary wall surrounded by distal cues (mountains, sun, and clouds) for orienting, two landmarks (beehives) placed in the grassland, and flowers that appeared one at a time in random locations. When participants picked a flower, they rated on a scale from 0-to-9 the likelihood of receiving a shock from the flower picked. After, the virtual character would enter a stationary period for 2000ms-8000ms. After the stationary, participants could move once again, and another flower would appear in the environment. The task included a total of 80 flowers—40 flowers were paired with a shock on 50% of the trials (danger) while encountering no shock with the remaining 40 (safe).

After every four flower trials, we included a spatial memory trial in which participants learned the location of one of four objects (wooden box, gas can, book, and clock); however, two objects appeared in each half of the environment. No shocks accompanied spatial memory stimuli. After the first four spatial memory trials, participants' memory for object locations was tested by asking participants to replace the objects where they originally found them (For Object task results see supplementary material).

At the end of the experiment, participants were asked to name the four objects and their locations, as well as explain the contingencies of danger and safety during threat learning, by answering if there was a pattern in the shocks. Participants who were unable to provide the objects' name and position, or explain the contingencies were excluded from the final analysis to ensure that participants were paying attention to the task.

Behavioral analysis. Data processing and analysis of electrodermal activity (EDA) were performed using MATLAB. EDA data were down-sampled to 200 Hz and then synchronized to the task. EDA was assessed during two periods of the threat learning task. First, mean skin conductance level during each approach quantified tonic skin conductance levels (SCL) as participants navigated towards the flower. SCL was quantified from the last three-quarters of the approach period from flower appearance until trial completion. Skin conductance level was calculated by measuring the mean skin conductance from the beginning of active approach until right before the flower was picked for each trial. Second, skin conductance responses (SCR) were analyzed during the stationary period to examine phasic changes in anticipation of the shock outcome. SCRs were calculated for every trial by subtracting the minimum skin conductance during the stationary period (baseline) from the maximum response (peak) before the stimulus onset. Any response difference under 0.03 micro-Siemens was scored zero. SCRs were log transformed (log [1+SCR]) to normalize the distribution, and then range correction ([SCR-SCRmin]/[SCRmax-SCRmin]) was applied to control for individual variation in responding.[37] The same correction was applied to the SCLs. For analyses, SCRs and SCLs were averaged into four equal blocks across the duration of the experiment, with each block including ten trials per condition (safe and danger).

Expectancy ratings taken at the beginning of each stationary period were analyzed similarly to skin conductance. Each rating (0-9) was averaged across trials to create four equal blocks separated by safe and danger conditions (10 trials in each block).

Finally, performance on the spatial memory task was analyzed by assessing distance error on each test trial. This distance error was calculated by taking the distance in virtual meters between the participant's response location when replacing the object and its correct location within the environment. Distance error was taken from each trial and averaged into four blocks (1 trial from each object in each block).

All results were analyzed using a General Linear Model (GLM) for repeated measures using 2x2x4 ANOVAs to test differences between zone (safe, danger), group (HC, ANX), and block (1 to 4). Based in *a*

priori hypothesis of ANX overgeneralization[12], the simple effects within each group were examined using 2x4 ANOVAs to characterize differences between conditions (safe, danger) and block (1 to 4).

fMRI acquisition. Blood oxygen level-dependent T2*-weighted functional images were acquired on a 3T Skyra system (Siemens, Germany) using echo-planar imaging (EPI) with a 32-channel head coil. Images were acquired with a 45° oblique angle with the following parameters: 3300 ms TR; 30 ms TE; 1 mm inter-slice gap, 192 mm field of view, and 48 axial slices with 2 mm slice thickness resulting in 3mm isotropic voxels. A single echo field map was recorded for distortion correction of the acquired EPI. After the functional scans, a T1-weighted 3-D structural image ($1mm^3$) was acquired to co-register and display the functional data.

fMRI analysis. Data processing and analysis were performed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm). EPI images were first preprocessed using a bias correction to control for within volume signal intensity difference, unwarping and realignment to correct for movement and slice-time correction. Images were then spatially normalized to the MNI template using parameter estimates from warping each participant's structural image to a T1-weighted average template image. All images were finally smoothed using an 8mm FWHM Gaussian kernel.

The analysis model included 16 regressors of interest. Four separate regressors of interest were created for approach periods by zone (safe or danger) and block (first or second half of experiment); these consisted of boxcar functions from the end of the first quarter of each approach period to the point in which the flower was reached. Four regressors of interest (also by zone and block) were created for the stationary period of each trial, consisting of a boxcar function starting after the participant rated their shock expectancy and lasting for the duration of the stationary period. Three regressors using a stick/impulse function modeled the end of each trial in safe conditions and in danger conditions split by whether participants received a shock or not. Finally, four regressors modeled when participants were replacing objects during the spatial memory task, using a boxcar function covering when a response was made and the approach period to the location where the object was picked (first and second half of the experiment). Six regressors of no interest were also added to the model representing movement parameters estimated during realignment. Frames with more than 0.5 mm frame-wise head motion were detected as outliers and modeled using Artifact detection tool (ART). Participants with outliers totaling 20% or more of their total scan were removed from analysis.

Statistical analyses occurred in two stages, individual (General Linear Model analysis) and group level (two sample t-test). At the individual level, an initial control analysis was conducted to examine emotional flower vs. unemotional object approaches to a cue. The neural responses to approach periods were compared between the threat learning task (approaching flowers) and the spatial memory task (approaching location to replace the object; See supplementary material). We created a model contrasting approach periods for threat learning (collapsing across safety and danger) with approach during spatial memory across the first and second half of the experiment using a 2x2 ANOVA (task, block). Approach periods during threat learning were then examined contrasting approach to flowers associated with

safety or danger and collected during the first or second half of the experiment using a 2x2 ANOVA (zone, block). Finally, stationary periods during threat learning, periods when the flower was picked, and participants were held stationary, were analyzed in a similar first-level analysis using a 2x2 ANOVA (zone, block).

At the group level, the contrasts of interest, described above, were entered into two-sample t-tests to compare between the ANX and HC. Within this model, t-contrasts were examined to investigate ANX>HC and HC>ANX differences. To aid and confirm interpretation of significant group effects, parameter estimates were extracted from significant peak coordinates for the contrasts ANX>HC and HC>ANX using the MarsBaR region of interest toolbox. Post-hoc analysis of these contrasts was conducted using repeated measures 2x2x2 ANOVA (zone or task, block, group) in SPSS.

Family-wise error (FWE; $p<0.05$) corrected effects across the whole brain are reported for all analyses. Given the *a priori* hypotheses and previous findings,[8] an ROI approach was employed, based on the Automated Anatomical Labeling atlas (AAL)[38] and the WFU Pickatlas toolbox in SPM12.[39] A bilateral mask was defined for the hippocampus, amygdala, and mPFC (orbitofrontal gyrus, ventromedial prefrontal cortext, dorsomedial prefrontal cortex, and anterior cingulate and medial cingulate cortex). Statistical threshold was defined by small volume correction (SVC; $p<0.05$ FWE).

For any significant interaction, the representative time-course was extracted through SPM12 MarsBaR (http://marsbar.sourceforge.net) toolbox, using a 6mm sphere centered on the peak of the activation in the regions of interest, using the first eigenvariate calculated from singular value decomposition. The extracted values were analyzed in SPSS 24 with a 2x2 ANOVA (task x block) and further analyzed using a one-sample t-test, which was Bonferroni corrected.

'Activation' was used throughout the manuscript to indicate an increase in presumed metabolic activity.

Functional connectivity analyses. Functional connectivity was assessed at the individual level using psychophysiological interactions (PPI) analysis with SPM12. The PPI compares functional connectivity from a single seed region across multiple task conditions. For exploratory analyses, seed regions were selected based on group differences identified in the main analysis, such as the vmPFC, OFC, dmPFC, amygdala, and hippocampus. Peak activation from these brain areas in the group level analysis, for approach and stationary periods, were used to create regions of interest (6 mm sphere centered on group level peak activation) for each participant. The seed time series activation for each participant was extracted at the center of the activation peak. The individual t-contrast images of the interaction from the PPI were examined using a group level one-sample t-test. As in a previous study,[8] the PPI were detected using t-test with a threshold of $p<0.001$ corrected for multiple comparison.

Declarations

Acknowledgments

This research was funded by the Medical Research Council and the Wellcome Trust, United Kingdom. The authors are grateful to the Welcome Trust Centre for Neuroimaging at UCL for providing facilities. We thank the National Institute of Mental Health, UCL Graduate Partnership Program, NIMH T32 grant MH015144, and NIMH K01 MH118428-01 for providing support to B.S.J.

Author Contributions

BSJ designed the experiment with contributions from JAK, DSP, CG, NB, and ME.; BSJ programmed the VR task.; BSJ collected the data with contributions from NLB, JCL, and AH; BSJ analyzed the data with contributions from NLB, JAB, NB, and ME; BSJ, JAB, DSP, CG, NB and ME interpreted the data.; BSJ wrote the manuscript with contributions from all authors.

Competing Interest Statement

The authors declare that they have no competing financial interest.

References

1. Davis, M., Walker, D. L., Miles, L. & Grillon, C. Phasic vs sustained fear in rats and humans: role of the extended amygdala in fear vs anxiety. *Neuropsychopharmacology* **35,** 105–35 (2010).
2. Lang, P. J., Davis, M. & Ohman, A. Fear and anxiety: animal models and human cognitive psychophysiology. *J Affect Disord* **61,** 137–59 (2000).
3. Sylvers, P., Lilienfeld, S. O. & LaPrairie, J. L. Differences between trait fear and trait anxiety: implications for psychopathology. *Clin Psychol Rev* **31,** 122–37 (2011).
4. Tovote, P., Fadok, J. P. & Lüthi, A. Neuronal circuits for fear and anxiety. *Nat. Rev. Neurosci.* **16,** 317–31 (2015).
5. LeDoux, J. E. & Pine, D. S. Using Neuroscience to Help Understand Fear and Anxiety: A Two-System Framework. *Am J Psychiatry* **173,** 1083–1093 (2016).
6. Britton, J. C., Lissek, S., Grillon, C., Norcross, M. A. & Pine, D. S. Development of anxiety: the role of threat appraisal and fear learning. *Depress Anxiety* **28,** 5–17 (2011).
7. Pine, D. S. Research review: a neuroscience framework for pediatric anxiety disorders. *J Child Psychol Psychiatry* **48,** 631–48 (2007).
8. Suarez-Jimenez, B. *et al.* Linked networks for learning and expressing location-specific threat. *Proc. Natl. Acad. Sci. U.S.A.* **115,** E1032–E1040 (2018).
9. Schwabe, L. *et al.* Stress modulates the use of spatial versus stimulus-response learning strategies in humans. *Learn. Mem.* **14,** 109–16 (2007).

10. Schwabe, L., Bohbot, V. D. & Wolf, O. T. Prenatal stress changes learning strategies in adulthood. *Hippocampus* **22,** 2136–43 (2012).
11. Schwabe, L., Dalm, S., Schächinger, H. & Oitzl, M. S. Chronic stress modulates the use of spatial and stimulus-response learning strategies in mice and man. *Neurobiology of learning and memory* **90,** 495–503 (2008).
12. Kheirbek, M. A., Klemenhagen, K. C., Sahay, A. & Hen, R. Neurogenesis and generalization: a new approach to stratify and treat anxiety disorders. *Nat. Neurosci.* **15,** 1613–20 (2012).
13. Fanselow, M. S. & Dong, H.-W. Are the dorsal and ventral hippocampus functionally distinct structures? *Neuron* **65,** 7–19 (2010).
14. Grillon, C. *et al.* Increased anxiety during anticipation of unpredictable aversive stimuli in posttraumatic stress disorder but not in generalized anxiety disorder. *Biol. Psychiatry* **66,** 47–53 (2009).
15. Sotres-Bayon, F., Sierra-Mercado, D., Pardilla-Delgado, E. & Quirk, G. J. Gating of fear in prelimbic cortex by hippocampal and amygdala inputs. *Neuron* **76,** 804–12 (2012).
16. Schiller, D., Levy, I., Niv, Y., LeDoux, J. E. & Phelps, E. A. From fear to safety and back: reversal of fear in the human brain. *J. Neurosci.* **28,** 11517–25 (2008).
17. Lissek, S. *et al.* Generalization of conditioned fear-potentiated startle in humans: experimental validation and clinical relevance. *Behav Res Ther* **46,** 678–87 (2008).
18. Herry, C. *et al.* Switching on and off fear by distinct neuronal circuits. *Nature* **454,** 600–6 (2008).
19. Orsini, C. A., Kim, J. H., Knapska, E. & Maren, S. Hippocampal and prefrontal projections to the basal amygdala mediate contextual regulation of fear after extinction. *J. Neurosci.* **31,** 17269–77 (2011).
20. Schwabe, L. *et al.* Stress modulates the use of spatial versus stimulus-response learning strategies in humans. *Learning \& memory* **14,** 109–116 (2007).
21. Kim, J. J., Lee, H. J., Han, J. S. & Packard, M. G. Amygdala is critical for stress-induced modulation of hippocampal long-term potentiation and learning. *J. Neurosci.* **21,** 5222–8 (2001).
22. Bach, D. R. *et al.* Human hippocampus arbitrates approach-avoidance conflict. *Curr. Biol.* **24,** 541–7 (2014).
23. Paulus, M. P. & Stein, M. B. An insular view of anxiety. *Biol. Psychiatry* **60,** 383–7 (2006).
24. Etkin, A. & Wager, T. D. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *American Journal of Psychiatry* **164,** 1476–1488 (2007).
25. Mechias, M.-L., Etkin, A. & Kalisch, R. A meta-analysis of instructed fear studies: implications for conscious appraisal of threat. *Neuroimage* **49,** 1760–8 (2010).
26. Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci. (Regul. Ed.)* **15,** 85–93 (2011).
27. Motzkin, J. C., Philippi, C. L., Wolf, R. C., Baskaya, M. K. & Koenigs, M. Ventromedial prefrontal cortex is critical for the regulation of amygdala activation in humans. *Biol. Psychiatry* **77,** 276–284 (2015).

28. Passamonti, L. *et al.* Effects of acute tryptophan depletion on prefrontal-amygdala connectivity while viewing facial signals of aggression. *Biol. Psychiatry* **71,** 36–43 (2012).
29. Milad, M. R. *et al.* Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert. *Biol. Psychiatry* **62,** 446–54 (2007).
30. Graham, B. M. & Milad, M. R. The study of fear extinction: implications for anxiety disorders. *American Journal of Psychiatry* **168,** 1255–1265 (2011).
31. Milad, M. R. *et al.* A role for the human dorsal anterior cingulate cortex in fear expression. *Biol. Psychiatry* **62,** 1191–4 (2007).
32. Quirk, G. J. & Beer, J. S. Prefrontal involvement in the regulation of emotion: convergence of rat and human studies. *Current opinion in neurobiology* **16,** 723–727 (2006).
33. Mobbs, D. *et al.* When fear is near: threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science* **317,** 1079–83 (2007).
34. Robinson, O. J., Charney, D. R., Overstreet, C., Vytal, K. & Grillon, C. The adaptive threat bias in anxiety: amygdala-dorsomedial prefrontal cortex coupling and aversive amplification. *Neuroimage* **60,** 523–9 (2012).
35. Vytal, K. E., Overstreet, C., Charney, D. R., Robinson, O. J. & Grillon, C. Sustained anxiety increases amygdala-dorsomedial prefrontal coupling: a mechanism for maintaining an anxious state in healthy adults. *J Psychiatry Neurosci* **39,** 321–9 (2014).
36. Robinson, O. J. *et al.* Towards a mechanistic understanding of pathological anxiety: the dorsal medial prefrontal-amygdala “aversive amplification” circuit in unmedicated generalized and social anxiety disorders. *Lancet Psychiatry* **1,** 294–302 (2014).
37. Lykken, D. Range correction applied to heart rate and to GSR data. *Psychophysiology* **9,** 373–379 (1972).
38. Rolls, E. T., Joliot, M. & Tzourio-Mazoyer, N. Implementation of a new parcellation of the orbitofrontal cortex in the automated anatomical labeling atlas. *Neuroimage* **122,** 1–5 (2015).
39. Maldjian, J. A., Laurienti, P. J., Kraft, R. A. & Burdette, J. H. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* **19,** 1233–9 (2003).

Figures

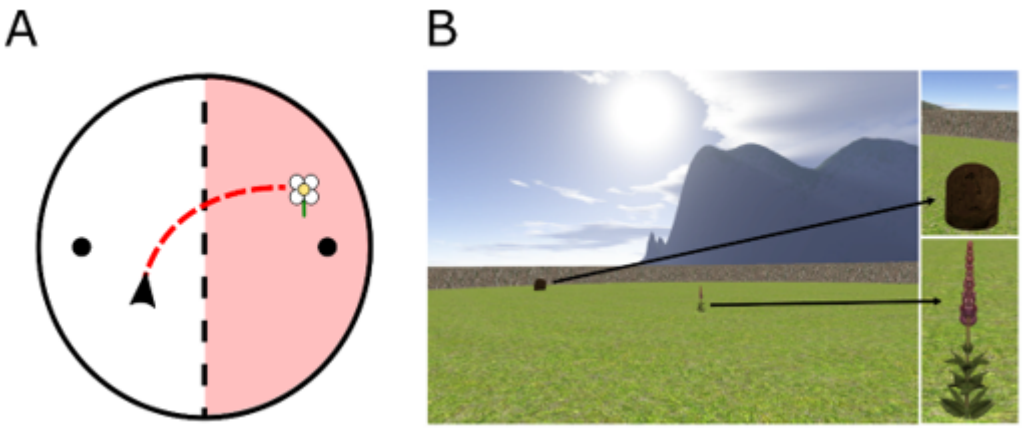

Figure 1

(A) Helicopter view of the circular environment that participants (black arrow) explored (red trace) showing how the environment was split into one-half associated with danger (red) and the other with safety. The environment included two beehives (black dots) located at opposite sides of the environment. Participants were required to collect flowers, which were generated randomly within the environment. (B) Example of the participant's viewpoint, showing a beehive and flower in the environment.

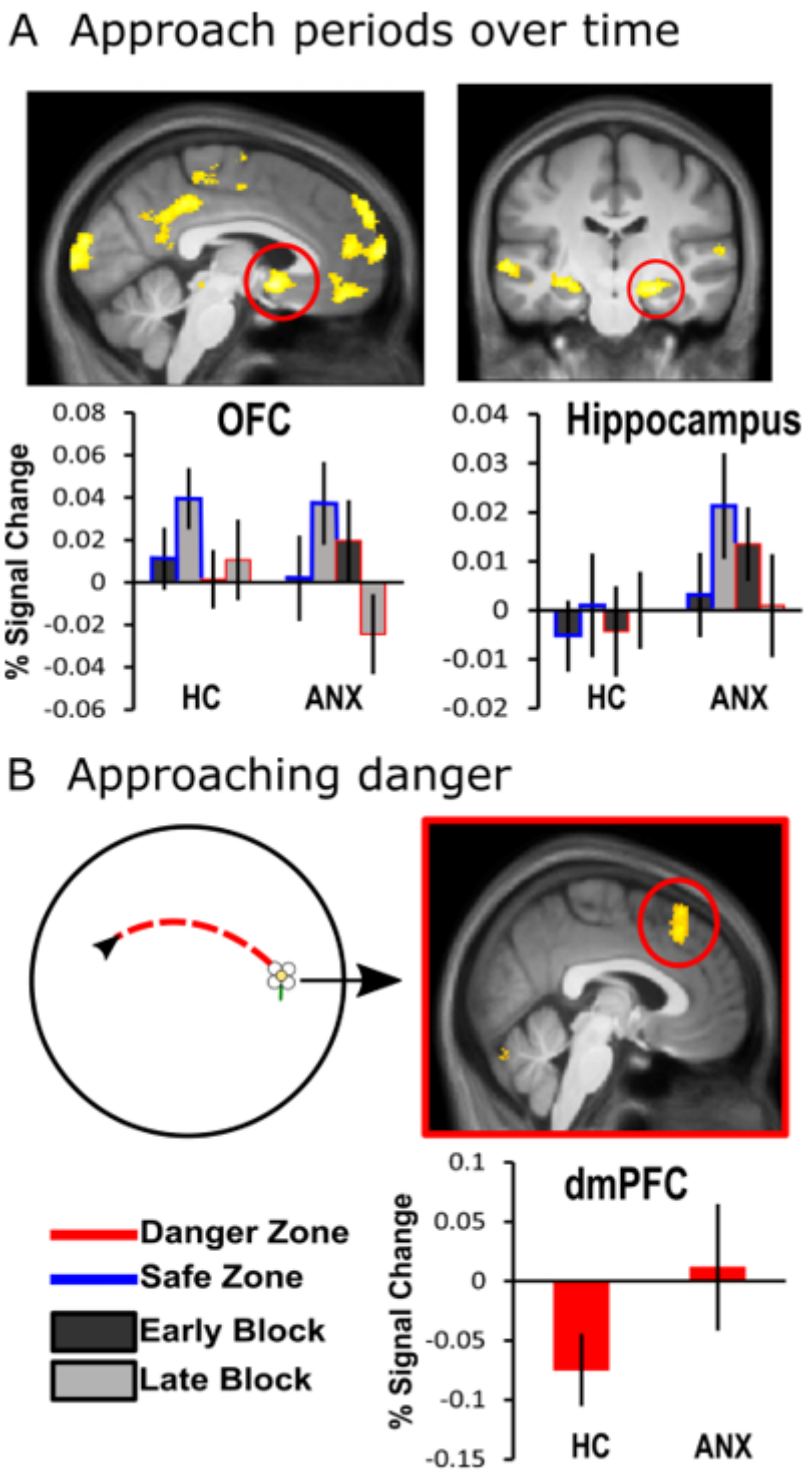

Figure 2

fMRI results of approaching flowers. (A) For ANX, activation increased from the early to late block of the experiment when approaching flowers in the safe zone compared to the danger zone in medial parietal areas (p<0.05 FWE), ventromedial prefrontal cortex (vmPFC),orbitofrontal cortex (OFC; top left panel), and anterior hippocampus (p<0.05 FWE SVC; top right panel). (B) Helicopter view of the circular environment that participants (black arrow) explored to approach the flower (red trace). For flowers in the danger compared to the safe zone, there was greater activation in dorsomedial prefrontal cortex (dmPFC) across

the whole test session in ANX compared to HC ($p<0.05$ FWE SVC; lower right panel). All images are presented at $p<0.001$ uncorrected for display purposes. Percentage signal changes for danger and safety across early and late periods of learning extracted from (A) OFC (MNI coordinates: 5, 9, -11; left panel) and anterior hippocampus (MNI coordinates: -27, -17, -14; right panel); and (B) dmPFC (MNI coordinates: 9, 26, 45; right panel). Error bars show standard error mean.

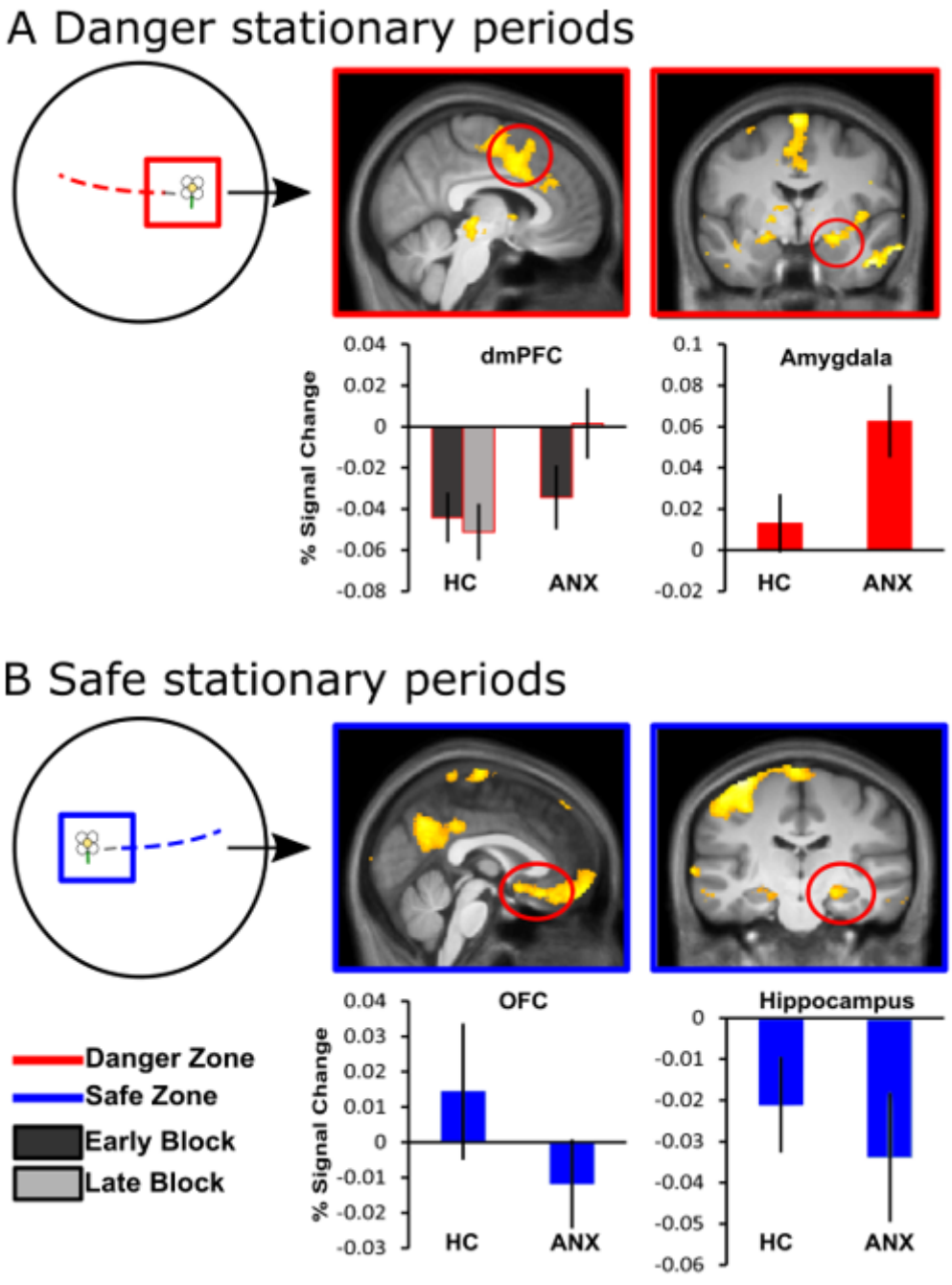

Figure 3

fMRI results of stationary periods. Circular illustrations: Helicopter view of the circular environment that participants explored. The stationary period is represented for the dangerous-flower as a red square and for the safe-flower as a blue square. (A) stationary period in the dangerous versus safe zone showed greater activation in periaqueductal gray, dorsomedial prefrontal cortex (dmPFC), dorsal anterior

cingulate cortex (dACC; middle panel), amygdala, and insula (p<0.05 FWE; right panel) in ANX. (B) Stationary period in the safe vs. dangerous zone showed greater activation in the posterior cingulate cortex (PCC; p<0.05 FWE), ventromedial prefrontal cortex (vmPFC), orbitofrontal cortex (OFC; lower middle panel), and anterior hippocampus (p<0.05 FWE SVC; lower right panel) in HC. Images are presented at p<0.001 uncorrected for display purposes. Percentage signal changes during stationary periods for danger and safety across early and late parts of learning extracted from (A) dmPFC (MNI coordinates: 0, -8, 71; middle panel) and amygdala (MNI coordinates: 26, -2, -15; right panel), and (B) OFC (MNI coordinates: -3, 54, -17; middle panel) and anterior hippocampus (MNI coordinates: -32, -29, 12; right panel). Error bars show standard error mean.

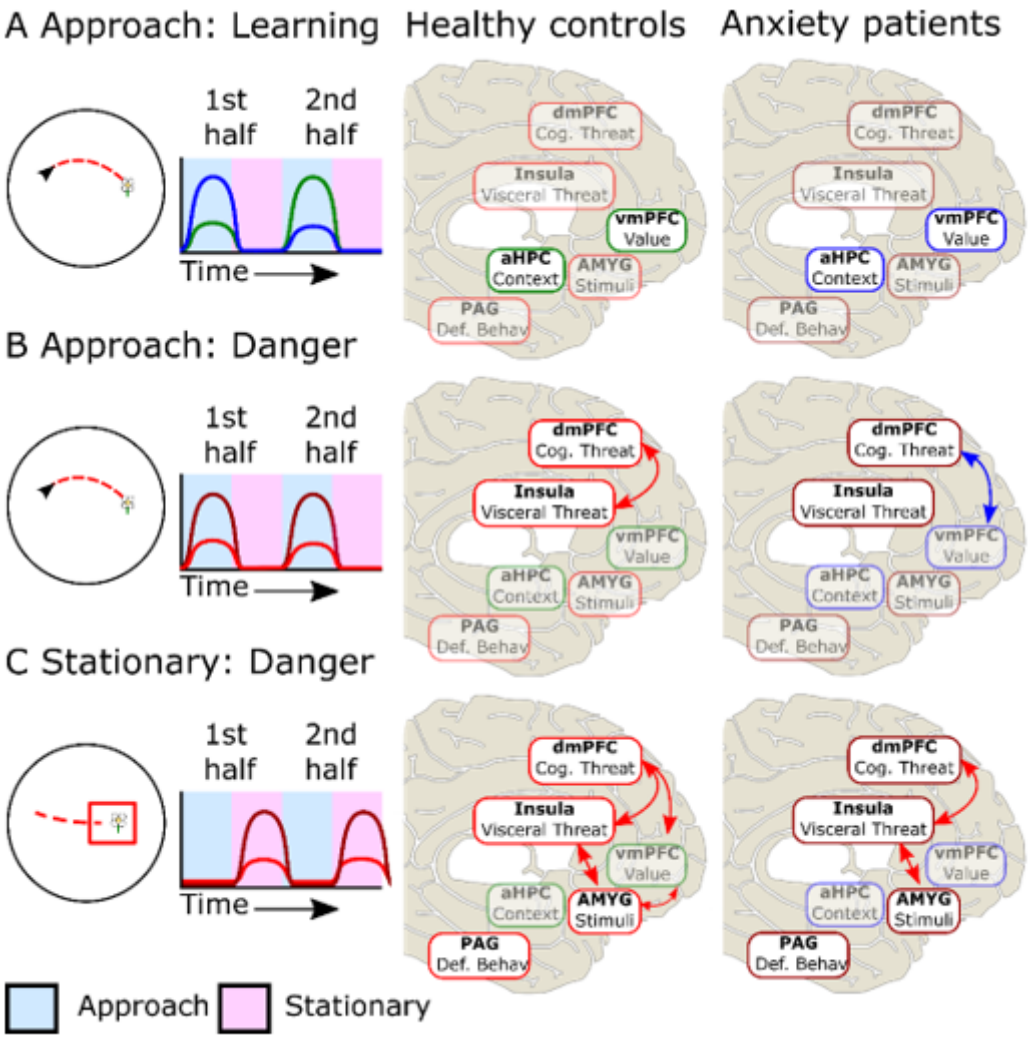

Figure 4

Illustration of sequential network activation in the flower task between groups. Left Panel: Helicopter view of task phase and schematic of activations over blocks. Middle and Right Panels: Brain activation and functional connectivity of HC (middle) and ANX (right). Green lines/boxes represent activation (and red arrows functional connectivity) that increased from the first to second half of the experiment. Blue lines/boxes represent activation (and blue arrows functional connectivity) that decreased from the first to second half of the experiment. Red lines/boxes represent activation (and red arrows functional

connectivity) that increased with danger, with darker red representing higher activation. (A) For HC, during approach of dangerous flower (left panel), activation in the anterior hippocampus (aHPC) and ventromedial prefrontal cortex (vmPFC) increased in the late-phase compared to the early-phase of learning (middle panel), while it decreased for ANX (right panel). (B) For HC, during approach of flowers predicting danger (left panel), activation in the insula and dorsomedial prefrontal cortex (dmPFC) showed positive connectivity (middle panel). For ANX, compared to HC, higher activation was evident in the insula and dmPFC and there was a negative connectivity of dmPFC-vmPFC (right panel). (C) For HC, when danger was imminent during stationary period (left panel), Insula-Amygdala (Amyg), Insula-dmPFC, vmPFC-dmPFC, vmPFC-Amyg were positively connectivity (middle panel). For ANX, compared to HC, higher activation was evident in dmPFC, insula, amygdala, and periaqueductal grey (PAG) and there was a positive connectivity in insula-dmPFC and insula-Amyg. See Tables S1-3 for a complete breakdown of regions across these analyses.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NIMHflowerpaper081020FinalSupplementaryMaterial.docx