Methods

**Participant Screening**

Participants were recruited through flyers, listings on UCLA campus newsletters, and advertisements on online boards (e.g. Nextdoor, Reddit, etc). Participants were screened using a modified Fear Survey Schedule (cite) and were required to endorse fear levels of a 3 or 4 for at least two animals available in our image dataset to be eligible. Participants meeting these criteria were then interviewed using the ADIS-V (cite) by a trained staff member and reviewed for consensus with the Principal Investigator. During the interview, participants were required to meet diagnostic criteria for at least two specific animal phobias with either a fear or avoidance rating of 4 or greater on a 0 – 8 point scale (0 = no fear/never avoids, 8 = very severe fear/always avoids). Participants were excluded if at any time they 1) were unable to understand or complete the informed consent process; 2) were unable to respond to screening questions; 3) did not have normal/corrected to normal vision/hearing; 4) had history of neurological disease/defect or any other serious and unstable medical conditions; 5) were diagnosed with current MDD, SUD, OCD, PTSD, bipolar disorder, or psychosis; 6) were currently taking psychotropic medication, or; 7) did not meet MRI scanning safety criteria.

23 participants (insert mean ages and gender distribution) meeting criteria for at least two animal phobias were enrolled to undergo multi-voxel neuro-reinforcement after undergoing the informed consent process as approved by the Institutional Review Board of the University of California, Los Angeles and were randomly assigned to 1, 3, or 5 days of multi-voxel neuro-reinforcement. From the 23 participants, two were unable to complete the full treatment. Out of the remaining 21 who completed treatment, one was excluded for nausea and two did not complete the “fear test” amygdala response task, leaving 18 participants with complete amygdala response data. However, seven of these 18 participants did not have complete pre-post resting state scan data (due to scanning time limitations) as required for our group ICA method, leaving a total of 11 participants analyzed.

**Resting State Scans**

Resting state scans were collected at participants’ initial visit for decoder construction as well as their final post-treatment assessment visit. MRI scans were acquired on a 3-T whole body scanner (Siemens, Prisma) with a 32-channel head coil. Data was collected using a multi-band sequence in the posterior (P) to anterior (A) direction with an acceleration factor of 8. Voxel size was 2.0x2.0x2.0mm3 with a field of view of 208x208mm2. Images were collected from 72 interleaved slices with a TR of 800ms, TE of 37.00 ms, and a flip angle of 52 degrees. Anatomical scans came from a T1-weighted sequence with volumetric navigators (vNAV) with prospective motion correction (TR: 2500ms/TI: 1000ms/Flip Angle: 8.0 degrees/Voxel Size: 0.8x0.8x0.8mm/Matrix Size: 256x256/Num. Slices: 208/Slice Thickness: 0.8mm).

*Processing*

Resting state scans were first distortion corrected using FSL topup (cite) and opposite direction phase-encoded spin echo field maps. Brain extraction of the anatomical image was conducted using bet (cite). Preprocessing and ICA-decomposition was conducted using FSL FEAT and AROMA (cite). During FEAT, data were motion corrected using mcflirt (cite) and brain extracted through bet (cite). Spatial smoothing occurred with a Gaussian kernel of FWHM of 4.0. No highpass/lowpass filtering occurred at this stage. The functional scans were then registered to the standard MNI space and refined with nonlinear registration using FLIRT and FNIRT (cite), respectively. A high-contrast single-band reference image collected at the beginning of each run allowed us to improve registration of multi-band images.

Next, data were denoised using FSL Automatic Removal Of Motion Artifacts (AROMA, cite) and anatomical segmentation was conducted using FSL FAST (cite). Afterwards, additional motion and physiological (i.e. white matter, CSF) noise were regressed out of the data using the 3dTproject function from the AFNI toolbox (cite). A bandpass filter was also applied to remove frequencies below .008 and above .08 and then the filtered data were fitted to the standard space using the FSL applywarp function (cite).

Following this, group-ICA was conducted using FSL MELODIC (cite) to generate independent components shared across all participants for pre-assessment and post-assessment each. We limited the number of components to 20 following previous research as well as to limit the number of comparisons and in order to prevent splitting of components into subcomponents (see [here](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6869246/) for citation). Group-level functional resting state networks were correlated between pre- and post-assessment using the fslcc function of FSL and were visually inspected to confirm similarity while thresholded at Z = 3.2, revealing four matching resting state networks. After group-ICA, we decomposed each group-level component into subject-specific components using dual regression in FSL (cite). Individual participant-specific spatial maps for each of the four networks were subtracted post-treatment to pre-treatment and subjected to a non-parametric one-sample t-test using FSL randomise with 5000 permutations, threshold-free cluster enhancement, and variance smoothing. Results were corrected for these four comparisons using a Bonferroni correction. This resulted in one network with significant connectivity differences pre- to post-assessment.

We created a mask of significant connectivity changes by taking the corrected p-value map, thresholding at a p-value of .9875, and binarizing. We then extracted the average ICA connectivity estimates from the masked subject level maps using FSL’S fslmeants, which were then used in a linear regression.

**Multi-voxel Neuro-reinforcement**

*Task*

Between the pre-treatment and post-treatment assessment, participants underwent either 1, 3, or 5 sessions of neuro-reinforcement. During these sessions, participants were asked to “[think about] whatever [they] can to try and get the best score”, during which a neuro-reinforcement method (cite) would reward them based on successful activation of their phobic image category. The score participants received directly corresponded to the real-time output of the decoder constructed for that participant’s specific phobias.

Each run of neuro-reinforcement began with a 50 second rest period to allow for the scanner image reconstruction processing to catch up to real time. Then, an additional 10 seconds of rest to capture a baseline level of BOLD signal, after which the 16 trials of neuro-reinforcement began. Each trial started with six seconds of rest and six seconds of “induction” where the participant engaged in various thinking strategies in order to maximize their feedback score. Once the induction period ended, the participant would rest for four seconds while the decoder calculated their score which was then shown to the participant as a green disk for two seconds. The size of the disk represented the likelihood estimate of a 100% likelihood being associated with a maximum disc size (i.e. the boundary of the disk) whereas a 0% likelihood was associated with no disk display. Participants were also told that the size of the disk corresponded to the amount of their monetary reward at the end of the run, with their reward based on the average feedback score during the run (e.g. 50% average score meant 50%, or $3.00 of the $6.00 maximum payout per run). If participants were able to maintain an average score of 70% or above for three trials in a row, they would receive a high score streak bonus, or an extra $2.00. These high score streak bonuses were also visually shown upon achievement to the participants as a blue disk accompanied by a congratulatory message.

**Pre-/Post-Neuro-Reinforcement Assessments**

Phobic participants completed pre- and post-treatment fMRI assessments in which they completed a fear test while their BOLD signals were recorded. Each trial of the task began with a 3 – 7 second presentation of the fixation cross followed by a static image for 6 seconds. The images shown were chosen from either their target phobic animal, control phobic animal, or a randomly selected non-phobic animal or object as determined by diagnostic interview. Afterwards, participants saw a blank screen for 4 – 12 seconds before being prompted to rate how fearful they found the image to be on a 7-point Likert scale. Each administration of the task consisted of two runs of 15 images each, composed of 5 target phobic images, 5 control phobic images, and 2 – 3 images from a neutral animal/object, with the remaining photos randomly selected from the aforementioned categories.

*Amygdala Processing*

Fear test fMRI data were distortion corrected using opposite phase-encoded direction spin echo field map sequences as well as FSL topup (cite). Next, brain extraction of the T1 anatomical image was completed using bet (cite). Preprocessing and ICA-decomposition were completed using FSL melodic and FEAT (cite), where data were motion corrected through mcflirt (cite), brain extracted using bet (cite), spatially smoothed with a Gaussian kernel of FWHM 4.0mm, then intensity normalized and highpass filtered with a gaussian-weighted least squares straight line fitting with sigma=50.0s. Registration to the standard MNI space was performed using FLIRT and refined through nonlinear registration using FNIRT (cite). A high-contract single-band reference image collected at the beginning of each run allowed for an improved registration of the multi-band images.

After ICA-decomposition, components were visually inspected and manually cleaned to remove movement artifacts and other noise. Artifact Detection Tools (ART, cite) was used to further account for movement by generating motion regressors and identify outlier timepoints to be censored. Next, first-level GLMs were calculated using SPM12 with a temporal derivative to control for slice-timing differences. The target phobia, control phobia, neutral animal, and neutral object images were fit as regressors with a duration of 0 seconds to model event-related responses. Similar to previous methods (cite), BOLD data from only the first two appearances of each target phobia and control phobia image were analyzed.

Freesurfer segmentation of the T1 anatomical image created bilateral amygdala masks which were then transformed into the participant’s native functional space. We used marsbar (cite) to extract average parameter estimates from the amygdala masks. The estimates were then corrected by subtracting the average amygdala response to the neutral animal from the target and control phobias within runs. Finally, these corrected amygdala responses were averaged across runs for both pre/post assessments.

Other things – not needed now but saving just in case:

**Decoder Construction – may not be necessary**

A between-subject machine learning decoder was constructed using hyperalignment, a functional alignment method, from healthy controls (N = 22) prior to neuro-reinforcement. Healthy controls viewed a standardized image set of 3600 comprising 40 categories of animals and objects (e.g. spiders, sharks, frogs, bees) while participants meeting for specific phobias viewed a modified image set that excluded images of their diagnosed phobic animals. For neuro-reinforcement, subject-specific decoders were constructed using surrogate data (see main DecNef paper cite for more detail).

as well as contains the Ventral Temporal Area, the region primarily targeted by our neuroreinforcement protocol.