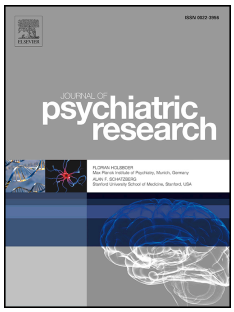
Accepted Manuscript



Neural mechanisms of expectancy-based placebo effects in antidepressant clinical trials

Sigal Zilcha-Mano, Zhishun Wang, Bradley Peterson, Melanie M. Wall, Ying Chen, Tor D. Wager, Patrick J. Brown, Steven P. Roose, Bret R. Rutherford

PII: S0022-3956(19)30340-1

DOI: https://doi.org/10.1016/j.jpsychires.2019.05.023

Reference: PIAT 3675

To appear in: *Journal of Psychiatric Research*

Received Date: 19 March 2019

Revised Date: 22 April 2019

Accepted Date: 23 May 2019

Please cite this article as: Zilcha-Mano S, Wang Z, Peterson B, Wall MM, Chen Y, Wager TD, Brown PJ, Roose SP, Rutherford BR, Neural mechanisms of expectancy-based placebo effects in antidepressant clinical trials, *Journal of Psychiatric Research* (2019), doi: https://doi.org/10.1016/j.jpsychires.2019.05.023.

Abstract word count: 250

Text word count: 3608

Number of figures: 3

Number of tables: 1

Supplementary material: 5 figures

Neural Mechanisms of Expectancy-Based Placebo Effects in Antidepressant Clinical Trials

Running title: Neural Mechanisms of Expectancy Augmentation

Keywords: antidepressants, placebo effect, clinical trials, pharmacotherapy, outcome expectancy

Sigal Zilcha-Mano, PhD*

Department of Psychology, University of Haifa

Mount Carmel, Haifa 31905, Israel

Telephone: 972-4-8249047, 972-52-3385944

E-mail: sigalzil@gmail.com

Zhishun Wang, PhD

Columbia University College of Physicians and Surgeons

New York State Psychiatric Institute

Zhishun.Wang@nyspi.columbia.edu

Bradley Peterson, MD

Keck School of Medicine

University of Southern California

bpeterson@chla.usc.edu

Melanie M. Wall, PhD

Columbia University College of Physicians and Surgeons

New York State Psychiatric Institute

Melanie.Wall@nyspi.columbia.edu

Ying Chen, PhD

Columbia University College of Physicians and Surgeons

New York State Psychiatric Institute

chenying8787@163.com

Tor D. Wager, PhD

Department of Psychology and Neuroscience and the Institute of Cognitive Science

University of Colorado at Boulder

torwager@gmail.com

Patrick J. Brown, PhD

Columbia University College of Physicians and Surgeons

New York State Psychiatric Institute

pb2410@cumc.columbia.edu

Steven P. Roose, MD

Columbia University College of Physicians and Surgeons

New York State Psychiatric Institute

spr2@cumc.columbia.edu

Bret R Rutherford, MD

Columbia University College of Physicians and Surgeons

New York State Psychiatric Institute

brr8@columbia.edu

*corresponding author

ABSTRACT

Background: Patient expectancy of therapeutic improvement is a primary mediator of placebo effects in antidepressant clinical trials, but its mechanisms are poorly understood. This study employed a novel antidepressant trial design, with integrated functional magnetic resonance imaging (fMRI), to manipulate patient outcome expectancy and examine its neural mediators.

Method: Twenty-three depressed outpatients, in a randomized controlled trial were assigned to either Open (high outcome expectancy) or Placebo-controlled (low outcome expectancy) treatment with citalopram for eight weeks. fMRI scans were acquired before and after the expectancy manipulation (before medication treatment), while participants performed a masked emotional face task. Focusing on an amygdala region-of-interest (ROI), we tested a model where reduction in amygdala activation mediated outcome expectancy effects on the slope of change in depressive symptoms.

Results: Following the manipulation, significant differences between conditions were found in neural activation changes in the amygdala, as well as in superior temporal gyrus, insula, and thalamus. Findings support the proposed mediation model according to which activation in the left amygdala ROI decreased significantly in the Open as opposed to the Placebo-controlled group following randomization ($p=0.009$) for sad vs. neutral face contrast. The reduced left amygdala activation, in turn, was a significant predictor of decreased depressive symptoms during the trial ($p=0.007$), and the mediation model was significant.

Conclusions: Results from this study, the first designed to identify the neural mechanisms of expectancy augmentation in an antidepressant randomized control trial, suggest that

therapeutic modulation of amygdala activity may be an important pathway by which patient outcome expectancy influences depressive symptoms.

ClinicalTrials.gov Identifier: NCT01919216; Trial name: Placebo Effects in the Treatment of Depression: Cognitive and Neural Mechanisms, URL: https://clinicaltrials.gov/ct2/show/NCT01919216

INTRODUCTION

Placebo response in antidepressant clinical trials has emerged in recent years as a complex phenomenon deserving scientific investigation (Rutherford and Roose, 2013). Meta-analyses and retrospective analyses have suggested that outcome expectancy, individuals' cognitive appraisal of whether and how much they will benefit from treatment, may account for a substantial portion of placebo-related improvement in depressed patients (Papakostas and Fava, 2009; Rutherford et al., 2009; Sinyor et al., 2010; Sneed et al., 2008). Recently, we showed in a prospective, randomized study that patient outcome expectancy is an important causal mechanism of placebo effects in antidepressant clinical trials (Rutherford et al., 2017). Specifically, expectancy was manipulated by instructions to participants about the probability of receiving active medication as opposed to placebo: one group was told that they were randomized to open trial antidepressant (100% chance of receiving active treatment); the other group was told that they were randomized to placebo-controlled antidepressant (50% chance of receiving active treatment). Self-reported expectancy post-manipulation was a significant mediator of the effect of expectancy manipulation on post-treatment depressive symptom reduction.

Studies of the neural mechanisms underlying placebo effects in antidepressant clinical trials largely have been limited to demonstrating objective differences in brain activity between responders and non-responders to placebo. One study (Mayberg et al., 2002) showed that placebo responses of hospitalized patients with depression were associated with regional metabolic increases in cortical areas (prefrontal, anterior and posterior cingulate, posterior insula) and decreases in limbic and paralimbic areas (thalamus, parahippocampus, subgenual cingulate). Another study found that placebo responders in an antidepressant clinical trial showed unique prefrontal changes on quantitative EEG compared to non-responders and medication responders (Hunter et al., 2006). Although these studies report important initial

findings, their naturalistic design cannot support identification of the causal mechanisms underlying placebo effects.

fMRI studies of expectancy-based placebo effects in non-depressed individuals have provided converging evidence suggesting that the brain areas associated with generating and maintaining expectancies include prefrontal cortex subregions, the orbitofrontal cortex, and the rostral anterior cingulate cortex (Petrovic et al., 2005). For example, in studies of placebo analgesia, Wager et al. (2004) found that the anticipation of pain relief was associated with activations in orbitofrontal (OFC), dorsolateral prefrontal (DLPFC), parietal, and pregenual anterior cingulate cortices, which modulated activity in parts of the insula, thalamus, and cingulate cortex associated with pain (Wager et al., 2004), possibly by potentiating pain-related opioid release (Wager et al., 2007). One study (Peciña et al., 2015) involving depressed patients, but which was not carried out as part of a randomized control trial, proposed that expectancy-based effects may be the result of changes in activation in the subgenual anterior cingulate cortex, nucleus accumbens, midline thalamus, and amygdala. Taken together, these findings may suggest that outcome expectancy-based placebo effects in antidepressant trials may also be a consequence of expectancy-related modulation of neural activity in perilimbic brain regions, such as the amygdala, which support both affective valuation processes in general and in MDD in particular (Keltner et al., 2006; Nitschke et al., 2006).

The structure and function of the amygdala have been a main focus of interest in studies investigating the neural changes at the basis of antidepressant treatments, in part because this region, whose structure and function are disrupted in MDD (Hamilton et al., 2012; Stuhrmann et al., 2011), is a primary node of emotional brain circuits (Williams and Gordon, 2007). For example, amygdala activation predicts trajectories of symptom change during antidepressant treatment, and it is especially sensitive to functional modulation (Fu et

al., 2004; Victor et al., 2010; Williams et al., 2015; Williams and Gordon, 2007). A robust way of probing amygdala activation by functional neuroimaging is with emotional faces (Ronchetti, 1990; Williams and Gordon, 2007), which have been used to define abnormalities in the processing of specific emotions. Hyperactivation of the amygdala in MDD has been observed during supraliminal and subliminal processing, especially of sad and fearful facial expressions (Arnone et al., 2012; Fu et al., 2004; Surguladze et al., 2005; Victor et al., 2010), and was shown to attenuate following treatment with antidepressants (Arnone et al., 2012; Fu et al., 2004; Victor et al., 2010; Williams et al., 2015). The advantage of subliminal conditions is that they help isolate the automatic processes that underpin amygdala activation from more elaborative supraliminal processes (Williams et al, 2006; Costafreda et al, 2008). Based on the literature, it may be hypothesized therefore that normalization of amygdala hyperactivation in depressed individuals may be the mechanism by which the outcome expectancy effect in antidepressant trials operates.

Given the evidence of amygdala dysfunction in patients with MDD, and its known roles in subserving outcome expectancy/appraisal processes and mediating antidepressant treatment effects, we sought to characterize the role of the amygdala in mediating expectancy augmentation effects in antidepressant treatment. Outpatients with MDD were randomly assigned to open administration (100% probability) of citalopram or placebo-controlled administration (50% probability) of citalopram. Outcome expectancy and depressive symptom scores were followed over 8 weeks of acute treatment. fMRI scans were acquired while patients performed a backward-masked emotional face task, designed to probe amygdala activation, before and after the expectancy manipulation. An amygdala-based ROI analysis, supplemented by voxel-wise whole brain analysis, was performed to identify the brain regions mediating clinical expectancy effects. We hypothesized that successful

modulation of MDD-related amygdala hyperactivation mediates outcome expectancy effects on response to treatment.

METHODS

Participants

The study was conducted in the Adult and Late Life Depression Research Clinic and MRI Laboratory at the New York State Psychiatric Institute (NYSPI). All procedures were approved by the NYSPI Institutional Review Board. Eligible participants were men and women aged 24-65 years, who met Diagnostic and Statistical Manual IV (DSM-IV) (American Psychiatric Association 2000) criteria for non-psychotic MDD, had a 24-item HRSD score ≥ 16, were right-handed, had no contraindications to MRI, gave informed consent, and complied with study procedures.

Study design

Study procedures are described in a previous report of clinical findings (Rutherford et al., 2017). Briefly, 50 patients were enrolled in an 8-week antidepressant clinical trial, randomizing participants to Placebo-controlled and Open groups. At baseline, patients underwent initial evaluation, eligibility was assessed, and pre-randomization HRSD scores and outcome expectancy (operationalized as their belief regarding the probability of receiving medication: 0 vs. 25% vs. 50% vs. 100%) were measured. fMRI scan 1 was performed as soon as possible after this visit, within 1 week. Following fMRI scan 1, patients' level of outcome expectancy was manipulated by randomization to either the Placebo-controlled group (50% chance of receiving active treatment) or the Open group (100% chance of receiving active treatment), and patients were informed of the results of randomization (which was the means of manipulating outcome expectancy). Outcome assessors were

blinded to group assignment. At the Week 0 visit, post-randomization outcome expectancy and depression scores were measured, with participants having this additional information. Participants in the Placebo-controlled group were blinded to treatment assignment within the group. fMRI scan 2 was then performed within 1 week of the Week 0 visit, after which either citalopram or a placebo pill was administered. Thus, both pre- and post-randomization outcome expectancy measurements and fMRI scans 1-2 were obtained before patients received any medication. HRSD was measured weekly over the 8-week clinical trial.

Materials

Masked Emotional Face task

In this task, participants viewed black and white pictures of human faces displaying fearful, sad, happy, or neutral emotional expressions taken from a standardized series (see **Figure 1**) (Ekman, 1976). Stimuli were masked so that an emotional face was presented for 33ms followed by 160ms presentation of a neutral face. Pilot testing and post-scan debriefing indicated that participants are only consciously aware of observing one face per trial. Following the face presentations, participants obtain affective ratings using a grid displaying the dimensions of valence (pleasant-unpleasant) and arousal (excited-sleepy) as visual analogue scales on the x- and y-axes, respectively, ranging from 1 to 100 in each dimension. Patients viewed 1 run of 120 trials comprising 30 presentations of each emotional valence (sad, happy, fearful, and neutral) followed by the neutral face. Each run scanned approximately 450 functional images (TR=2000ms).

Image acquisition

Images were obtained on a GE Signa 3-T whole body scanner (Milwaukee, WI) operating the E2-M4 platform using a quadrature head coil in receive mode. T1-weighed sagittal localizing images were used to position axial functional images parallel to the anterior-posterior commissure (AC-PC) line. A 3D spoiled gradient recall (SPGR) image was acquired for coregistration with axial echoplanar images and a reference brain from the Montreal Neurological Institute (MNI). Axial echoplanar images (TR = 2000 ms, TE = 28 ms, 77° flip angle, single excitation per image, slice thickness 3.54 mm, 1.0 mm gap, 24 cm × 24 cm field of view, 64 × 64 matrix) were obtained to provide an effective resolution of 3.75 mm × 3.75 mm × 3.5 mm and whole brain coverage, with 35 slices in each imaging volume and 452 volumes per run.

Image pre-processing

SPM8 (http://www.fil.ion.ucl.ac.uk/spm/) under MATLAB 2014B was used to preprocess the functional imaging data. The preprocessing procedure included the following steps: (a) slice-timing correction using the middle slice of each run as the reference image; (b) motion correction for three translational directions and rotations using a rigid-body transform; (c) spatial normalization to the standard MNI template using a hybrid algorithm of affine transform and nonlinear warping. Each participant's high-resolution structural image (fSPGR) was normalized to the template, and these subject-specific warping parameters were then used to normalize the functional images to the same template; (d) reformatting of the normalized functional images to 3x3x3 mm voxels; (e) Gaussian spatial filtering with a FWHM of 8 mm. A discrete cosine transform-based high-pass filter with a basis function length of 128s was also used to remove low-frequency noise, such as scanner drift, from the baseline image intensity.

Functional Image Analyses

Using SPM8, we performed an individual-level analysis (first-level) to detect task-related (face stimulation-related) activity within each participant. We then performed group-level analysis (second-level) to detect random effects of task-related activity. We conducted the first-level analysis using the general linear model (GLM), as implemented in SPM8, to model the data for each participant, with 4 independent functions and a constant for each run. The first 2 independent functions corresponded to 2 events recorded in the task, each generated by convolving a canonical hemodynamic response function (HRF) with a boxcar function (BCF) derived from the onsets and durations of each event, facial presentation, and participant rating. The second 2 independent functions were generated by a separate amplitude modulation of the facial stimulation function with each rating score, arousal score, and valence score. The model was estimated using the Restricted Maximum Likelihood (ReML) algorithm. Task-related T contrast images were generated using SPM8.

We implemented a Bayesian posterior inference approach (Surguladze et al., 2005) for the second-level analysis of the contrast images generated from the first-level GLM-based analysis to detect the random effects of task-related activity within and between the groups. We used a posterior probability of 97.50% as the threshold of significant posterior probability maps (PPMs), a rigorous threshold in Bayesian inference, to ensure that reported findings are true positives (Friston and Penny, 2003). We extracted ROI BOLD data based on the PPM images within those regions, showing significant group effects (open vs. PC group) in the differences between scan 2 and scan 1 on the contrast images of sad vs. neutral faces. The amygdala ROI was defined based on a brain atlas (Amunts et al., 2005), and the signal was extracted from the ROI by averaging BOLD signals across all voxels within the ROI for each contrast, for each patient. We entered the ROI data into further mediation analyses.

Data Analyses and Hypothesis Testing

First, we used a Mann-Whitney-Wilcoxon Exact Test to assess whether the outcome expectancy manipulation in the MRI subsample produced significant changes in outcome expectancy from pre- to post-randomization. Second, we used Spearman correlation to assess whether changes in outcome expectancy from pre- to post-randomization significantly correlated with changes in neural activation within specified amygdala ROIs from pre- to post-randomization. Of the three contrasts examined (sad vs. neutral, fearful vs. neutral, and happy vs. neutral faces), we focused the ROI analyses on the contrasts that showed substantial changes in amygdala activation in the whole-brain analyses.

Following Preacher and Hayes (2004), we assessed mediation by testing whether the expectancy manipulation was associated with changes in amygdala activation, and whether change in amygdala activation in turn correlated with the slope of change in HRSD, controlling for expectancy manipulation (**Figure S1**). Next, we repeated these mediation analyses using robust linear models based on M-estimators, as implemented in the WRS2 package of R software. Robust inferential methods perform well with relatively small sample sizes (Ronchetti, 1990; Wilcox, 2011), assigning a weight to each observation based on its Mahalanobis distance, so that observations in the tail of the distribution receive lower weights. Finally, we estimated and tested the significance of the mediation effect using 10,000 bootstrap samples combined with a robust estimation routine (Zu and Yuan, 2010). We calculated the proportion of the explained variance (R^2) of the robust regression effects, as previously described (Willett and Singer, 1988).

RESULTS

Demographic and clinical characteristics

Of the patients participating in the RCT, 23 met imaging criteria (no MRI-contraindications, etc.), and created the effective sample for this secondary analysis. Of these patients, 9 were randomized to the Open group and 14 to the Placebo-controlled group (11 received medication and 3 received placebo). No significant differences in demographic data or baseline clinical characteristics were found between participants who were and were not scanned (Table S1), or between participants in the Placebo-controlled and Open groups.

Behavioral effects

Valence and arousal ratings of the masked fearful, sad, happy, and neutral faces are presented in Table S2. Across groups, at Scan 1, sad and fearful faces tended to be rated as more arousing and less pleasant than neutral faces, but given the relatively small sample size, significant differences between emotional and neutral faces were observed only on valence ratings of fearful faces ($t_{(24)}$=-2.38, p=.02, Cohen's d = 0.97). Participants randomized to the Open group experienced numerically larger decreases in arousal/valence ratings measured before and after randomization to group than did participants randomized to the Placebo-controlled group, but group differences were not statistically significant.

Whole-brain and ROI analyses

Compared to the Placebo-controlled group (**Figures 2b, 2Sb, and 3Sb**), the Open group (**Figures 2a, 2Sa, and 3Sa**) showed significant neural activation changes on sad vs. neutral, fearful vs. neutral, and happy vs. neutral contrasts following the outcome expectancy manipulation (for the comparisons between the groups, see **Figures 2c, 2Sc, and 3Sc**, for sad, fearful and happy faces, respectively). **Figure 2** (for the sad vs. neutral face contrast) and **Figures S2-3 (**for the fearful/happy vs. neutral contrasts) show significant group differences in numerous brain regions, including amygdala (peak posterior probability (PPP)>99.99%),

superior temporal gyrus (STG) (PPP>99.99%), postcentral gyrus (PoG) (PPP >99.99%), insula (PPP>99.99%), and thalamus (PPP>99.88%). Between-groups contrasts on all three emotional face contrasts (sad vs. neutral, fearful vs. neutral, happy vs. neutral) demonstrated significant deactivations in the superior temporal gyrus and DLPFC bilaterally, but bilateral amygdala deactivation was observed on the sad vs. neutral face contrast only.

Focusing on the pre-post-randomization activation maps between the Open and Placebo-controlled groups in the amygdala ROI revealed greater decreases in amygdala activation on the sad vs. neutral face contrast in the Open group (**Figure 2c**). At baseline, participants in the Open group demonstrated activation in the left amygdala on this contrast, which, as expected, decreased following randomization, producing a significant left amygdala deactivation from Scan 1 to Scan 2 (**Figure 2a**). In contrast, as expected, a significant activation increase was observed from Scan 1 to Scan 2 in the Placebo-controlled group, resulting in a significant between-groups difference in left amygdala activation change from pre- to post-randomization (from Scan 1 to Scan 2; **Figure 2b**). **Figure 3** depicts these activation changes on the sad vs. neutral face contrast on the same coronal slice for each group, enlarging the left amygdala ROI. Peak coordinates are presented in **Table 1**.

Neural mediation of outcome expectancy-based placebo effects

We tested whether activation change in the amygdala ROI on the sad vs. neutral face contrast mediated the observed clinical effect of patient outcome expectancy on depressive symptom severity. First, we re-confirmed (in this smaller neuroimaging subset of our larger clinical sample) that the experimental randomization to Open vs. Placebo-controlled group resulted in between-groups outcome expectancy differences. The Mann-Whitney-Wilcoxon Exact Test demonstrated significant outcome expectancy differences between groups (W=31.5, p=.007), with patients in the Open group showing significantly greater increase in

outcome expectancy from pre- to post-manipulation (Mean=1.75, SD=0.50) than patients in the Placebo-controlled group (Mean=-0.12, SD=0.64). Second, we found a significant association between changes in left amygdala activation and in outcome expectancy from pre- to post-randomization (r=-.74, p=.006): there was a greater decrease in randomization-induced amygdala activation for the left amygdala as outcome expectancy scores increased (became more positive; **Figure S4a**). This association was not observed for the right amygdala (r=-.11, p=.24; **Figure S4b**).

Finally, we tested the mediation model. Any baseline differences were accounted for in the mediation analysis by the use of delta scores. Because there were no significant differences between activation conditions in the right amygdala for sad faces (p=.35), these data were not included in further analyses, and all analyses reported from here onward refer to the left amygdala. The first model revealed a significant effect of outcome expectancy manipulation on changes in left amygdala activation (B=5.94, S.E.=2.07, t=2.86, p=0.009, R^2=0.28). Patients in the Open group showed significantly greater reduction in activation from pre- to post-randomization (Mean=-2.34, SD=3.16) than did patients in the Placebo-controlled group (Mean=3.59, SD=5.66). The second model revealed a significant ability of changes in amygdala activation to predict the slope of change in depressive symptom severity (HRSD scores) from Week 0 to endpoint, controlling for outcome expectancy manipulation (Open vs. Placebo-controlled) (**Figure S5**, B=-0.09, S.E.=0.03, t=-3.01, p=0.007). The significant effect suggested that HRSD scores declined over time at a faster rate for patients demonstrating greater reduction in amygdala activation from pre- to post-randomization. The change in HRSD from week 0 to week 8 for the patients in the upper quartile of amygdala activation reduction was 15.0 (SD = 13), whereas for those in the lower quartile was 7.4 (SD = 11.5). The indirect effect was significant (IE=-0.52 CI95% [-1.3078, -0.0311], R^2=0.21). The total effect of outcome expectancy manipulation on HRSD slope was

significant when not controlling for changes in amygdala activation (*B*=-0.82, S.E.=.33, *t*=-2.48, *p*=.02), and the direct effect was non-significant when controlling for changes in amygdala activation (*B*=-0.29, S.E.=.33, *t*=-0.88, *p*=.39); 63.41% [0.07,0.71] of the total effect of outcome expectancy on slope change in HRSD is mediated by changes in amygdala activation. These findings support a mediation model in which outcome expectancy manipulation predicts changes in amygdala activation, which in turn predict the slope of change in HRSD.

Repeating the analyses using robust linear models based on M-estimators further supported the proposed mediation model (IE=-0.52 CI 95% [-1.3,-0.03], p=.027); 84% of the total effect of outcome expectancy on HRSD slopes was mediated by changes in amygdala activation. The first model revealed that patients in the Open group showed significantly greater reduction in activation from pre- to post-randomization than did patients in the Placebo-controlled group (*B*=5.83, $F_{(1,21)}$=8.41, *p*=0.009, R^2=0.34). The second model revealed that HRSD scores declined over time at a faster rate for patients demonstrating greater reduction in amygdala activation from pre- to post-randomization (*B*=-0.09, $F_{(1,20)}$=6.88, *p*=0.016, R^2=0.18). The total effect of outcome expectancy manipulation on HRSD slope was significant when not controlling for changes in amygdala activation (*B*=-0.62, $F_{(1,20)}$=4.05, *p*=.05, R^2=0.29), and the direct effect was non-significant when controlling for changes in amygdala activation (*B*=-0.28, $F_{(1,20)}$=0.59, *p*=.45). Repeating the analyses without patients receiving placebo (N = 3) resulted in similar findings, and the mediation effect remained significant (IE=-0.72 CI95% [-1.6692,-0.0132], *p*=.042).

DISCUSSION

The principal findings of this study were that expectancy augmentation in this antidepressant clinical trial, which we previously showed to be mediated by self-reported outcome expectancy, is partially mediated at a neural level by reduced amygdala activation. Manipulating outcome expectancy through increased probability of receiving active medication (as opposed to placebo) was associated with decreased amygdala activation in response to sad emotional faces, which in turn was associated with more rapid reduction in depressive symptoms during the course of antidepressant treatment. Further modeling revealed that the influence of outcome expectancy manipulation on depressive symptoms was partially mediated by change in amygdala activation, measured before patients received antidepressant medication. To the best of our knowledge, this is the first study to demonstrate that manipulating outcome expectancy in antidepressant trials results in modulating amygdala activation, making these findings an important step in elucidating the neural mechanism of the placebo effect in antidepressant clinical trials.

These results are among the first to provide evidence of the causal mechanisms by which placebo effects operate in antidepressant clinical trials. Our findings are consistent with neuroimaging investigations across a range of emotional experiences, from physical pain (Wager et al., 2004) to taste (O'Doherty et al., 2002), suggesting modulation of amygdala activation as a means by which expectancy regulates mood. The findings are also consistent with a recent report investigating the neural correlates of response to a 1-week placebo lead-in phase and the association of placebo response during lead-in with response to brief antidepressant treatment (Peciña et al., 2015). Increased placebo-induced μ-opioid neurotransmission in a network of regions implicated in the pathophysiology of MDD, including the amygdala, was associated with better antidepressant treatment response. Consistent with previous reports (Williams et al., 2015), the present findings suggest that the normalization of amygdala activity goes hand-in-hand with the normalization of symptoms.

The finding of a mediation model in the left rather than right amygdala is also consistent with previous reports demonstrating left amygdala hyperarousal in patients with MDD (Sheline et al., 2001).

Using whole-brain analyses, we were also able to explore other brain regions (e.g., thalamus, insula, the left and right superior temporal gyrus [STG], and the left postcentral gyrus [PoG]), demonstrating significant activation differences following the experimental manipulation. These findings are consistent with prior data and systematic reviews suggesting that the thalamus and insula play a role in transducing placebo response across disorders and symptom types (Ashar et al., 2017). The STG has been implicated in impaired affective appraisal effect (Ashar et al., 2017; Scott et al., 2007). The postcentral gyrus includes the primary somatosensory cortex, which is structurally and functionally connected to the thalamus, and plays a role in controlling and modulating associatively learned behaviors (Chau et al., 2013; Galvez et al., 2006). Abnormal function of this area and its connectivity with the thalamus have been suggested as a potential biomarker for MDD, given their association with core clinical MDD symptoms (Kang et al., 2018).

Determining the neural correlates of outcome expectancy provides important information about changes in the brain associated with improvement in depressive symptoms, and may help distinguish placebo response from improvement due to specific medication or psychotherapy effects. Although previous studies ascribed the brain changes observed during open medication treatment or open psychotherapy to the specific treatments, it is crucial to differentiate the brain changes associated with drug-specific or psychotherapy-specific factors from those due to expectancy. This is a critical shortcoming of previous research because the changes are in part the result of placebo effects. The data reported here help efforts to reveal the neural correlates of treatment effects by identifying the neural mechanisms of outcome expectancy.

The most significant limitation of the present study is the small sample size. Another limitation is that because of ethical considerations, it was not possible to use a high-outcome expectancy placebo group (i.e., informing participants that they were assigned to open trial but providing them with placebo). Additionally, although task-related neuroimaging approaches are of great importance, their findings should be complemented with resting state fMRI data, because each has its own advantages and disadvantages, and the literature reveals that consistent vs. inconsistent findings between the two approaches can add meaningful new knowledge (Di et al., 2013). Finally, although the findings provide important support for the proposed mediation effect, we did not examine a two-mediators model according to which expectancy manipulation predicts reduction in amygdala hyperactivation, which then predict changes in expectancy, which in turn predicts changes in depression.

This study is the first to manipulate and prospectively study outcome expectancy, deploying serial functional neuroimaging, and careful measurement of outcome expectancy and depressive symptoms. The principal findings of the study are that placebo effects in this antidepressant clinical trial, which we previously showed to be mediated by outcome expectancy, are partially mediated at a neural level by reduced amygdala activation. To the best of our knowledge, this is the first study to demonstrate that manipulating outcome expectancy in antidepressant trials results in modulating amygdala activation, making these findings an important step in elucidating the neural mechanism of the placebo effect in antidepressant clinical trials.

Acknowledgments

Work on this paper was supported by NIMH grant K23 MH085236 (to Dr. Rutherford).

Disclosures

Drs. Zilcha-Mano and Rutherford had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Drs.

Rutherford, Zilcha-Mano, Wager, Peterson, Wang, Wall, Brown, Roose, and Chen have no disclosure or conflict of interest to report. This paper has not been previously presented.

REFERENCES

Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N.J., Habel, U., Schneider, F., Zilles, K., 2005. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. Anat. Embryol. (Berl). 210, 343–352.

Arnone, D., McKie, S., Elliott, R., Thomas, E.J., Downey, D., Juhasz, G., Williams, S.R., Deakin, J.F.W., Anderson, I.M., 2012. Increased amygdala responses to sad but not fearful faces in major depression: relation to mood state and pharmacological treatment. Am. J. Psychiatry 169, 841–850.

Ashar, Y.K., Chang, L.J., Wager, T.D., 2017. Brain mechanisms of the placebo effect: an affective appraisal account. Annu. Rev. Clin. Psychol. 13, 73–98.

American Psychiatric Association, 2000. Diagnostic and statistical manual of mental disorders (revised 4th ed.). Washington, DC Author.

Chau, L.S., Davis, A.S., Galvez, R., 2013. Neocortical synaptic proliferation following forebrain-dependent trace associative learning. Behav. Neurosci. 127, 285.

Costafreda SG, Brammer MJ, David AS, Fu CH (2008). Predictors of amygdala activation during the processing of emotional stimuli: a meta-analysis of 385 PET and fMRI studies. Brain Res Rev 58: 57–70.

Di X, Gohel S, Kim EH, Biswal BB. Task vs. rest-different network configurations between the coactivation and the resting-state brain networks. *Front Hum Neurosci* 17, 493.

Ekman, P., 1976. Pictures of facial affect. Consult. Psychol. Press.

Friston, K.J., Penny, W., 2003. Posterior probability maps and SPMs. Neuroimage 19, 1240–1249.

Fu, C.H.Y., Williams, S.C.R., Cleare, A.J., Brammer, M.J., Walsh, N.D., Kim, J., Andrew,

C.M., Pich, E.M., Williams, P.M., Reed, L.J., 2004. Attenuation of the neural response to sad faces in major depressionby antidepressant treatment: a prospective, event-related functional magnetic resonance imagingstudy. Arch. Gen. Psychiatry 61, 877–889.

Galvez, R., Weiss, C., Weible, A.P., Disterhoft, J.F., 2006. Vibrissa-signaled eyeblink conditioning induces somatosensory cortical plasticity. J. Neurosci. 26, 6062–6068.

Hamilton, J.P., Etkin, A., Furman, D.J., Lemus, M.G., Johnson, R.F., Gotlib, I.H., 2012. Functional neuroimaging of major depressive disorder: a meta-analysis and new integration of baseline activation and neural response data. Am. J. Psychiatry 169, 693–703.

Hunter, A.M., Leuchter, A.F., Morgan, M.L., Cook, I.A., 2006. Changes in brain function (quantitative EEG cordance) during placebo lead-in and treatment outcomes in clinical trials for major depression. Am. J. Psychiatry 163, 1426–1432.

Kang, L., Zhang, A., Sun, N., Liu, P., Yang, C., Li, G., Liu, Z., Wang, Y., Zhang, K., 2018. Functional connectivity between the thalamus and the primary somatosensory cortex in major depressive disorder: a resting-state fMRI study. BMC Psychiatry 18, 339.

Keltner, J.R., Furst, A., Fan, C., Redfern, R., Inglis, B., Fields, H.L., 2006. Isolating the modulatory effect of expectation on pain transmission: a functional magnetic resonance imaging study. J. Neurosci. 26, 4437–4443.

Mayberg, H.S., Silva, J.A., Brannan, S.K., Tekell, J.L., Mahurin, R.K., McGinnis, S., Jerabek, P.A., 2002. The functional neuroanatomy of the placebo effect. Am. J. Psychiatry 159, 728–737.

Nitschke, J.B., Dixon, G.E., Sarinopoulos, I., Short, S.J., Cohen, J.D., Smith, E.E., Kosslyn, S.M., Rose, R.M., Davidson, R.J., 2006. Altering expectancy dampens neural response to aversive taste in primary taste cortex. Nat. Neurosci. 9, 435.

O'Doherty, J.P., Deichmann, R., Critchley, H.D., Dolan, R.J., 2002. Neural responses during

anticipation of a primary taste reward. Neuron 33, 815–826.

Papakostas, G.I., Fava, M., 2009. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. Eur. Neuropsychopharmacol. 19, 34–40.

Peciña, M., Bohnert, A.S.B., Sikora, M., Avery, E.T., Langenecker, S.A., Mickey, B.J., Zubieta, J.-K., 2015. Association between placebo-activated neural systems and antidepressant responses: neurochemistry of placebo effects in major depression. JAMA psychiatry 72, 1087–1094.

Petrovic, P., Dietrich, T., Fransson, P., Andersson, J., Carlsson, K., Ingvar, M., 2005. Placebo in emotional processing—induced expectations of anxiety relief activate a generalized modulatory network. Neuron 46, 957–969.

Preacher, K. J., Hayes, A. F., 2004. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Instrum Comput, 36(4), 717-731.

Ronchetti, E., 1990. Small sample asymptotics: A review with applications to robust statistics. Comput. Stat. Data Anal. 10, 207–223.

Rutherford, B.R., Roose, S.P., 2013. A model of placebo response in antidepressant clinical trials. Am. J. Psychiatry 170, 723–733. https://doi.org/10.1176/appi.ajp.2012.12040474

Rutherford, B.R., Sneed, J.R., Roose, S.P., 2009. Does study design influence outcome? Psychother. Psychosom. 78, 172–181.

Rutherford, B.R., Wall, M.M., Brown, P.J., Choo, T.H., Wager, T.D., Peterson, B.S., Chung, S., Kirsch, I., Roose, S.P., 2017. Patient expectancy as a mediator of placebo effects in antidepressant clinical trials. Am. J. Psychiatry 174, 135–142. https://doi.org/10.1176/appi.ajp.2016.16020225

Scott, D.J., Stohler, C.S., Egnatuk, C.M., Wang, H., Koeppe, R.A., Zubieta, J.-K., 2007. Individual differences in reward responding explain placebo-induced expectations and

effects. Neuron 55, 325–336.

Sheline, Y.I., Barch, D.M., Donnelly, J.M., Ollinger, J.M., Snyder, A.Z., Mintun, M.A., 2001. Increased amygdala response to masked emotional faces in depressed subjects resolves with antidepressant treatment: an fMRI study. Biol. Psychiatry 50, 651–658.

Sinyor, M., Levitt, A.J., Cheung, A.H., Schaffer, A., Kiss, A., Dowlati, Y., Lanctôt, K.L., 2010. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials? Results from pooled and meta-analyses.

Sneed, J.R., Rutherford, B.R., Rindskopf, D., Lane, D.T., Sackeim, H.A., Roose, S.P., 2008. Design makes a difference: a meta-analysis of antidepressant response rates in placebo-controlled versus comparator trials in late-life depression. Am. J. Geriatr. Psychiatry 16, 65–73.

Stuhrmann, A., Suslow, T., Dannlowski, U., 2011. Facial emotion processing in major depression: a systematic review of neuroimaging findings. Biol. Mood Anxiety Disord. 1, 10.

Surguladze, S., Brammer, M.J., Keedwell, P., Giampietro, V., Young, A.W., Travis, M.J., Williams, S.C.R., Phillips, M.L., 2005. A differential pattern of neural response toward sad versus happy facial expressions in major depressive disorder. Biol. Psychiatry 57, 201–209.

Victor, T.A., Furey, M.L., Fromm, S.J., Öhman, A., Drevets, W.C., 2010. Relationship between amygdala responses to masked faces and mood state and treatment in major depressive disorder. Arch. Gen. Psychiatry 67, 1128–1138.

Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M., Cohen, J.D., 2004. Placebo-induced changes in FMRI in the anticipation and experience of pain. Science (80-.). 303, 1162–1167.

Wager, T.D., Scott, D.J., Zubieta, J.-K., 2007. Placebo effects on human μ-opioid activity

during pain. Proc. Natl. Acad. Sci. 104, 11056–11061.

Wilcox, R.R., 2011. Introduction to robust estimation and hypothesis testing. Academic press.

Willett, J.B., Singer, J.D., 1988. Another cautionary note about R 2: Its use in weighted least-squares regression analysis. Am. Stat. 42, 236–238.

Williams, L.M., Gordon, E., 2007. Dynamic organization of the emotional brain: responsivity, stability, and instability. Neurosci. 13, 349–370.

Williams LM, Das P, Liddell BJ, Kemp AH, Rennie CJ, Gordon E (2006). Mode of functional connectivity in amygdala pathways dissociates level of awareness for signals of fear. J Neurosci 26: 9264–9271.

Williams, L.M., Korgaonkar, M.S., Song, Y.C., Paton, R., Eagles, S., Goldstein-Piekarski, A., Grieve, S.M., Harris, A.W.F., Usherwood, T., Etkin, A., 2015. Amygdala reactivity to emotional faces in the prediction of general and medication-specific responses to antidepressant treatment in the randomized iSPOT-D trial. Neuropsychopharmacology 40, 2398.

Zu, J., Yuan, K.-H., 2010. Local influence and robust procedures for mediation analysis. Multivariate Behav. Res. 45, 1–44.

Table 1. Peak coordinates for activation differences on the sad vs. neutral face contrast from Scan 1 to Scan 2 between the Open and Placebo-controlled groups

Region	Hemisphere	Peak location			Z score
		x	y	z	
Superior Temporal Gyrus	R	60	-13	4	-7.40
Postcentral Gyrus	L	-60	-7	25	-6.96
Superior Temporal Gyrus	L	-56	-13	2	-4.19
Insula	R	39	-9	7	-4.26
Thalamus	R	8	-9	-3	-3.05
Amygdala	L	-17	-8	-24	-3.70

Table S1. Clinical and demographic characteristics of patients who were and were not scanned

Characteristic	**Scanned** (N = 23)	**Not-scanned** (N = 27)	**Comparison** (p value)
Age (years)	40.04 ± 8.29	44.63 ± 13.27	.16
Gender (% male)	43.5	40.7	.53
Ethnicity (% Hispanic)	21.7	14.8	.39
Baseline 24-item HRSD	26.73 ± 4.49	25.04 ± 4.94	.21
Baseline QIDS SR	20.69 ± 6.51	19.53 ± 5.79	.51

Note. HRSD = Hamilton Rating Scale for Depression, QIDS SR = Quick Inventory of Depressive Symptomatology Self-report.

Table S2. Behavioral data for fear, sadness and neutral priming for pre- and post-randomization for the Placebo-Controlled Group and the Open Group. Participants rated arousal and valence on a two-dimensional grid by moving an arrow controlled by an MRI-compatible computer mouse. The screen remained visible until the participant clicked the mouse button, up to a maximum of 5 seconds

		Arousal		**Valence**	
Emotional face type	**Group assignment**	**Scan 1**	**Scan 2**	**Scan 1**	**Scan 2**
Fear	Placebo-controlled	8.0 (20.8)	11.6 (22.96)	-20.61(20.33)	-14.18(22.73)
	Open Group	9.63(10.30)	2.95(21.90)	-8.91(18.56)	-1.54(20.89)
Sad	Placebo-controlled	12.17(26.12)	11.53(23.48)	-16.91(24.65)	-13.58(22.2)
	Open Group	9.75(14.01)	1.37(22.35)	-10.72(16.03)	-5.91(18.77)
Happy	Placebo-controlled	11.1(22.7)	11.9(23.2)	-16.1(20)	-10.0(22.3)
	Open Group	12.4(10.7)	4.6(21.8)	-10.1(11.4)	.11(22.8)
Neutral	Placebo-controlled	6.43(22.28)	6.98(21.90)	-13.67(22.19)	-9.93(24.43)
	Open Group	7.92(17.56)	1.46(24.15)	-.93(19.85)	-.82(23.23)

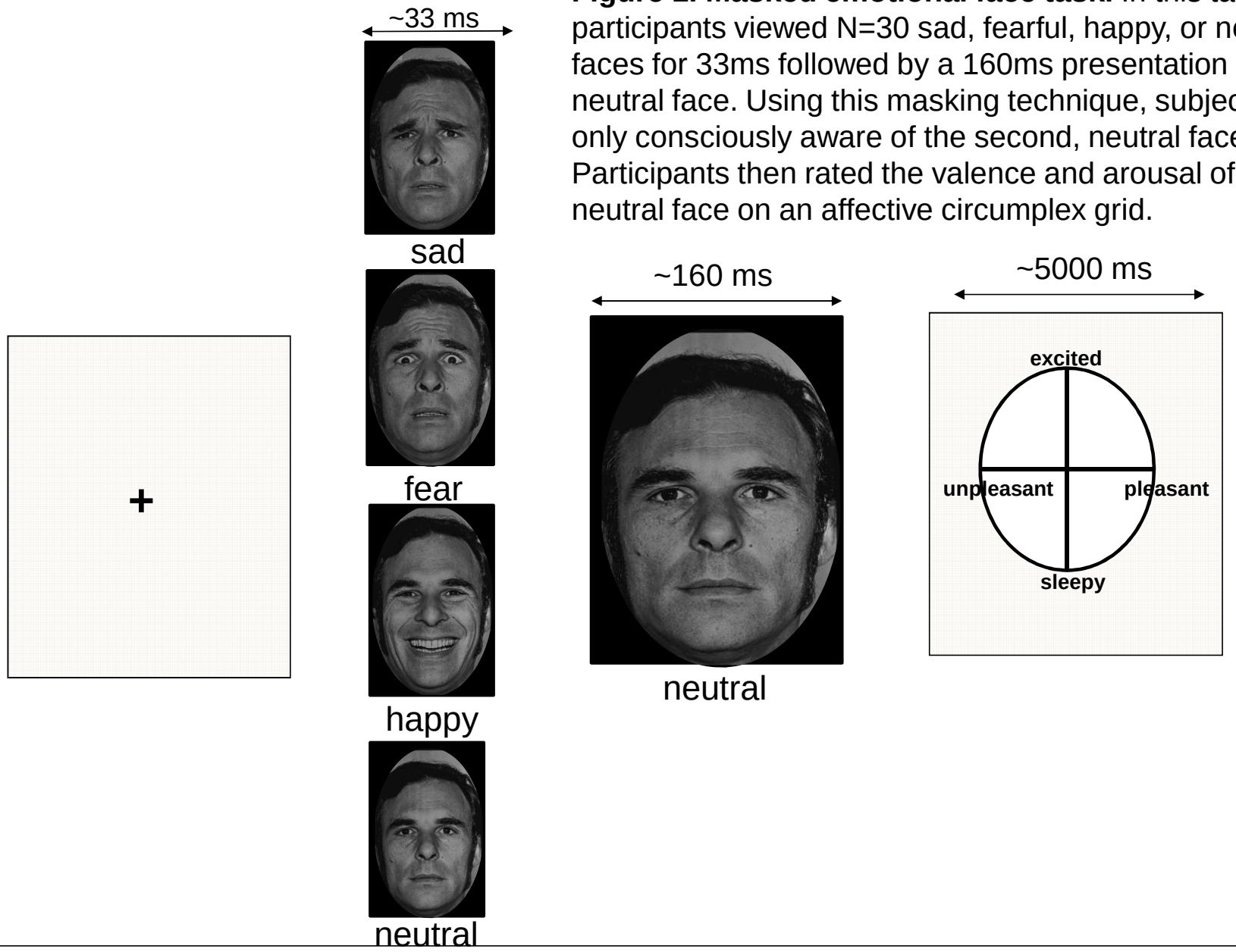


Figure 1. Masked emotional face task. In this task, participants viewed N=30 sad, fearful, happy, or neutral faces for 33ms followed by a 160ms presentation of a neutral face. Using this masking technique, subjects were only consciously aware of the second, neutral face. Participants then rated the valence and arousal of the neutral face on an affective circumplex grid.

Figure 2. Within- and Between-group neural activation maps for the sad vs. neutral face contrast. Panels a-b present Scan 1, Scan 2, and their difference (Scan 2 – Scan 1) for the Open and Placebo-controlled groups, respectively. Panel c presents the between-group difference in neural activation change from Scan 1 to Scan 2.

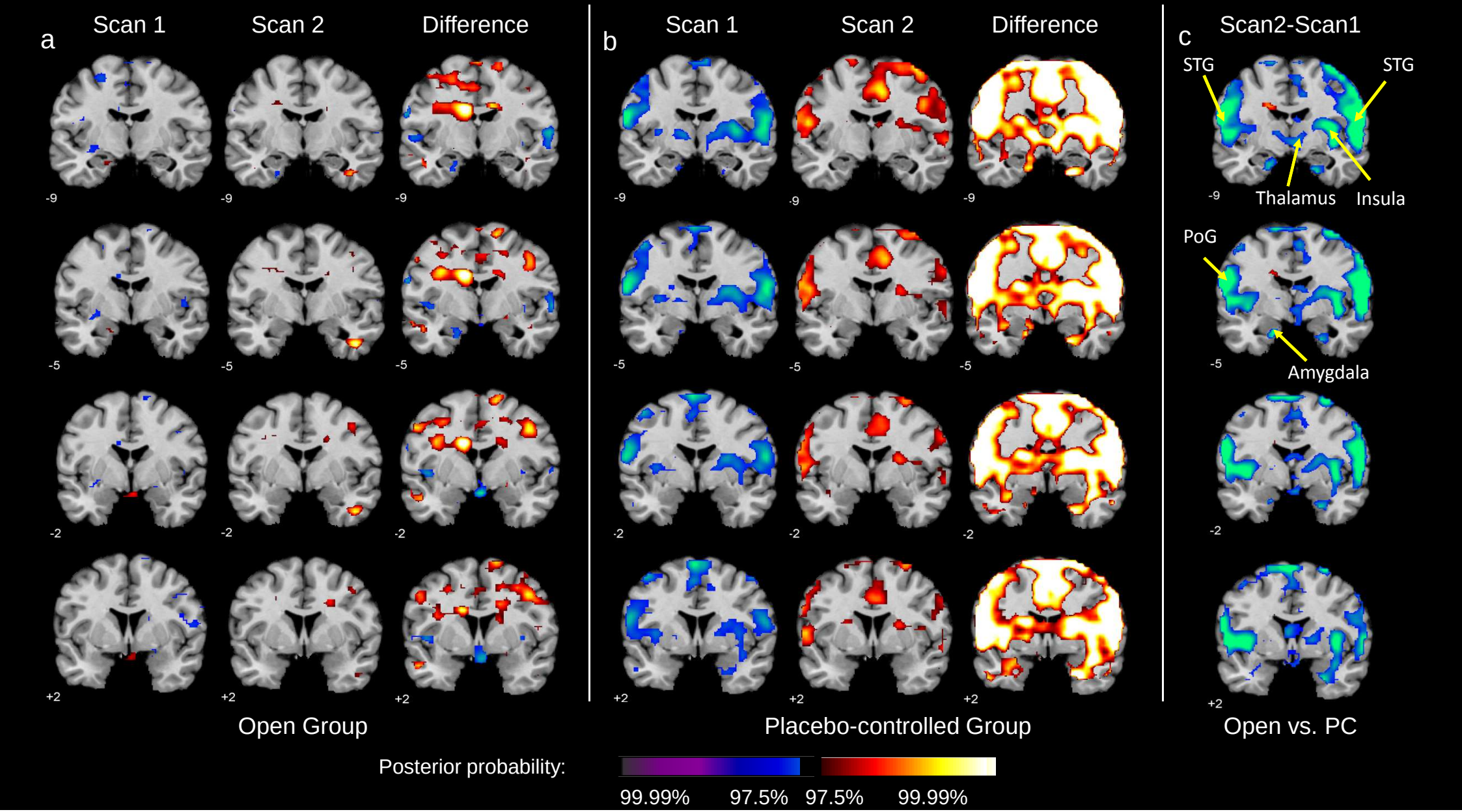
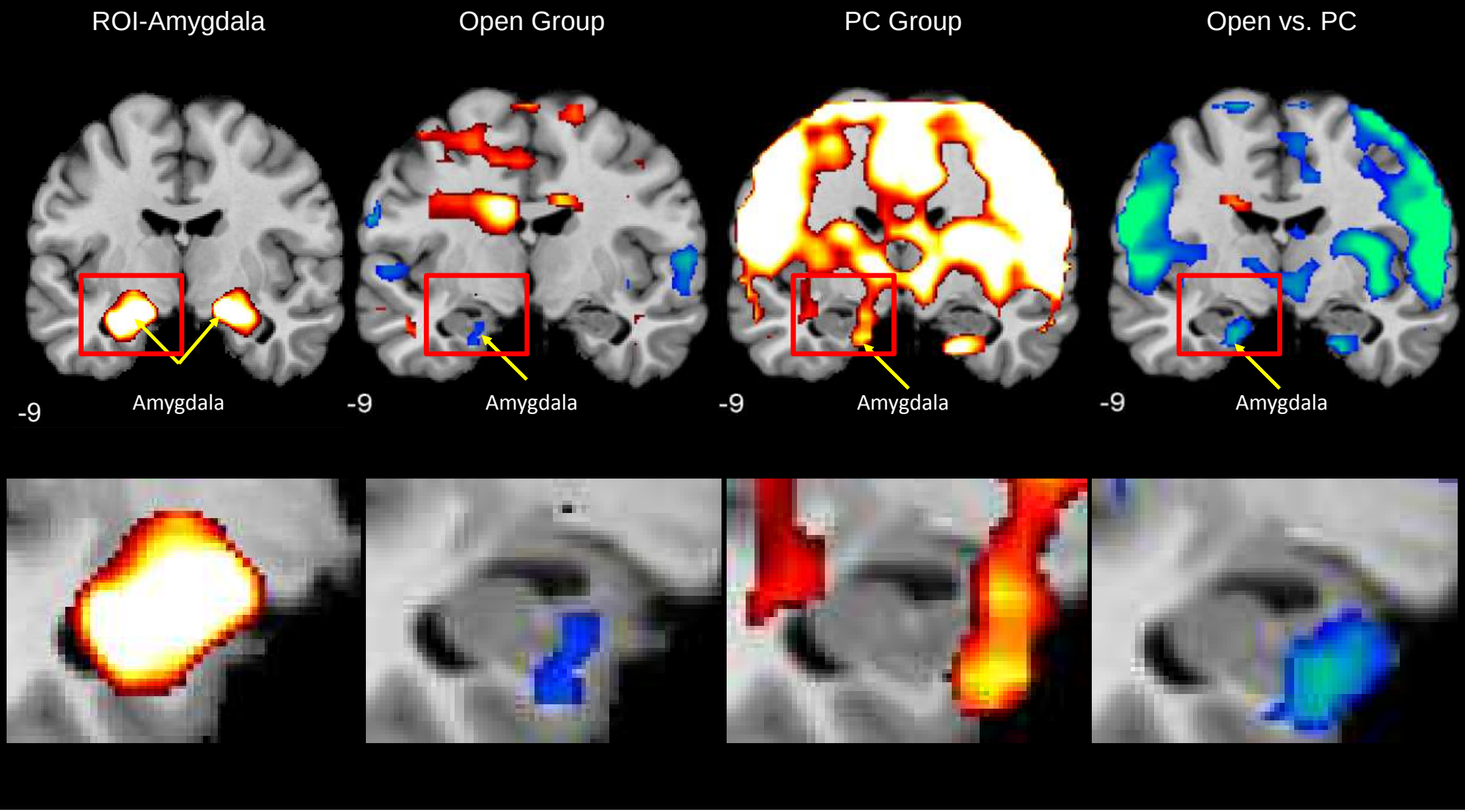


Figure 3. Within- and Between-group neural activation maps for the sad vs. neutral face contrast. To the left is depicted the amygdala region-of-interest, and the panels moving left to right depict within- and between-group change in neural activation from pre- to post-randomization.



Conflict of interest

Drs. Zilcha-Mano and Rutherford had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Drs. Rutherford, Zilcha-Mano, Wager, Peterson, Wang, Wall, Brown, Roose, and Chen have no disclosure or conflict of interest to report. This paper has not been previously presented.

Acknowledgments

Work on this paper was supported by NIMH grant K23 MH085236 (to Dr. Rutherford).