Submission Number: 4704

Measuring Negative Attention Bias in Depression Using Differential Brain Decoding

Pisner[1,2], D., Shumake, J.[1,2], Beevers[1,2], C., Schnyer[1,2,3] D.

[1]The University of Texas at Austin Department of Psychology
[2]The University of Texas at Austin Institute for Mental Health Research (IMHR)
[3]The University of Texas at Austin Imaging Research Center (IRC)

INTRODUCTION

- Preferential attendance towards negative affect is a key maintaining factor in depression[1].
- Characterizing this 'negative attention bias' on an individual basis, however, has proven challenging if not outright unreliable[2]. To fill this gap, the attentional brain decoding capabilities of Multivoxel Pattern Analysis (MVPA) might be used.
- In the present study, we demonstrate MVPA decoding of negative attention bias using data from an fMRI attention task.
- By auto-tuning pairs of MVPA classifiers to decode attention to scenes vs. faces, we derive a novel index of negative attentional bias based on paired differential classifier performance when the face stimuli are sad as compared to neutral-valenced.

METHODS

Participants

- Data from N=61 depressed young adults ages 18-39 with 48 females was analyzed as part of an ongoing treatment study of depression[3].

Imaging Acquisition Parameters and Preprocessing

- fMRI data was acquired from a 3T Siemens Skyra MRI with 32-channel head coil (TR/TE=1500/30ms, 66 axial slices, MB factor = 3, FOV = 22mm, flip angle = 71, voxel size = 2x2x2 mm).
- Data was preprocessed using fMRIprep[4]. Each scan was subsequently standardized, detrended, and scrubbed of motion.

In-Scanner Sustained Attention to Response Task (SART)

- During fMRI acquisition, participants underwent a task used to probe attention to scenes vs. faces under conditions of high attentional load. The task involved eight separate runs, across two visits, of 40 alternating blocks (60 seconds each) of images (2 seconds each) with overlapping scenes and faces from the Karolinska Directed Emotional Faces (KDEF)[5].
- Participants were asked to attend to scenes (50%) and faces (50%), pressing a button each time an image pair included an image matching a target category (e.g. outside/inside places, male/female faces)(*See Figure 1*).

A: Neutral Classification:
- 1 = Scene targets with neutral face distractors
- 0 = Neutral face targets with scene distractors

B: Sad Classification:
- 1 = Scene targets with sad face distractors
- 0 = Neutral face targets with scene distractors

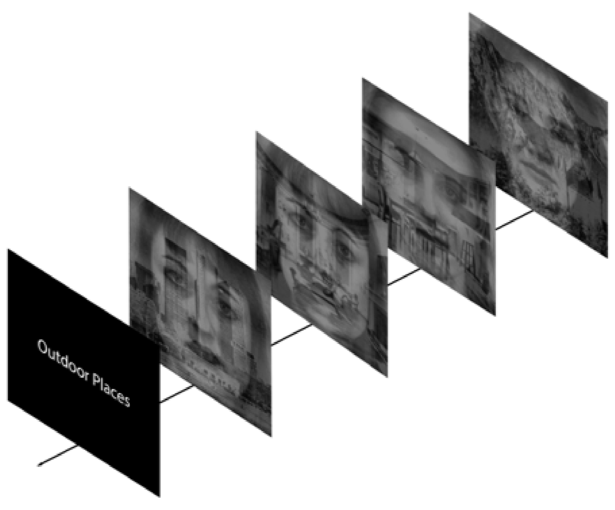


Figure 1: SART fMRI task

- For each condition, a D' sensitivity Index was calculated from the task behavioral data as a measure of task performance based on number of hits, false alarms, misses, and correct rejections.

Behavioral Measures

- Hamilton Depression Inventory (HDRS)[6]—a standard interview-based assessment of depression symptom severity over the past two weeks.
- Eye tracking dot-probe task[7] – 192 trials with 12 pairs of sad/neutral images and 12 pairs of happy/neutral images from KDEF were randomly presented four times each within each of two 96-trial blocks. Mean bias was calculated from eye tracking data based on the percentage of trials where gaze time for sad stimuli > gaze time for neutral stimuli.

Multivoxel Pattern Analysis (MVPA) for Decoding Attention to Scenes Versus Faces

- The time-series from each subject's set of eight fMRI runs were extracted into 2D feature-vectors, restricted to grey-matter and a composite ROI mask generated from task-relevant contrast images obtained from NeuroVault[8,9,10,11] (*See figure 2*).

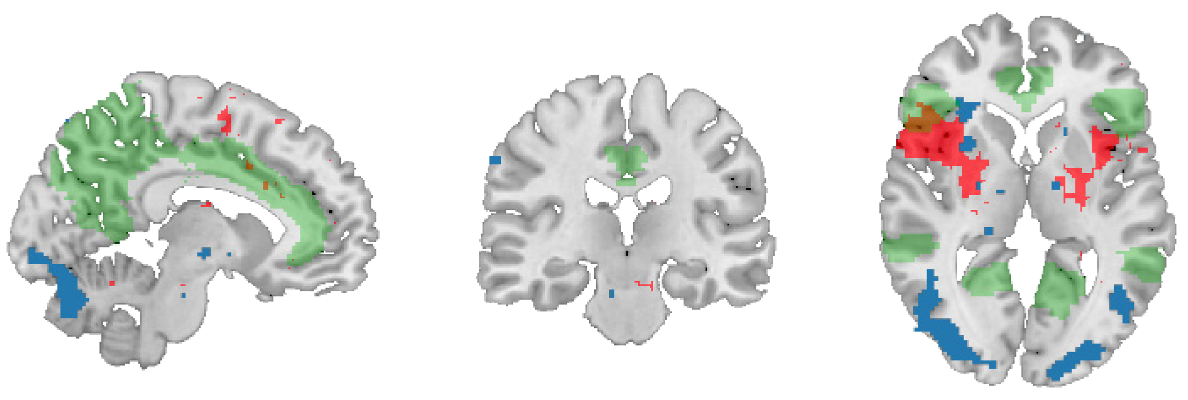


Figure 2: ROI's included a FFA-PPA scene-face network (blue), a sadness network (red), and the Attention Control Network (ACN) (green).

- For each subject, the feature vectors were fragmented into 40-trial time-series blocks which were then randomly shuffled and concatenated into feature vectors for classifier training using a pipeline developed from tools in Scikit-Learn[12].
- Using a stratified 80-20 train-test split, we trained two L2-logistic Regression classifiers (one for each of conditions A and B) per subject. A 10-fold cross-validation with grid-search and F1-scoring was used to tune C and tolerance based on a 1 SE rule. The k=200 most informative features were selected based on F-ANOVA[13] (*See Figure 3*), and class-weighted accuracy was reported.

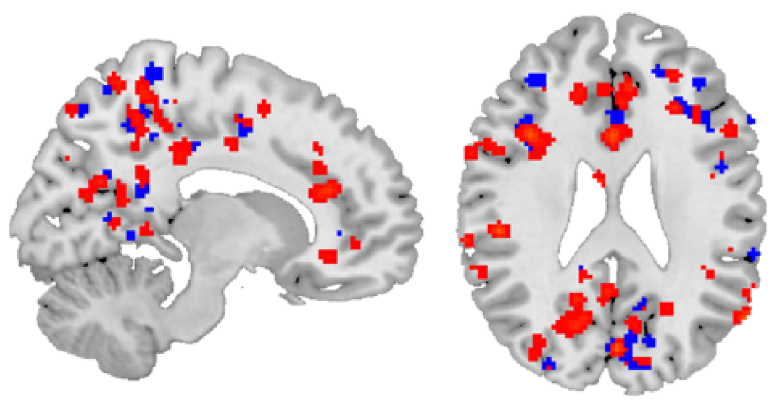


Figure 3: Mean feature importance map across subjects where features most relevant for condition A and B are depicted in red and blue, respectively.

Brain-Behavioral Synchrony Differential Across Sad Relative to Neutral Face Conditions

- In R 3.3.1, paired T-tests were used to determine whether the mean differences in task D' and classifier accuracy, across conditions A and B, was (not) equal to zero.
- To then derive a measure for negative attention bias, we calculated the residual paired difference[14] from each of the classifier accuracy and D' measures, which is the paired difference minus the mean of all paired differences across conditions A and B.
- Using the residual paired difference measures, we ultimately explored associations with depression severity and the dot-probe measure of negative attention bias.

RESULTS

- Classification accuracy was higher for sad (M=93%, SD=0.03) relative to neutral-face (M=92%, SD=0.04) conditions with mean paired differences that were significantly different from zero ($t(61) = 2.02$, $p < 0.05$) (*See Figure 4*).

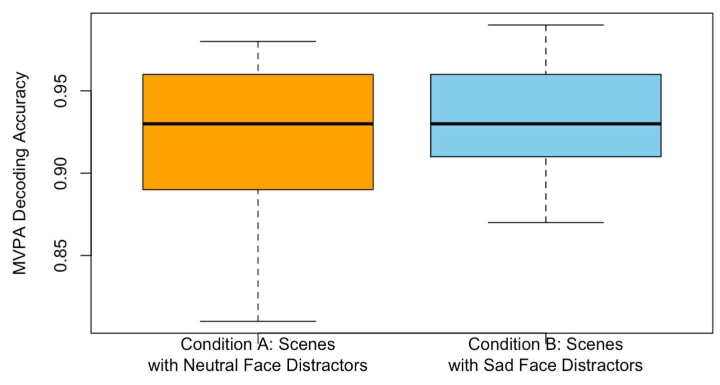
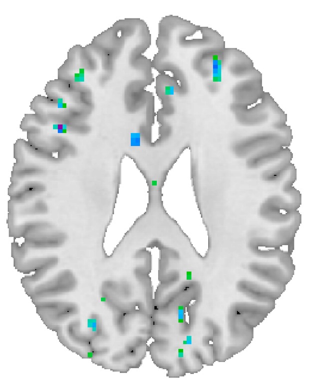


Figure 4: Paired difference in F-scores between classifier A (orange) and classifier B (blue)

- Relative MVPA classification accuracy across conditions A and B was driven predominantly by a fronto-parietal attention network (*See Figure 5*).



Regional Feature Importance
Middle Frontal Gyrus
Frontal Pole
Precentral Gyrus
Postcentral Gyrus
Precuneous Cortex:
Anterior Cingulate Gyrus:
Superior Lateral Occipital Cortex
Posterior Cingulate Gyrus:
Occipital Fusiform Gyrus:
Paracingulate Gyrus:

Figure 5: Sad-Face Feature Importance Map

- Classification accuracy was marginally correlated with D' for the neutral-face but not sad-face condition ($F(1,60)=5.27$, $R^2=0.08$, $p<0.05$).
- The residual paired difference scores based on classification accuracies followed a normal distribution (*See Figure 6*).
- Classification residual paired difference scores were also negatively correlated with task behavioral performance ($F(1,59)=4.63$, $R^2=0.07$, $p<0.05$).

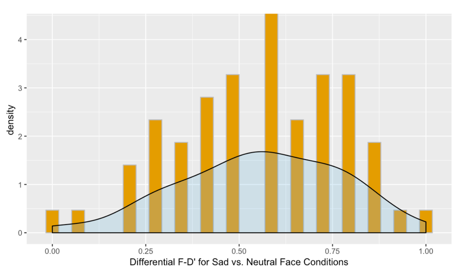


Figure 6: Distribution of residual paired Difference scores in classification accuracy

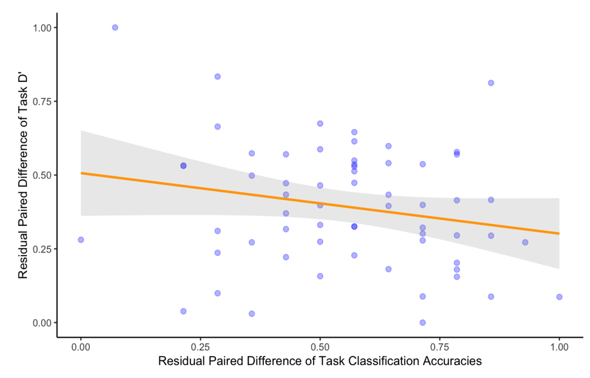


Figure 7

- When exploring convergent validity with depression severity and a more traditional dot-probe measure of negative attention bias, we found positive correlations with residual paired difference of classification accuracies, but not with that of task behavioral D' ($F(1,60)=4.16$, $R^2=0.06$, $p<0.05$) (*See Figure 8*).

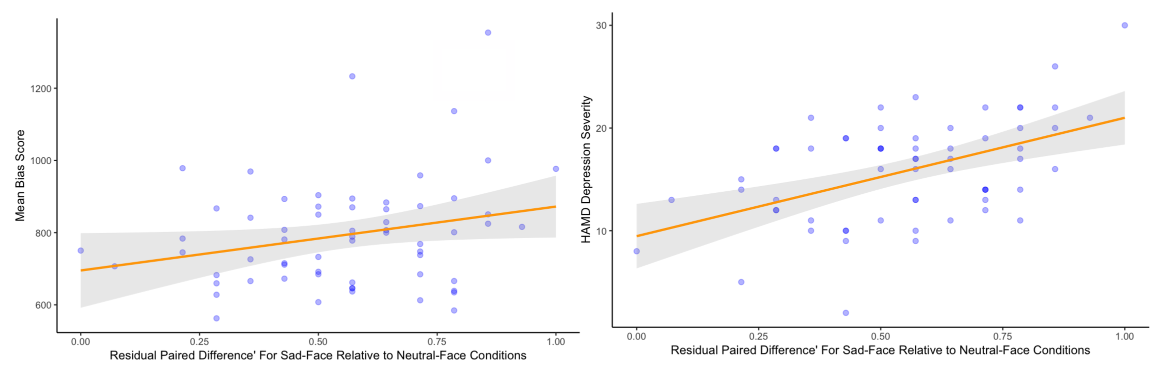


Figure 8: Residual paired difference scores in classification accuracy across conditions were positively correlated with mean bias (left) and depression severity (right).

CONCLUSION

- The present study employed MVPA to decode attention to scenes versus faces based on the brain activation patterns of depressed individuals during an fMRI task.
- A 'negative attention bias' score was calculated for each subject based on the differential performance of classifier pairs for which the face stimuli were either neutral or sad valenced, respectively.
- The difference in feature performance across conditions was driven predominantly by nodes of a fronto-parietal attention control network.
- The MVPA-based measure of negative attention bias, which represents the magnitude of attentional distractibility in response to sad-neutral valence variation perceived during the task, demonstrated convergent validity with measures of depression severity and a dot-probe measure of negative attention bias.