Neural changes with attention bias modification for anxiety: a randomized trial

Jennifer C. Britton,1,2 Jenna G. Suway,3,4 Michelle A. Clementi,1,5 Nathan A. Fox,3 and Yair Bar-Haim6

1Section on Development and Affective Neuroscience, National Institute of Mental Health, Bethesda, MD, 20892 USA, 2Department of Psychology, University of Miami, Coral Gables, FL, 33146 USA, 3Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, 20742 USA, 4Joint Doctoral Program in Clinical Psychology, San Diego State University/University of California, San Diego, San Diego, CA, 92120 USA, 5Department of Psychology, University of Houston, Houston, TX, 77204 USA, and 6School of Psychological Sciences and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, 69978 Israel

Attention bias modification (ABM) procedures typically reduce anxiety symptoms, yet little is known about the neural changes associated with this behavioral treatment. Healthy adults with high social anxiety symptoms (n = 53) were randomized to receive either active or placebo ABM. Unlike placebo ABM, active ABM aimed to shift attention away from threat. Using the dot-probe task, threat-related attention bias was measured during magnetic resonance imaging before and after acute and extended training over 4 weeks. A subset of participants completed all procedures (n = 30, 15 per group). Group differences in neural activation were identified using standard analyses. Linear regression tested predictive factors of symptom reduction (i.e., training group, baseline indices of threat bias). The active and placebo groups exhibited different patterns of right and left amygdala activation with training. Across all participants irrespective of group, individuals with greater left amygdala activation in the threat-bias contrast prior to training exhibited greater symptom reduction. After accounting for baseline amygdala activation, greater symptom reduction was associated with assignment to the active training group. Greater left amygdala activation at baseline predicted reductions in social anxiety symptoms following ABM. Further research is needed to clarify brain-behavior mechanisms associated with ABM training.

Keywords: attention training; dot-probe; fMRI; amygdala; anxiety; treatment

INTRODUCTION

Threatening cues capture the attention of individuals with anxiety to a greater extent than for non-anxious individuals (Bar-Haim et al., 2007; Armstrong and Olatunji, 2012). Attention bias modification (ABM), which targets threat bias in attention, typically reduces anxiety symptoms (Hakamata et al., 2010; Beard et al., 2012). However, changes in underlying neural circuitry following ABM remain unclear. This functional magnetic resonance imaging (fMRI) study examined changes in neural circuitry before and after ABM in adults with elevated symptoms of social anxiety.

To date, few studies have examined neural changes following ABM designed to shift attention away from threat. In an event-related potential (ERP) study, adults with high anxiety scores showed decreased P2 and P3 and increased N2 amplitudes following such ABM training (Eldar and Bar-Haim, 2010). Although these ERP results speak to the chronometry of neural changes, they are limited in spatial resolution. fMRI studies in healthy adults shed light on this issue. One study found greater dorsolateral and ventrolateral prefrontal cortex (PFC) activation after ABM training relative to placebo training (Browning et al., 2010); the other reported reduced right and left amygdala, insula and subgenual anterior cingulate (sgACC) activation as well as increased PFC and visual cortex activation following active ABM training (Taylor et al., 2013). Although the results should be interpreted with caution based on the statistical thresholds used, another fMRI study in adults with social anxiety disorder reported increased right and left amygdala activation following internet-delivered ABM relative to baseline assessment, and increased left amygdala activation following treatment was correlated with anxiety reduction (Mansson et al., 2013). Although this work suggests that ABM might influence fronto-amygdala function, the results are difficult to interpret due to either lack of a baseline scan (Browning et al., 2010) or lack of an ABM control group (Mansson et al., 2013; Taylor et al., 2013). Here, we examined changes from before to after training in both active ABM and placebo training groups.

We examined the degree to which ABM specifically influences brain regions that support top-down or bottom-up mechanisms (Bishop, 2008; Heeren et al., 2013). Top-down mechanisms involve voluntary or effortful control (Ochsner et al., 2012), processes thought to engage the PFC. Bottom-up mechanisms involve threat sensitivity or threat salience (Davis and Whalen, 2001; Phan et al., 2002), processes thought to engage the amygdala. Thus, prior fMRI studies implicating the PFC in threat-attention interactions (Monk et al., 2006) and the amygdala in threat processing (Phan et al., 2006; Monk et al., 2008) suggest a role for top-down and bottom-up mechanisms, respectively, in threat-related attention in individuals with anxiety. It remains unclear whether ABM alters top-down attention control and/or bottom-up attention processes with the supporting neural architecture. Prior research examining neural changes on the dot-probe task before and after cognitive behavioral therapy (CBT) or pharmacological treatment found changes in the ventrolateral PFC and right and left amygdala, suggesting that various cognitive mechanisms and brain structures could play a role in ABM (Maslowsky et al., 2010). Finally, individual differences in top-down or bottom-up mechanisms, as well as their neural correlates, also appear to predict response to CBT and pharmacological treatment (Faria et al., 2012;...
This raises the possibility that baseline functioning and neural activation in the PFC and amygdala would also predict response to ABM.

In this fMRI study, attention bias of healthy adults with high self-reported social anxiety was assessed before and after randomization to active ABM or placebo training. These data were used to test two specific hypotheses. First, we hypothesized that attention bias and neural activation would more robustly change with active ABM than with placebo training. Specifically, compared with highly anxious adults receiving placebo training, those individuals receiving active training were expected to show greater behavioral and neural changes in the PFC and amygdala. Second, we predicted that baseline attention bias and neural activation would predict symptom reduction post-training. Finally, because previous studies were mixed in relation to the laterality of amygdala activation underlying ABM-related changes in anxiety, we separately tested our predictions using right and left amygdala activation.

**MATERIALS AND METHODS**

**Participants**

Healthy socially anxious adults (Liebowitz Social Anxiety Scale (LSAS) (Fresco et al., 2001) \( \geq 50 \)) were recruited from the University of Maryland and the National Institute of Mental Health (NIMH) subject pools from October 2010 to February 2012. The LSAS cutoff approximated the upper quartile of scores collected from college student samples. LSAS scores within this range are more likely to characterize the generalized concerns of social anxiety (Mennin et al., 2002; Rytwinski et al., 2009). The trial ended when complete MRI data from 15 individuals in each group was achieved to reach minimal sample sizes typically used for neuroimaging research (Carter et al., 2008).

Written informed consent was obtained from all participants and procedures were conducted in accordance with both University of Maryland and NIMH Institutional Review Boards.

All participants were paid volunteers 18–30 years old, English speaking, and had normal or corrected-to-normal vision. To be eligible, each participant had an IQ > 70 (Wechsler Abbreviated Scale of Intelligence, WASI (Wechsler, 1999)), and was free of chronic medical illness, any psychotropic medication and contraindications to MRI. Any current Axis I psychiatric disorder was exclusionary. The rationale for studying non-diagnosed but symptomatic subjects related to ethical concerns about withholding effective treatment from subjects with an anxiety disorder. To assess psychopathology, experienced clinicians administered the Structured Clinical Interview for DSM-IV (First et al., 2002).

In addition, indices of anxiety and depressive symptom severity were measured using the LSAS, Spielberger’s Trait Anxiety Inventory (STAI) (Spielberger, 1983), Fear of Negative Evaluations Scale (Watson and Friend, 1969), and Beck Depression Index (Beck et al., 1996).

Of 138 potential participants responding to initial requests for participation and meeting eligibility criteria to participate in the study based on phone screen, 74 were screened at the NIMH for psychopathology and 53 enrolled. After blocking for gender and LSAS score within ranges 50–65, 66–80 and above 80, participants were randomly assigned to an active ABM group (\( n = 26 \)) or placebo training group (\( n = 27 \)). Within each category, an online random number generator repeatedly created a closed sequence of eight values assigned to either active or placebo (J.C.B.). NIMH staff enrolled participants and assigned the intervention based on the random assignment (M.A.C.).

Participants and staff administering ABM training were blind to the treatment condition. As in previous work (Britton et al., 2012, 2013), participants were excluded for non-compliance. Participants were excluded if accuracy rates on the dot-probe task were below 75% or if motion was excessive (more than 75% of MRI data had >3 mm translation/3° rotation in any direction from initial position). The latter criterion attempts to limit the overall movement to a voxel. Final analyses included those individuals who completed all procedures and had usable fMRI data from both assessments. These criteria yielded 15 participants in each group (Figure 1). Group characteristics for the final sample are presented in Table 1. Groups were well matched on demographics, symptoms and task performance at baseline (all \( P > 0.1 \)). See Supplemental Information for more details.

**Procedures**

During the first MRI visit at NIMH, participants completed a dot-probe task (Amir et al., 2009) to measure attention bias before (pre-training) and after acute-ABM training while in the MRI scanner. Between attention bias measurements, participants completed training according to group assignment (i.e., active ABM or placebo) while remaining in the MRI scanner. Following the scan, participants completed extended ABM training outside of the scanner at the University of Maryland (twice weekly over 4 weeks). After the extended training, participants completed a second MRI scan at the NIMH, and attention bias and anxiety were again measured. Therefore, all three attention bias assessments (i.e., pre-training, post-acute training and post-extended training) were conducted in the MRI scanner. Acute training occurred during the first MRI visit and extended training was conducted behaviorally over a 4-week period.

**Dot-probe task**

In the dot-probe task, an angry and a neutral face of the same individual were presented simultaneously. In addition, trials with pairs of neutral faces (neutral trials) were used as a control condition. The face stimuli were presented vertically in black and white and consisted of 12 identities selected from the NimStim Face Stimulus set (Tottenham et al., 2009). Immediately following the face pair, a probe replaced either the angry face (threat congruent trial) or the neutral face (threat incongruent trial). Participants indicated the probe letter (E or F) via button press as quickly and as accurately as possible.

The pre-training, post-acute training and post-extended training dot-probe task consisted of 144 trials (48 threat congruent, 48 threat incongruent, 48 neutral trials). Each trial began with a central fixation cross (+) for 500 ms, then a pair of faces (angry–neutral, neutral–neutral) appeared vertically for 500 ms. A 400 ms probe (E or F) replaced the angry or neutral face with equal probability. To reduce anticipation effects in the behavioral response, trials were separated by a variable length inter-stimulus interval, which averaged 1100 ms and ranged from 900 to 1300 ms. In addition, 48 fixation trials per run were presented to be able to deconvolve the hemodynamic signal.

**Attention training task**

Using the dot-probe task, active ABM was designed to train participants to attend away from threat (Bar-Haim, 2010). Thus, the probe always replaced the neutral face. In the placebo version, the probe replaced the angry face and neutral face with equal probability. Pairs of neutral faces also were presented in both versions as a control condition (neutral trials). Of note, the four face stimuli in ABM/placebo task (Matsumoto and Ekman, 1989) were different from the face stimuli presented during assessment to potentially demonstrate generalization of training.

During acute training while in the scanner, 480 training trials were completed over three runs (160 trials each). ABM consisted of 384 threat incongruent trials and 96 neutral trials, whereas the placebo training had 192 threat congruent, 192 threat incongruent, and 96 neutral trials. In addition, there were 96 fixation trials. During extended training conducted behaviorally, eight shorter training sessions with
Assessed for eligibility (n=74)
  Excluded (n=21)
    • Not meeting inclusion criteria (n=20)
    • Study ended (n=1)
Randomized (n=53)
Allocated to active ABM intervention (n=26)
  • Received allocated intervention (n=26)
Allocated to placebo intervention (n=27)
  • Received allocated intervention (n=27)
Follow-Up
  • Discontinued intervention (n=3)
Analysis
  • Discontinued intervention (n=6)
Analyzed (n=15)
  • Excluded for poor accuracy (n=4)
  • Excluded for motion (n=3)
  • Excluded for technical reasons (n=1)
  • Excluded for poor accuracy (n=2)
  • Excluded for motion (n=6)

Fig. 1 Study flow.

Table 1 Group characteristics of individuals included in fMRI analysis

<table>
<thead>
<tr>
<th></th>
<th>Active training</th>
<th>Placebo training</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Age (years)</td>
<td>21.8 ± 2.9</td>
<td>22.1 ± 3.2</td>
</tr>
<tr>
<td>Gender (males)</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>IQ</td>
<td>119.1 ± 8.4</td>
<td>118.5 ± 9.5</td>
</tr>
<tr>
<td><strong>Baseline symptoms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSAS average screen</td>
<td>52.0 ± 16.3</td>
<td>54.6 ± 17.3</td>
</tr>
<tr>
<td>STAI trait</td>
<td>34.3 ± 8.8</td>
<td>36.4 ± 10.5</td>
</tr>
<tr>
<td>BDI</td>
<td>3.8 ± 3.8</td>
<td>4.6 ± 4.6</td>
</tr>
<tr>
<td>FNE</td>
<td>13.1 ± 8.1</td>
<td>11.0 ± 8.0</td>
</tr>
<tr>
<td><strong>Dot-probe performance</strong></td>
<td>Pre-training</td>
<td>Post-acute training</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>94.8 ± 3.4</td>
<td>91.2 ± 4.8</td>
</tr>
<tr>
<td>Threat bias</td>
<td>−2.5 ± 22.4</td>
<td>−5.4 ± 27.0</td>
</tr>
<tr>
<td>RT congruent</td>
<td>658.3 ± 85.3</td>
<td>637.3 ± 90.3</td>
</tr>
<tr>
<td>RT incongruent</td>
<td>655.8 ± 84.6</td>
<td>631.9 ± 91.6</td>
</tr>
<tr>
<td>RT neutral</td>
<td>657.4 ± 93.2</td>
<td>632.6 ± 93.5</td>
</tr>
</tbody>
</table>

Means and S.D. IQ = intelligence quotient based on Wechsler Abbreviated Scale of Intelligence (WASI); LSAS, Liebowitz Social Anxiety Scale; STAI, Spielberger Trait Anxiety Inventory; BDI, Beck Depression Index; FNE, Fear of Negative Evaluations Scale; RT, reaction time (ms).
160 trials each were completed twice weekly over 4 weeks (effectively yielding another 1280 trials).

**Apparatus and acquisition**

Stimuli were presented via standardized software (E-Prime, Inc., 2.0, Pittsburgh, PA) on a PC computer. In the MRI scanner, images were projected onto a screen and viewed by participants through a mirror mounted on the head coil.

For each participant, MRI data were collected with a 3.0 T General Electric Signa system scanner (Waukesha, WA) and an eight-channel gradient head coil. After shimming, a T2*-sensitive gradient echo pulse sequence was used for functional imaging (TR = 2300 ms, TE = 25 ms, flip angle = 90°, FOV = 24 cm, 96 x 96 matrix, 36 contiguous 2.6 mm interleaved axial slices). To reach longitudinal magnetization equilibrium, four initial acquisition images were collected prior to task onset. Finally, a high-resolution T1-weighted volumetric scan of the whole brain was acquired using a magnetization prepared gradient echo sequence (124 1.2 mm axial slices; FOV = 220 mm; NEX = 1; TR = TE = min; matrix = 256 x 192; TI = 725 ms; bandwidth = 31.25 kHz for 256 pixels). This anatomical scan was used for coregistration and normalization procedures.

**Data analysis**

**Changes in social anxiety symptoms**

Changes in LSAS scores across the procedures were analyzed using a Repeated Measures Analysis of Variance (ANOVA) to examine interactions between Group (active, placebo) and Time (averaged screening, scan 2). Significance was determined using $\alpha = 0.05$.

**Attention bias measures**

Attention bias was measured before training, after acute training and after extended training. For each measurement, bias scores were calculated by subtracting the reaction time on threat congruent trials from threat incongruent trials. Positive bias scores reflect a bias towards threat and negative bias scores reflect a bias away from threat. Attention bias calculations excluded trials with incorrect responses, trials with reaction times >150 ms or >2000 ms, and trials with reaction times >2.5 s.d. outside of the mean (based on correct trials with reaction times between 150 and 2000 ms). These criteria were applied separately to each condition (threat congruent, threat incongruent, neutral) within each time point (pre-training, following acute training, following extended training).

Behavioral indices were analyzed using a Repeated Measures ANOVA to test interactions between Group (active, placebo) and Time (pre-training, following acute-training and following extended training). Significance was determined using $\alpha = 0.05$ and Bonferroni correction was applied, when appropriate.

**Brain activation**

fMRI data were preprocessed using Analysis of Functional Neuroimages (AFNI, http://afni.nimh.nih.gov/afni). Pre-processing included slice- timing correction and realignment. The reference image for realignment corresponded to the first image of each series. The functional series was coregistered to the anatomical image and normalized into Talairach space using automated routines (Talairach and Tournoux, 1988). The data were smoothed using a 6 mm full-width-at-half maximum isotropic Gaussian filter and scaled, converting the data to percent signal change relative to baseline (i.e., fixation). After all pre-processing steps, the resulting images contained isotropic 2.5-mm voxels.

Separate general linear models were created for each scan visit. Individual models contained only bias assessment data. For all models, regressors coded for threat congruent, threat incongruent and neutral conditions for each assessment run. These regressors of interest included only correct trials. Incorrect trials or trials with reaction times outside of the criteria limits were included as a separate regressor of no interest. Finally, six motion regressors corresponding to translation and rotation in each xyz direction and regressors modeling low-frequency drift were included. A series of estimated betas, one for each regressor, was generated to minimize the error term within the model.

A whole-brain voxel-wise analysis was conducted using random effects analysis. The beta coefficients generated from individual analyses were entered into a second-level model (i.e., AFNI GroupAna). The model examined interactions between Group (active, placebo), Time (pre-training, following acute-training and following extended training) and Condition (threat congruent, threat incongruent, neutral).

To correct for multiple comparisons in our omnibus analysis, we performed a family-wise error approach at the cluster level through 1000 Monte Carlo simulation using AFNI’s AlphaSim program. The cluster threshold was set at $\alpha = 0.05$ with a voxel-wise $P$-value of 0.005 and estimated smoothness values in our data, that is, $\sim 8.6$ mm. Statistical significance was defined as 80 voxels (1250 mm$^3$) for whole brain (92165 voxels, 1440078 mm$^3$). To restrict the search territory, regional masks were used for a priori regions [i.e., amygdala (117 voxels, 1828 mm$^3$), inferior frontal gyrus (733 voxels, 11453 mm$^3$)]. Based on these calculations, cluster sizes of 2 voxels (31 mm$^3$) for the amygdala and 14 voxels (219 mm$^3$) for the inferior frontal gyrus were needed to yield a corrected $p < 0.05$ level. Coordinates are presented using Left-Posterior-Inferior (LPI) coordinate system.

Although statistical inference is based on the whole-brain voxel-wise analysis, *post-hoc* analyses were conducted to decompose the Group x Time x Condition interaction within each significant cluster. Similar approaches have been used widely as a means to describe the effects of complex interactions in whole-brain analysis (Shin et al., 2011; Killgore et al., 2014; White et al., 2014; Wiggins et al., 2014). Based on the group level analysis, significant voxels in each cluster were identified. These significant voxels included all voxels above a threshold of $p < 0.005$ that lay in a cluster surpassing the more stringent, corrected statistical threshold. Across these voxels, the percent signal-change values, relative to fixation, were averaged for each condition for each participant. *Post-hoc* analyses were conducted on these averaged values within each condition, based on the threat bias contrast (i.e., threat incongruent–threat congruent). These analyses aimed to investigate the patterns across conditions as a function of group, time and condition. These extracted values also served to predict changes in social anxiety.

**Predicting reduction in social anxiety using pre-training baseline measures**

Using a step-wise linear regression, group and baseline measures (i.e., pre-training attention bias scores and pre-training right or left amygdala activation in response to threat bias) were used to predict overall symptom reduction (i.e., change in LSAS scores between screening and post-extended training). All variables were mean centered across participants. Separate analyses were conducted using the right and left amygdala.

After conducting group analyses and appropriate *post-hoc* analyses on the imaging data, a group difference was found in both right and left amygdala activation at baseline. The active group had less right and left amygdala activation in the threat-bias contrast (threat incongruent > threat congruent) at baseline compared with the placebo group [ts(28) > 4.4, $p < 0.001$]. As a result, right or left baseline amygdala activation for this same threat bias contrast was entered into the
prediction models in step 1 to minimize this initial pre-ABM group difference. In step 2, main effects of group (active, placebo) and baseline attention bias were entered. Two-way interactions (Group*Pre-training Attention Bias, Group*Pre-training Amygdala Activation, Pre-training Attention Bias*Pre-training Amygdala Activation) were entered in step 3, and the three-way interactions (Group*Baseline Attention Bias*Pre-training Amygdala Activation) were entered in step 4. Significant effects were determined using $\alpha = 0.05$.

RESULTS

Changes in social anxiety symptoms

Although no group differences or group interactions were found ($p > 0.4$), LSAS scores significantly reduced from screening (active: $52.0 \pm 16.3$, placebo: $54.6 \pm 17.3$) to post-extended training (active: $42.9 \pm 22.5$, placebo: $49.4 \pm 23.6$) in the sample as a whole [$F(1, 28) = 4.3$, $p < 0.05$, $\eta^2_p = 0.13$].

Attention bias measures

No group differences or interactions were detected when examining accuracy and attention bias at any point ($p > 0.6$). No changes in attention bias across time were detected ($p > 0.6$), and neither group showed an attention bias towards or away from threat after acute or extended training ($p > 0.1$) (Table 1).

Brain activation

Whole brain, random-effects analyses indicated three-way interactions (Group $\times$ Time $\times$ Condition) in the right and left amygdala: [right: $(16, -1, -9)$, $54$ voxels ($844 \text{ mm}^3$), $F(4,112) = 4.3$, $p < 0.05$ corrected; left: $(19,1, -11)$, $49$ voxels ($766 \text{ mm}^3$), $F(4,112) = 5.4$, $p < 0.05$ corrected] (Figure 2). Post-hoc analyses were used to decompose this interaction. These analyses showed that, following extended training, right and left amygdala activation in response to the threat bias contrast (threat incongruent > threat congruent) increased in the active ABM group and decreased in the placebo group [Group $\times$ Time interaction: right: $F(2,56) = 12.1$, $p < 0.001$, $\eta^2_p = 0.30$, left: $F(2,56) = 15.9$, $p < 0.001$, $\eta^2_p = 0.36$; follow-up contrasts in each group, $p < 0.001$]. Importantly, as noted above, interpretation of this three-way interaction was complicated by the presence of significant group differences in amygdala activation to threat bias prior to training. This baseline difference reflects the fact that the active group, prior to training, had less right and left amygdala activation in the threat bias contrast than the placebo group. Separate models for right and left amygdala were used to predict symptoms changes. Therefore, this pre-treatment group difference in amygdala activation was minimized by including baseline right or left amygdala activation as step 1 in separate prediction models using left or right amygdala data.

Of note, no significant effects in amygdala activation were detected when examining angry trials (i.e., collapsing angry incongruent and congruent trials) relative to neutral trials (all $p > 0.1$). Furthermore, amygdala activation in response to neutral trials was not significantly different between active and placebo training groups during any of the three assessments (pre-training, post-acute training, post-extended training) (all $p > 0.3$).

No other regions (i.e., prefrontal cortical regions) survived a priori cluster-level correction. However, a Group $\times$ Time $\times$ Condition interaction was complicated by the presence of significant group differences in amygdala activation to threat bias prior to training. This baseline difference reflects the fact that the active group, prior to training, had less right and left amygdala activation in the threat bias contrast than the placebo group. Separate models for right and left amygdala were used to predict symptoms changes. Therefore, this pre-treatment group difference in amygdala activation was minimized by including baseline right or left amygdala activation as step 1 in separate prediction models using left or right amygdala data.

Of note, no significant effects in amygdala activation were detected when examining angry trials (i.e., collapsing angry incongruent and congruent trials) relative to neutral trials (all $p > 0.1$). Furthermore, amygdala activation in response to neutral trials was not significantly different between active and placebo training groups during any of the three assessments (pre-training, post-acute training, post-extended training) (all $p > 0.3$).

No other regions (i.e., prefrontal cortical regions) survived a priori cluster-level correction. However, a Group $\times$ Time $\times$ Condition interaction was complicated by the presence of significant group differences in amygdala activation to threat bias prior to training. This baseline difference reflects the fact that the active group, prior to training, had less right and left amygdala activation in the threat bias contrast than the placebo group. Separate models for right and left amygdala were used to predict symptoms changes. Therefore, this pre-treatment group difference in amygdala activation was minimized by including baseline right or left amygdala activation as step 1 in separate prediction models using left or right amygdala data.

Of note, no significant effects in amygdala activation were detected when examining angry trials (i.e., collapsing angry incongruent and congruent trials) relative to neutral trials (all $p > 0.1$). Furthermore, amygdala activation in response to neutral trials was not significantly different between active and placebo training groups during any of the three assessments (pre-training, post-acute training, post-extended training) (all $p > 0.3$).

Fig. 2 Neural changes associated with ABM training. Whole brain, random-effects analyses indicated three-way interactions (Group $\times$ Time $\times$ Condition) in the right and left amygdala: [right: $(16, -1, -9)$, $54$ voxels ($844 \text{ mm}^3$), $F(4,112) = 4.3$, $p < 0.05$ corrected; left: $(19,1, -11)$, $49$ voxels ($766 \text{ mm}^3$), $F(4,112) = 5.4$, $p < 0.05$ corrected]. To decompose this three-way interaction, the percent signal-change values, relative to fixation, across significant voxels in each whole-brain cluster were averaged for each condition for each participant. Post-hoc analysis examined the threat bias contrast (i.e. threat incongruent > congruent).
interaction was detected in a non-\textit{a priori} region, the cerebellum [(9, −74, −16)], 119 voxels (1859 mm$^3$), $F(4112) = 4.3, p < 0.05$ corrected.

\textbf{Predicting reduction in social anxiety symptoms using pre-training measures}

\textit{Right amygdala}

The predictive model using right amygdala activation was not significant (all $p > 0.1$).

\textit{Left amygdala}

As shown in Table 2, the model using left amygdala activation was significant [$F(6,29) = 2.8, p < 0.04$] and accounted for 42.0% of the variance in LSAS symptom reduction. After accounting for baseline amygdala activation (step 1), group assignment significantly predicted reduction in social anxiety symptoms. As shown in Figure 3, greater symptom reduction was associated with greater baseline left amygdala activation irrespective of group assignment (Unstandardized $\beta = 94.0$, SE = 42.0, $p < 0.04$), and with assignment to the ABM relative to placebo condition (Unstandardized $\beta = 9.5$, SE = 4.1, $p < 0.03$).

Baseline attention bias score did not significantly predict symptom reduction ($p > 0.8$) and individuals showing avoidant biases at baseline did not significantly differ in terms of symptom reduction or changes in amygdala activation (all $p > 0.1$).

\textbf{DISCUSSION}

This study examined both neural and behavioral predictors of social anxiety symptom reduction following ABM training. Two findings emerged. First, greater left amygdala activation to the threat bias contrast at baseline was associated with greater symptom reduction across both training groups. Second, after accounting for group differences in baseline amygdala activation, assignment to one or the other training condition further predicted changes in symptoms, with greater symptom reduction in the active ABM group.

\textbf{Brain activation}

Based on the whole-brain analysis, the active ABM group exhibited increased bilateral amygdala activation to the threat bias contrast following extended training compared with before training; whereas, the placebo group exhibited decreased bilateral amygdala activation following extended training compared with before training. These results in the active ABM training group contrast with recent findings showing pre-to-post training reductions in bilateral amygdala activation in an emotional matching task following active ABM training (Taylor et al., 2013). These potentially conflicting results could relate to differences in the tasks used. Unlike previous work using non-dot-probe tasks, the threat bias contrast used in the current study compared threat incongruent and threat congruent conditions. As attention is trained away from threat in the active ABM group, the threat incongruent condition may become increasingly salient in this trained group, leading to increased amygdala activation over time. Furthermore, threat-incongruent trials in the active ABM group may signal the need to shift attention away from threat, which might be facilitated by amygdala engagement. In addition, the stimulus type that becomes most salient may differ by ABM group, which may lead to opposite patterns of amygdala activation. No such effects are expected.

\begin{table*}[h]
\centering
\begin{tabular}{lcccccc}
\hline
Variable & $R^2$ & $\Delta R^2$ & $F$ & Significant & Unstandardized & $t$ & $p$
\hline
Step 1: & & & & & & & \\
Constant & 0.04 & 0.04 & 1.13 & 0.30 & & & \\
Pre-training left amygdala activation to threat bias contrast & 0.14 & 0.10 & 1.35 & 0.28 & & & \\
Step 2: & & & & & & & \\
Group & & & & & & & \\
Pre-training threat bias & & & & & & & \\
Step 3: & & & & & & & \\
Group $\times$ Pre-training Threat bias & 0.42 & 0.29 & 2.78 & 0.04 & & & \\
Group $\times$ Pre-training Left Amygdala Activation & & & & & & & \\
Pre-training threat bias $\times$ Pre-training left amygdala activation & & & & & & & \\
\hline
\end{tabular}
\caption{Parameters from step regression model predicting symptom reduction change using screening measures}
\end{table*}

Inclusion of the three-way interaction in step 4 was non-significant ($p > 0.4$); therefore, unstandardized $B$, SE of $B$, $t$- and $p$-values associated Step 3 are reported.

\begin{figure}[h]
\centering
\begin{subfigure}{0.45\textwidth}
\centering
\includegraphics[width=\textwidth]{A}
\caption{A}
\end{subfigure}
\hfill
\begin{subfigure}{0.45\textwidth}
\centering
\includegraphics[width=\textwidth]{B}
\caption{B}
\end{subfigure}
\caption{Predictors of symptom reduction. Panel A: Across both groups, left amygdala activation to threat bias contrast prior to training predicted symptom reduction (scores on LSAS from screening to post-extended training). Panel B: The active group showed greater symptom reduction than the placebo group.}
\end{figure}
in the emotional matching task used in prior work (Taylor et al., 2013). Importantly, however, regardless of the direction of change in neural activation the amygdala emerges as a key player in ABM induced threat-related attentional changes across the two studies.

**Predicting reduction in social anxiety symptoms**

Across both the placebo and ABM groups, individuals with greater left amygdala activation to the threat bias contrast prior to training exhibited the greatest reduction in anxiety symptoms following extended training. The current findings of amygdala activation predicting social anxiety symptom reduction were specific to the left amygdala. Several studies report amygdala changes with treatment of anxiety (Furmark et al., 2002; Faria et al., 2012; Hauner et al., 2012; Mansson et al., 2013; Phan et al., 2013; Taylor et al., 2013). Some work shows reduced amygdala activation bilaterally after treatment (Furmark et al., 2002; Faria et al., 2012; Taylor et al., 2013), while others have found lateralized effects either in the left (Hauner et al., 2012; Mansson et al., 2013) or right amygdala (Phan et al., 2013). In one study, amygdala activation decreased more for treatment responders than non-responders, particularly in the right hemisphere (Furmark et al., 2002). In yet another study, different subregions of the amygdala in each hemisphere were associated with treatment response (Faria et al., 2012). Clearly, more work is needed to specify the laterality of these effects.

Surprisingly, the association between left amygdala activation and reductions in symptoms in the current study occurred regardless of training group. At the neural activation level, it is possible that rather than observing specific effects of training attention ‘away’ from threat in the active relative to the placebo conditions, we in fact detect ABM’s ability to more robustly train general attention flexibly, which is evident in both conditions. Therefore, repeated trials that alternate between face pairs containing a threat and target probes may benefit both training groups. Alternatively, both active and placebo training groups are exposed to repeated presentations of mildly threatening stimuli and greater between-session habituation may be associated with greater symptom reduction (Fischer et al., 2003) and supported by changes in amygdala reactivity. Further investigation is needed to test these alternatives.

It is important to note that our study possessed both advantages and disadvantages, relative to prior imaging studies of ABM. Unlike prior work finding PFC changes associated with ABM (Brownning et al., 2010; Eldar and Bar-Haim, 2010; Taylor et al., 2013), our whole-brain analysis identified interactions only in the amygdala, and not in the PFC. Previous fMRI studies used emotional matching and selective attention tasks to assess the neural changes associated with ABM (Brownning et al., 2010; Mansson et al., 2013; Taylor et al., 2013), which have advantages of examining how attention training generalizes to tasks beyond the task on which subjects had been trained. Thus, procedures in our study cannot identify how training generalized to different cognitive-affective tasks (i.e. far transfer). Rather, the current study examined activation to the threat bias contrast during assessment and training using similar dot-probe tasks (i.e. near transfer). By doing so, our findings may more closely characterize the specific neural mechanism of dot probe-based ABM, at the expense of failing to elucidate effects on other important cognitive processes that might contribute to any effects of ABM. Future studies should attempt to examine both far and near transfer by including both dot probe based and other fMRI tasks probing emotional and attention processes.

Given previous meta-analyses of ABM studies showing reductions in anxiety symptoms (Hakamata et al., 2010; Beard et al., 2012), it is surprising that changes in symptoms were not statistically different across groups. It is important to note that the state of ABM research continues to change, and there is a need for large-scale randomized controlled trials (RCTs). In light of our negative findings, the changing landscape of research in this area, and the need for such large RCTs, questions on the degree to which ABM can produce changes in symptoms and produce far or near transfer effects on associated psychology processes remain. Here, our negative findings were complicated by between-group differences in baseline levels of amygdala activation, which were covaried in our prediction models. In this latter statistical model, assignment to the active group (i.e. training to attend away from threat) did in fact predict larger reductions in symptom reduction than the placebo group. This pattern does suggest that, in the context of ABM training delivered alongside neuroimaging, the ability to detect training-related reductions in anxiety symptoms may be improved by accounting for amygdala activation. Indeed, MRI scans can induce stress, particularly in anxious participants (Monk et al., 2006; Perez-Edgar and Bar-Haim, 2010), and stress has been shown to alter threat-related attention patterns (Wald et al., 2011; Shechner et al., 2012). Nevertheless, interpretation of this latter finding from our prediction analysis should be interpreted with caution, due to the potential problems in controlling for confounding variables (Miller and Chapman, 2001).

**Limitations**

These findings should be considered in light of several limitations. First, our sample included healthy volunteers with high social anxiety symptoms. Healthy volunteers may retain abilities to flexibly alter attention; therefore, future studies should examine effects in a clinically anxious population. Second, random assignment using small samples yielded group differences in bilateral amygdala activation prior to training. To minimize the differences across groups, baseline amygdala activation to the threat bias contrast was included in the first step of our regression models. However, this is not an ideal solution, since caution is warranted when using a covariate to explain baseline group differences (Miller and Chapman, 2001). Importantly, random assignment to active and placebo ABM groups is unlikely to have caused baseline differences in amygdala activation. However, small studies such as ours, where multiple between-group baseline variables are examined, run the risk of Type I errors. Thus, larger samples may be needed to overcome such unintended consequence of randomization (Ioannidis et al., 2014).

**CONCLUSIONS**

Despite limitations, this neuroimaging study is the first to examine predictive factors of symptoms change associated with both ABM and placebo training. These findings have important implications for treatment. For example, traditional CBT strategies may enhance top-down processes through explicit strategies intended to reduce threat appraisals and increase emotion regulation through repeated exposure. Alternatively, here we have shown that ABM may utilize bottom-up processes to implicitly train biased threat-related attention (Bar-Haim, 2010). Understanding the underlying mechanisms of treatments may help optimize treatment approach, producing maximum benefit to individuals with particular neurobehavioral profiles. For example, individuals with exaggerated amygdala activation may benefit more from ABM. Further, for certain patients, perhaps with high left amygdala activation at baseline, ABM could be applied alongside other treatment options such as CBT that target top-down cognitive processes. Further research is needed to understand the brain-behavior relations associated with ABM training and to refine this treatment for individuals with anxiety disorders.
SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

Conflict of Interest

None declared.

REFERENCES


Behavioral Pharmacology, 39, 1254–61.


