

SHORT TITLE: Effect of neurofeedback on emotional processing in MDD

**Real-Time fMRI Amygdala Neurofeedback Changes Positive Information Processing in
Major Depressive Disorder**

Kymberly D. Young Ph.D.^{ab*}, Masaya Misaki Ph.D.^a, Catherine J. Harmer D.Phil.^c, Teresa Victor
Ph.D.^a, Vadim Zotev, Ph.D.^a, Raquel Phillips B.S.^a, Greg J Siegle Ph.D.^b, Wayne C. Drevets M.D.^d,
Jerzy Bodurka Ph.D.^{ae}

^a Laureate Institute for Brain Research, Tulsa, OK

^b University of Pittsburgh School of Medicine, Pittsburgh, PA

^c Department of Psychiatry, University of Oxford, Medical Sciences Division, Oxford, UK

^d Janssen Research and Development, LLC, of Johnson & Johnson, Inc., Titusville, NJ

^e Stephenson School of Biomedical Engineering, University of Oklahoma College of Engineering,
Norman, OK

*Corresponding author: Western Psychiatric Institute and Clinic, 3811 O'Hara Street, Pittsburgh,
PA 15213; phone number:(412)-648-6179; fax number(412) 586-59255; email youngk@pitt.edu

Key Words: neurofeedback; emotional processing; depression; amygdala; backward masking;
fMRI

Abstract word count: 249

Manuscript word count: 3974

Number of Figures: 4

Number of Tables: 1

Supplementary Information: 6

Abstract

Background: In participants with major depressive disorder(MDD) trained to upregulate their amygdala hemodynamic response during positive autobiographical memory recall with real-time fMRI neurofeedback(rtfMRI-nf) training, depressive symptoms diminish. This study tested whether amygdala rtfMRI-nf also changes emotional processing of positive and negative stimuli in a variety of behavioral and imaging tasks.

Method: Patients with MDD completed two rtfMRI-nf sessions(18 received amygdala rtfMRI-nf, 16 received control parietal rtfMRI-nf). One week prior-to and following rtfMRI-nf training participants performed tasks measuring responses to emotional stimuli including a backward masking task(BMT), which measures the amygdala hemodynamic response to emotional faces presented for traditionally subliminal duration and followed by a mask, and the Emotional Test Battery(ETB) in which reaction times and performance accuracy are measured during tasks involving emotional faces and words.

Results: During the BMT, amygdala responses increased while viewing masked happy faces, but decreased to masked sad faces in the experimental versus control group following rtfMRI-nf. During the ETB, reaction times decreased to identification of positive faces and during self-identification with positive words, and vigilance scores increased to positive faces and decreased to negative faces during the face dot-probe task in the experimental versus control group following rtfMRI-nf.

Conclusions: rtfMRI-nf training to increase the amygdala hemodynamic response to positive memories was associated with changes in amygdala responses to happy and sad faces and improved processing of positive stimuli during performance of the ETB. These results may suggest that amygdala rtfMRI-nf training alters responses to emotional stimuli in a manner similar to antidepressant pharmacotherapy.

The research protocol is registered as a clinical trial on [clinicaltrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT02079610): NCT02079610.

Introduction

Major depressive disorder(MDD) is characterized by altered processing of positive and negatively valenced stimuli(1). These findings have been reported in behavioral studies in which memory and attention are enhanced for negative stimuli and reduced for positive stimuli in MDD relative to healthy control participants(2-5), as well as in neurophysiological studies in which hemodynamic responses of limbic brain regions are increased to negative and decreased to positive stimuli in MDD relative to control participants(6-10). It is hypothesized that this pattern of responding to emotional information in MDD is critical to the onset and maintenance of the disorder(11-13). Emerging evidence suggests that antidepressant drug treatments exert therapeutic effects, in part, by normalizing processing of emotional stimuli, resulting in increased memory for and attention to positive information(4, 14), increased amygdala activity to positive stimuli(8, 15), and decreased amygdala activity to negative stimuli(8, 16). Furthermore, antidepressant drugs have been shown to alter processing towards positive valence prior to changes in mood(17), consistent with changing emotional information processing being a potential mechanism of action. Cognitive therapies have also reported changes in limbic activity to positive and negative stimuli(18, 19) as well as reductions in negative cognitive biases(20), suggesting that changing emotional information processing is important for remission from MDD regardless of the type of therapy(cognitive or pharmacological).

Additional literature suggests the existence and importance of a normative *positive* emotional processing bias(21, 22) that may be supported by subcortical brain networks in healthy individuals(23). We recently developed an intervention for MDD that directly targets a neurobiological mechanism believed to contribute to this bias, real-time fMRI neurofeedback(rtFMRI-nf) training(24) of amygdala hemodynamic activity during positive memory recall. In a pilot study using this intervention, individuals with MDD successfully learned to increase their amygdala hemodynamic response during positive memory recall, and this learning was associated with improvements in state-dependent measures of mood(25). In the first randomized double-blind clinical trial of amygdala rtFMRI-nf for MDD we found that two rtFMRI-nf sessions

resulted in large and significant clinical improvements relative both to baseline and to a control group receiving rtfMRI-nf from a parietal region putatively not involved in emotion processing(26).

The current study, conducted as a secondary analysis on data from a clinical trial of neurofeedback that has already been published(26), tested the specificity and generalizability of the intervention by examining the effects of rtfMRI-nf training on tasks of emotional processing known to be affected by antidepressant treatments—the backward masking task(BMT) and the Emotional Test Battery(ETB). Testing generalizability and specificity is a critical step for determining the clinical potential of this procedure(27). We hypothesized that the amygdala's response to subliminally presented happy faces during the BMT would increase following rtfMRI-nf training and that accuracy would increase while reaction times decrease during tasks involving positive faces and words in the ETB.

Materials and Methods

Participants

Thirty-four right-handed, unmedicated adults ages 18-55 who met the Diagnostic and Statistical Manual of Mental Disorders(DSM-IV-TR)(28) criteria for MDD in a current major depressive episode completed the study. Volunteers, recruited from the community via advertisements, underwent screening evaluations at the Laureate Institute for Brain Research, including the Structural Clinical Interview for *DSM-IV* disorders(29). Exclusion criteria included current pregnancy, general MRI exclusions, serious suicidal ideation, psychosis, major medical or neurological disorders, exposure to any medication likely to influence cerebral function or blood flow within three weeks, and meeting *DSM-IV* criteria for drug/alcohol abuse within the previous one year or for alcohol/drug dependence(excepting nicotine) within the lifetime. All volunteers were naive to rtfMRI-nf. All subjects gave written informed consent to participate in the study and received financial compensation. The research protocol was approved by the Western Institutional Review Board.

Procedure

Participants were randomly assigned under double-blind conditions to the experimental group receiving amygdala rtfMRI-nf($n=18$) or the control group receiving rtfMRI-nf from the left horizontal segment of the intraparietal sulcus($n=16$) and completed four study visits. During the Baseline Visit(Visit 1) participants were administered the Montgomery-Asberg Depression Rating Scale(MADRS)(30), selected as the primary outcome, and performed the BMT during fMRI and the ETB immediately following completion of fMRI. Between 5-7 days later participants completed their first rtfMRI-nf session(Visit 2), followed 5-7 days later by their second rtfMRI-nf session(Visit 3) during which participants were instructed to recall positive autobiographical memories in order to increase a red bar representing activity of the assigned region(amygdala or parietal). As the focus of the current analysis is on the change observed in emotional processing, details of the rtfMRI-nf procedure can be found in our previous publications(24-26) and in the Supplement. The Follow-Up Visit(Visit 4) was completed 5-7 days following Visit 3 and was identical to the baseline Visit 1. The 21 item Hamilton Depression Rating Scale(HAM-D)(31) and Beck Depression Inventory(BDI-II)(32) were also completed at each visit(results in Table ST1)

Backward Masking Task(BMT)

Participants performed a BMT while undergoing fMRI that has been previously validated as a paradigm that results in altered hemodynamic activity to face stimuli that are not consciously perceived in both healthy and depressed participants(8). Prior to each of two 9min 8sec runs, participants were shown two neutral target faces and instructed to remember them and respond as quickly as possible to indicate if the presented face matched one of the target faces based on identity(not emotional expression). Two faces were presented for each stimulus, a “masked” face for 26ms, immediately followed by a “masking face” for 107ms, a fixation cross for 1866ms, and an 10-13sec interstimulus interval. In total, 48 stimuli(6 types [sad/ neutral, happy/neutral, neutral/sad, neutral/happy, neutral/neutral female, neutral/neutral male] x 8 presentations each) were presented for each run in a pseudo-randomized, mixed-trial design(Figure 1). Each event type was gender matched. Two different stimulus face sets equivalent on ratings of valence were counterbalanced

and used for the baseline and follow-up sessions. Faces were selected from the NimStim Set of Facial Expressions(33).

Image preprocessing and analysis were performed using AFNI(<http://afni.nimh.nih.gov>) and consisted of despiking, slice timing correction, and within-subject realignment. The anatomical image was registered to the first functional image then spatially normalized to the TT_N27 template using Advanced Normalization Tools with SyN method(34). The estimated warping parameter was used to normalize the functional images. The template image was resampled to 1.75 mm³ isotropic voxels, resulting in the spatially normalized image voxel size of 1.75mm³. Images were spatially smoothed using a 4-mm full-width half-maximum Gaussian kernel. Voxel-wise, the fMRI signal was scaled to percent signal change relative to the mean signal across time. A general linear model analysis was performed using 3dDeconvolve. For each participant, the hemodynamic response to each event type was modeled with a delta function at the event onset and convolved with the gamma-variate hemodynamic response function. Regressors modeling the task, motion parameters, and 4th-order polynomial regressors were used in the model. Because the masked and unmasked faces for each stimulus pair were presented too closely in time to model the response to each component separately, the data were modeled as event-related correlates of the combined stimulus pairs. The main effects of interest were presentation of sad/neutral(SN), happy/neutral(HN), and neutral/neutral(NN) faces. Events with target faces of sad and happy in the unmasked position were also modeled separately and included in the design matrix. 3dREMLfit was then used to generate the generalized least squares time series fit for each participant. The output from this program was used for group comparisons.

We defined regions-of-interest(ROI) for the amygdala using the left and right Talairach amygdala masks provided within AFNI. For each participant, the 3dREMLfit output was resampled and the amygdala mask applied to calculate the percent signal change within the ROI for the SN-NN and HN-NN conditions. The resulting percent signal change values were entered into a linear mixed effects model with the fixed factors of visit(Baseline, Follow-up), condition(HN-NN, SN-NN), lateralization(Left, Right), and group(Experimental, Control) using SPSS 23(IBM Corp., USA).

Autocorrelations were modeled with the ARMA1 structure, as this minimized Akaike's information criterion, with participant as the random effect. Associated t-tests were performed to characterize significant differences underlying main effects and interactions. For exploratory reasons, a supplementary whole-brain analysis was performed using 3dLME with the significance criterion set at $p < 0.05$ corrected (determined using AFNI 3dClustSim at voxel $p < 0.001$ and the Spatial AutoCorrelation Function to address recent criticisms of the cluster method(35), cluster size > 32 ; Table S3).

Emotional Test Battery(ETB)

The P1vital® Oxford Emotional Test Battery consists of four tasks designed to measure emotional processing. These tasks were completed outside of the fMRI following completion of the BMT. For each task, two different stimulus face/word sets equivalent on ratings of valence were counterbalanced and used for the baseline and follow-up sessions. Significance testing for this battery was performed using a linear mixed effects model and the ARMA1 covariance structure set at $p < 0.003$, Bonferroni corrected for 16 comparisons.

Facial Expression Recognition Task(FERT). The FERT assesses interpretation of facial expressions. Faces with six different emotions(happiness, surprise, sadness, fear, anger, disgust) were taken from 10 individual characters from the Pictures of Facial Affect series(36) and morphed between each prototype and neutral in 10% steps(37). Four examples of each emotion at varying intensity levels were presented(6 emotions x 10 intensities x 4 examples). Each face was also presented with a neutral expression(10 stimuli) giving a total of 250 stimulus presentations. Faces were presented in random order for 500ms then replaced by a blank screen, during which participants had an unlimited time to select the expression on the face by pressing one of seven labeled keys. Happy and surprise faces were combined to create a score for Positive faces(as they show similar effects opposite to those of the negative faces both in previous studies(38, 39) and in the current dataset; Table ST2), and sadness, fear, anger, and disgust were combined to create a score for Negative faces, as has been previously used as primary outcome measures(40). Accuracy and reaction times for correct responses were measured and entered into a linear mixed

effects model with the fixed factors of visit(Baseline, Follow-up), emotion(Positive, Negative), and group(Experimental, Control).

Emotional Categorization and Emotional Recall Tasks(ECAT, EREC). The ECAT assesses speed to respond to positive and negative self-referent personality descriptors. Sixty personality characteristics selected to be disagreeable(e.g., hostile, untidy) or agreeable(e.g., honest, cheerful)(41) were presented for 500ms each. Participants categorized each trait as something they would like or dislike to overhear someone else referring to them as possessing, as quickly and accurately as possible. As the task is designed for ceiling levels of accuracy only reaction times for correct identification were measured and entered into a linear mixed effects model with the fixed factors of visit(Baseline, Follow-up), emotion(Positive, Negative), and group(Experimental, Control).

The EREC measures encoding of the emotional stimuli during the ECAT. Fifteen minutes after completion of the ECAT, participants had 4min to recall as many of the personality traits as possible. The number of positive and negative words recalled were computed and analyzed for both correct and false responses. A linear mixed effects model was conducted using the fixed factors of visit(Baseline, Follow-up), emotion(Positive, Negative), and group(Experimental, Control) to assess the effects of the rtfMRI-nf intervention on accuracy and intrusions.

Faces Dot Probe Task(FDOT). The FDOT measures attention to positive versus negative stimuli using reaction times. Pairs of photographs were presented comprising of either one emotional(happy or fearful) and one neutral facial expression or two neutral expressions. Faces appeared above and below a central fixation point. For unmasked trials, the pair was displayed for 100ms, and for masked trials the pair was displayed for 16ms then replaced with a scrambled face-mask for 84ms. For all trials, images were then replaced with a probe, located in the position of one of the stimuli. The probe consisted of a pair of dots either in a vertical(:) or horizontal(··) orientation, and participants selected the correct orientation. There were 32 happy+neutral, 32 fearful+neutral, and 32 neutral+neutral pairs for both unmasked and masked trials. Vigilance scores were calculated by subtracting reaction times from trials when probes appeared in the same position as

the emotional face(congruent) from trials where probes appeared in the opposite position to the emotional face(incongruent). Higher vigilance scores(more positive) indicate more attention toward the emotional relative to the neutral face. A linear mixed effects model with the fixed factors of visit(Baseline, Follow-up), emotion pair(Positive/Neutral, Negative/Neutral), masking(Masked, Unmasked), and group(Experimental, Control) to assess the effect of the rtfMRI-nf intervention on vigilance scores.

Correlation Analysis

For all variables found to significantly differ between groups at follow-up, Pearson partial correlations were calculated between these variables and the change in MADRS score as well as neurofeedback success, controlling for baseline performance on these measures. Neurofeedback success was defined as the mean percent signal change in the amygdala ROI from the Baseline Neurofeedback Run 2 to the final Transfer Neurofeedback Run. Higher scores indicate more amygdala activity following training relative to baseline. Results are Bonferroni corrected for 12 comparisons, with significance set at $p < 0.004$. We decided to examine the correlation in both groups, instead of just the experimental group, to increase the variance in our data and power to detect associations, and because some participants in the control group were able to increase their amygdala response. It is important to understand the relationship with changing amygdala BOLD activity, regardless of whether or not participants received explicit feedback from the amygdala.

Results

Demographic and Clinical Characteristics

The clinical results have been published previously(26). Briefly, the experimental and control groups did not differ at baseline on measured demographic or clinical characteristics(Supplementary Table S1). At follow-up the experimental group had significantly

lower MADRS scores than the control group, with six participants in the experimental group and one in the control group meeting criteria for remission at study end, yielding a Number Needed to Treat=4.

Backward Masking Task Amygdala ROI Analysis

For the linear mixed model The Condition x Visit x Group interaction was significant($F(1,182)=5.63$, $p=0.02$; Figure 2). There was no main effect or interaction with laterality($F_s<1.05$, $p_s>0.31$), and therefore the signal from each side of the amygdala was averaged together for each contrast. Groups did not differ at baseline on amygdala activity during the SN-NN or HN-NN conditions($t_s(32)<1.41$, $p_s>0.17$, $d_s<0.16$). At follow-up, the experimental group had significantly lower amygdala activity during the SN-NN condition relative to the control group($t(32)=3.35$, $p=0.002$, $d=0.52$) and higher amygdala activity during the HN-NN condition($t(32)=3.89$, $p=0.001$, $d=1.13$). Paired-samples t-tests confirmed that amygdala activity did not change from baseline to follow-up in the control group($t_s(15)<0.44$, $p_s>0.67$, $d_s<0.20$), but decreased significantly in the experimental group during the SN-NN condition($t(17)=3.13$, $p=0.003$, $d=0.91$) and increased significantly during the HN-NN condition($t(17)=3.44$, $p=0.002$, $d=1.47$).

Emotional Test Battery

Table 1 provides group means for each visit and each task of the ETB.

Facial Expression Recognition Task. When examining accuracy for the linear mixed model there were no interactions with Group($F_s>2.33$, $p_s>0.14$). Rather, there was a main effect of Emotion($F(2,126)=72.6$, $p<0.001$), with neutral faces correctly classified more often than negative or positive faces($t_s(33)>6.77$, $p_s<0.001$, $d_s>0.75$), and positive faces correctly classified more often than negative faces($t(33)=8.50$, $p<0.001$, $d=0.91$). There was also a main effect of Visit($F(1,32)=16.5$, $p<0.001$) with performance during follow-up better than that during baseline($t(33)=4.01$, $p<0.001$, $d=0.40$), showing a training effect.

When examining reaction time in the linear mixed model, there was a significant Emotion x Visit x Group interaction($F(2,134)=10.9$, $p<0.001$). Groups did not differ from each other at baseline in reaction time to respond to faces of any emotion($t_s(32)<1.03$, $p_s>0.31$, $d_s<0.26$). At follow-up,

the experimental relative to the control group had significantly reduced reaction times in response to positive faces($t(32)=3.64$, $p=0.001$, $d=0.83$), but did not differ in reaction time to negative or neutral faces($t(32)<1.26$, $ps>0.22$, $ds<0.32$). Paired-samples t-tests confirmed that reaction times did not change from baseline to follow up in the control group($t(15)<1.66$, $ps>0.12$, $ds<0.35$). In the experimental group, reaction times decreased significantly to positive faces($t(17)=4.60$, $p<0.001$, $d=1.22$), but did not significantly change to negative or neutral faces($t(17)<1.57$, $ps>0.14$, $ds<0.22$).

Emotional Categorization and Recall Tasks. For the linear mixed model, there was a significant Emotion x Visit x Group interaction($F(1,82)=15.9$, $p<0.001$). Groups did not differ from each other at baseline($t(32)<0.99$, $ps>0.33$, $ds<0.24$). At follow-up the experimental group had decreased reaction times relative to the control group to positive words($t(32)=3.57$, $p=0.002$, $d=0.87$) but no response difference to negative words($t(32)=0.92$, $p=0.37$, $d=0.23$). Paired-samples t-tests confirmed that reaction times did not change from baseline to follow up in the control group($t(15)<0.95$, $ps>0.36$, $ds<0.14$), but decreased significantly to positive words in the experimental group($t(17)=4.95$, $p<0.001$, $d=1.26$), and did not change to negative words($t(17)=0.65$, $p=0.52$, $d=0.13$).

When examining the recall of these words later during the session, for the linear mixed model there was a main effect of Emotion for both accuracy and intrusions($F_s(1,96)>18.6$, $ps<0.001$) with negative words recalled more than positive($t(33)=4.41$, $ps<0.001$, $ds>0.65$; both accurately and for false positive intrusions). There was also a main effect of Visit for accuracy($F(1,96)=14.6$, $p<0.001$) with performance at follow-up better than that at baseline($t(32)=4.25$, $p<0.001$, $d=0.46$). There was no main effect of or interaction with Group for accuracy or intrusions($F_s<0.82$, $ps>0.37$).

Faces Dot Probe Task. For the linear mixed model, the Emotion x Visit x Group interaction was significant($F(1,169)=11.9$, $p=0.001$; Figure 3). Groups did not differ from each other on vigilance scores at baseline($t(32)>1.26$, $ps>0.22$, $ds<0.10$). At follow-up, regardless of masking, vigilance to positive faces increased in the experimental relative to control group($t(32)=4.72$,

$p=0.002$, $d=1.13$), and decreased to negative faces($t(32)=4.50$, $p=0.003$, $d=0.64$). Paired-samples t -tests confirmed that in the control group, vigilance scores did not change from baseline to follow-up($t(15)<0.98$, $ps>0.34$, $ds<0.06$) while in the experimental group vigilance scores increased to positive faces($t(17)=4.78$, $p<0.001$, $d=1.15$) and decreased to negative faces($t(17)=4.12$, $p=0.001$, $d=0.64$).

Correlations with clinical change and neurofeedback success

Controlling for all baseline scores, the change in amygdala activity to the HN-NN condition was significantly correlated both with the change in MADRS scores($r=-0.41$, $p=0.002$ Figure 4a) and neurofeedback success($r=-0.39$, $p=0.003$ Figure 4b), while there was no correlation between the change in amygdala activity to the SN-NN condition and either of these measures($rs>0.11$, $ps>0.57$). With respect to the ETB, the change in MADRS score was significantly correlated with the change in reaction time to positive words during the ECAT($r=0.39$, $p=0.004$ Figure 4c) and with the change in vigilance scores to positive faces during the FDOT task($r=-0.44$, $p=0.002$; Figure 4e), controlling for all baseline scores. Neurofeedback success was also significantly correlated with the change in reaction time to positive words during the ECAT($r=0.44$, $p=0.002$; Figure 4d). No other correlation was significant with MADRS scores or neurofeedback success after controlling for baseline performance($rs<0.28$, $ps>0.15$).

Discussion

Results of this study suggest that in patients with MDD, neurofeedback training to increase the amygdala hemodynamic response to positive autobiographical memories results in a selectively altered pattern of amygdala responding to other emotional stimuli, and also affects multiple indicators of positive information processing. Our neurofeedback training only targets increasing the amygdala hemodynamic response during positive memory recall. Despite this focus, the amygdala response assessed in the BMT normalized following training, evinced by an *increased* response to subliminally presented *happy* faces and a *decreased* response to subliminally presented *sad* faces. This finding suggests that training the amygdala response in one direction (upregulate to positive) at least partly generalized to the processing of other types of emotional stimuli in the amygdala, and that patients are learning to control their amygdala response rather than nonspecifically increase it. We previously showed a crucial role for the mPFC in the successful up-regulation of amygdala BOLD activity(43). This suggests that bidirectional control may be due to increased amygdala-medial prefrontal cortex(mPFC) connectivity. In addition, the mPFC can increase amygdala activity in response to salient stimuli in order to orient attention(44), and decrease amygdala activity in response to negative stimuli(45). Collectively, these findings suggest that our amygdala rtfMRI-nf procedure provides an effective means of strengthening cortical connections to the amygdala, which can alter its function in either direction. Furthermore, the change in amygdala activity to implicitly presented sad and happy faces following amygdala rtfMRI-nf appears analogous to that seen following antidepressant pharmacotherapy, which has reported increased amygdala activity in response to masked happy faces and decreased amygdala activity in response to masked sad faces following 8 weeks of treatment with sertraline(8).

Performance on ETB tasks also changed in a manner similar to that seen following antidepressant pharmacotherapy, with decreased reaction times following neurofeedback training to positive faces and words during the FERT and ECAT along with increased attention towards positive faces, and decreased attention towards negative faces, during the FDOT in the

experimental relative to the control group. Decreased reaction times to positive words during the ECAT has been observed following short-term antidepressant drug administration in healthy(42) and depressed(4) individuals. Decreased vigilance to negative stimuli and increased vigilance to positive stimuli during the FDOT task has also been observed following short-term antidepressant treatment in healthy volunteers(46, 47). Reduced reaction time to identify positive emotional faces during the FERT has been found in depressed patients following treatment with transcranial direct current stimulation of the dorsolateral prefrontal cortex(48). The majority of these studies employed a cross-sectional design. The longitudinal within-subjects design of the current study allows us to extend this literature by replicating previous findings in a pre-post test design and further supporting the hypothesis that treatments for MDD changes positive emotional processing(17, 49).

Though we replicated previous effects of antidepressant treatments on *reaction times* to emotional stimuli, we did not find any difference in accuracy or memory recall following neurofeedback. This could be evidence that neurofeedback training changes the automatic processing of emotional stimuli, as reaction times to positive stimuli decreased without affecting accuracy. However, this lack of effect could also be due to the test-retest practice effects in the longitudinal design employed in the current study, as accuracy during the FERT and EREC tasks increased in both the experimental and control group at follow-up. A previous study which used a within-subjects design to examine changes in FERT performance following antidepressant drug administration found an overall increase in accuracy. However the lack of a placebo control group precludes determination of whether increased accuracy was a practice effect that would have been observed in a placebo control group as well as the antidepressant groups(50). Future studies examining emotional processing changes using the ETB may benefit from employing placebo-controlled longitudinal designs to fully characterize the changes in emotional processing.

Though changes in the processing of negative stimuli were evident following neurofeedback training, with decreased amygdala activity to masked sad face presentations during the BMT and decreased vigilance scores to negative faces during the FDOT, only changes to positive stimuli were associated with measures of clinical improvement and neurofeedback

success. This suggests that it is the enhancement of the processing of positive emotional information, rather than suppression of the processing of negative, that underlies the clinical effects of amygdala neurofeedback.

Several limitations to the current study merit comment. Our entrance criteria resulted in a large proportion of patients being excluded (primarily due to medication status or not meeting DSM-IV-TR criteria for MDD), limiting the generalizability of our findings. Additionally, our relatively small sample lacked the statistical power to examine behavioral/symptom/demographic/brain factors that might moderate neurofeedback success. Further testing in larger more heterogeneous samples that include medicated individuals is necessary to determine the sub-populations or characteristics for whom this intervention is best suited.

In conclusion, amygdala neurofeedback training in individuals with MDD increases emotional processing of positive information and normalizes the amygdala response to emotional stimuli in a manner similar to antidepressant treatments. This result is consistent with the hypothesis that the mechanism through which amygdala neurofeedback training produces an antidepressant effect involves a change in positive emotional processing. It is possible that the efficacy of conventional therapies for MDD may be enhanced by neurocircuitry-based approaches that more specifically assist depressed individuals to enhance the processing of positive stimuli and events (12, 51). These results also support the clinical utility of fMRI amygdala neurofeedback training by showing that this training generalizes to other emotional stimuli and results in changes in amygdala activity and emotional processing that persists beyond the neurofeedback training sessions.

Acknowledgements:

This work was supported by the National Institutes of Health/ National Institute of Mental Health under Award Number K99MH101235 and by a NARSAD Young Investigator Grant from the

Brain & Behavior Research Foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Financial Disclosures

WCD is currently an employee of Janssen Research & Development, LLC, of Johnson & Johnson, Inc., and also holds equity in Johnson & Johnson, Inc. CJH has received consulting fees from P1Vital and Lundbeck, and grant income from Janssen Research & Development, LLC,

Sunovion Pharmaceuticals, Inc., and UBC, Inc. The other authors have no financial conflicts of interest or disclosures to report. Dr. Young takes responsibility for the integrity of the data and accuracy of the data analysis. All authors had full access to all the data in the study.

References

1. Chamberlain SR, Sahakian BJ. The neuropsychology of mood disorders. *Current psychiatry reports*. 2006;8:458-463.
2. Bradley BP, Mogg K, Williams R. Implicit and explicit memory for emotion-congruent information in clinical depression and anxiety. *Behav Res Ther*. 1995;33:755-770.
3. Gotlib IH, Krasnoperova E, Yue DN, Joormann J. Attentional biases for negative interpersonal stimuli in clinical depression. *J Abnorm Psychol*. 2004;113:121-135.

4. Harmer CJ, O'Sullivan U, Favaron E, Massey-Chase R, Ayres R, Reinecke A, Goodwin GM, Cowen PJ. Effect of acute antidepressant administration on negative affective bias in depressed patients. *Am J Psychiatry*. 2009;166:1178-1184.
5. Young K, Bellgowan P, Bodurka J, Drevets WC. Behavioral and Neurophysiological Correlates of Autobiographical Memory Deficits in Patients With Depression and Individuals at High Risk for Depression. *Journal of the American Medical Association psychiatry*. 2013;70:698-708.
6. Canli T, Sivers H, Thomason ME, Whitfield-Gabrieli S, Gabrieli JD, Gotlib IH. Brain activation to emotional words in depressed vs healthy subjects. *Neuroreport*. 2004;15:2585-2588.
7. Suslow T, Konrad C, Kugel H, Rumstadt D, Zwitterlood P, Schoning S, Ohrmann P, Bauer J, Pyka M, Kersting A, Arolt V, Heindel W, Dannlowski U. Automatic mood-congruent amygdala responses to masked facial expressions in major depression. *Biol Psychiatry*. 2010;67:155-160.
8. Victor TA, Furey ML, Fromm SJ, Ohman A, Drevets WC. Relationship between amygdala responses to masked faces and mood state and treatment in major depressive disorder. *Arch Gen Psychiatry*. 2010;67:1128-1138.
9. Surguladze S, Brammer MJ, Keedwell P, Giampietro V, Young AW, Travis MJ, Williams SC, Phillips ML. A differential pattern of neural response toward sad versus happy facial expressions in major depressive disorder. *Biol Psychiatry*. 2005;57:201-209.
10. Young KD, Siegle GJ, Bodurka J, Drevets WC. Amygdala activity during autobiographical memory recall in depressed and vulnerable individuals; Association with symptom severity and autobiographical overgenerality. *Am J Psychiatry*. 2016;173:78-89.
11. Bouhuys AL, Geerts E, Gordijn MC. Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: a longitudinal study. *J Nerv Ment Dis*. 1999;187:595-602.
12. Beck AT, Rush AJ, Shaw BF, Emery G: *Cognitive Therapy of Depression*. New York, N.Y., The Guilford Press; 1979.
13. Mathews A, MacLeod C. Cognitive vulnerability to emotional disorders. *Annu Rev Clin Psychol*. 2005;1:167-195.
14. Harmer CJ. Serotonin and emotional processing: does it help explain antidepressant drug action? *Neuropharmacology*. 2008;55:1023-1028.
15. Norbury R, Taylor MJ, Selvaraj S, Murphy SE, Harmer CJ, Cowen PJ. Short-term antidepressant treatment modulates amygdala response to happy faces. *Psychopharmacology*. 2009;206:197-204.
16. Fu CH, Williams SC, Cleare AJ, Brammer MJ, Walsh ND, Kim J, Andrew CM, Pich EM, Williams PM, Reed LJ, Mitterschiffthaler MT, Suckling J, Bullmore ET. Attenuation of the neural response to sad faces in major depression by antidepressant treatment: a prospective, event-related functional magnetic resonance imaging study. *Arch Gen Psychiatry*. 2004;61:877-889.
17. Harmer CJ, Goodwin GM, Cowen PJ. Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *Br J Psychiatry*. 2009;195:102-108.
18. Ritchey M, Dolcos F, Eddington KM, Strauman TJ, Cabeza R. Neural correlates of emotional processing in depression: changes with cognitive behavioral therapy and predictors of treatment response. *J Psychiatr Res*. 2011;45:577-587.
19. Siegle GJ, Ghinassi F, Thase ME. Neurobehavioral Therapies in the 21st Century: Summary of an Emerging Field and an Extended Example of Cognitive Control Training for Depression. *Cognit Ther Res*. 2007;31:235-262.

20. Lang TJ, Blackwell SE, Harmer CJ, Davison P, Holmes EA. Cognitive Bias Modification Using Mental Imagery for Depression: Developing a Novel Computerized Intervention to Change Negative Thinking Styles. *Eur J Pers.* 2012;26:145-157.
21. Schwartz RM, Caramoni GL. Cognitive balance and psychopathology: Evaluation of an information processing model of positive and negative states of mind. *Clinical Psychology Review.* 1989;9:271-294.
22. Fredrickson BL. The broaden-and-build theory of positive emotions. *Philos Trans R Soc Lond B Biol Sci.* 2004;359:1367-1378.
23. Chavez RS, Heatherton TF, Wagner DD. Neural Population Decoding Reveals the Intrinsic Positivity of the Self. *Cereb Cortex.* 2016.
24. Zotev V, Krueger F, Phillips R, Alvarez RP, Simmons WK, Bellgowan P, Drevets WC, Bodurka J. Self-regulation of amygdala activation using real-time fMRI neurofeedback. *PLoS One.* 2011;6:e24522.
25. Young K, Zotev V, Phillips R, Misaki M, Yuan H, Drevets W, Bodurka J. Real-time fMRI neurofeedback training of amygdala activity in patients with major depressive disorder. . *PLoS One.* 2014;8:e79184.
26. Young KD, Siegle GJ, Zotev V, Phillips R, Misaki M, Yuan H, Drevets WC, Bodurka J. Randomized Clinical Trial of Real-Time fMRI Amygdala Neurofeedback for Major Depressive Disorder: Effects on Symptoms and Autobiographical Memory Recall. *Am J Psychiatry.* In Press.
27. Sulzer J, Haller S, Scharnowski F, Weiskopf N, Birbaumer N, Blefari ML, Bruehl AB, Cohen LG, DeCharms RC, Gassert R, Goebel R, Herwig U, LaConte S, Linden D, Luft A, Seifritz E, Sitaram R. Real-time fMRI neurofeedback: progress and challenges. *Neuroimage.* 2013;76:386-399.
28. (APA) APA: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. . Washington, DC, American Psychiatric Association; 2000.
29. First MB, Spitzer RL, Gibbon M, Williams JBW: Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition(SCID-I/P). New York, NY, New York State Psychiatric Institute, Biometrics Research; 2002.
30. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry.* 1979;134:382-389.
31. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960;23:56-62.
32. Beck AT, Steer RA, Brown GK: Manual for the Beck Depression Inventory-II. San Antonio, TX, Psychological Corporation; 1996.
33. Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, Marcus DJ, Westerlund A, Casey BJ, Nelson C. The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res.* 2009;168:242-249.
34. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis.* 2008;12:26-41.
35. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci U S A.* 2016.
36. Ekman P, Wallace VF: Pictures of Facial Affect, Consulting psychologists Press; 1976.
37. Young AW, Rowland D, Calder AJ, Etcoff NL, Seth A, Perrett DI. Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition.* 1997;63:271-313.

38. Harmer CJ, O'Sullivan U, Favaron E, Massey-Chase R, Ayres R, Reinecke A, Goodwin GM, Cowen PJ. Effect of acute antidepressant administration on negative affective bias in depressed patients. *Am J Psychiatry*. 2009;166:1178-1184.
39. Alves-Neto WC, Guapo VG, Graeff FG, Deakin JF, Del-Ben CM. Effect of escitalopram on the processing of emotional faces. *Braz J Med Biol Res*. 2010;43:285-289.
40. Post A, Smart TS, Krikke-Workel J, Dawson GR, Harmer CJ, Browning M, Jackson K, Kakar R, Mohn R, Statnick M, Wafford K, McCarthy A, Barth V, Witkin JM. A Selective Nociceptin Receptor Antagonist to Treat Depression: Evidence from Preclinical and Clinical Studies. *Neuropsychopharmacology*. 2016;41:1803-1812.
41. Anderson NH. Likableness ratings of 555 personality-trait words. *Journal of personality and social psychology*. 1968;9:272-279.
42. Harmer CJ, Shelley NC, Cowen PJ, Goodwin GM. Increased positive versus negative affective perception and memory in healthy volunteers following selective serotonin and norepinephrine reuptake inhibition. *Am J Psychiatry*. 2004;161:1256-1263.
43. Zotev V, Phillips R, Young KD, Drevets WC, Bodurka J. Prefrontal control of the amygdala during real-time fMRI neurofeedback training of emotion regulation. *PLoS One*. 2013;8:e79184.
44. Liddell BJ, Brown KJ, Kemp AH, Barton MJ, Das P, Peduto A, Gordon E, Williams LM. A direct brainstem-amygdala-cortical 'alarm' system for subliminal signals of fear. *Neuroimage*. 2005;24:235-243.
45. Quirk GJ, Likhtik E, Pelletier JG, Pare D. Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *J Neurosci*. 2003;23:8800-8807.
46. Browning M, Reid C, Cowen PJ, Goodwin GM, Harmer CJ. A single dose of citalopram increases fear recognition in healthy subjects. *J Psychopharmacol*. 2007;21:684-690.
47. Murphy SE, Longhitano C, Ayres RE, Cowen PJ, Harmer CJ. Tryptophan supplementation induces a positive bias in the processing of emotional material in healthy female volunteers. *Psychopharmacology*. 2006;187:121-130.
48. Nitsche MA, Koschack J, Pohlers H, Hulleman S, Paulus W, Happe S. Effects of frontal transcranial direct current stimulation on emotional state and processing in healthy humans. *Front Psychiatry*. 2012;3:58.
49. Zotev V, Yuan H, Misaki M, Phillips R, Young KD, Feldner MT, Bodurka J. Correlation between amygdala BOLD activity and frontal EEG asymmetry during real-time fMRI neurofeedback training in patients with depression. *Neuroimage Clin*. 2016;11:224-238.
50. Tranter R, Bell D, Gutting P, Harmer C, Healy D, Anderson IM. The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. *J Affect Disord*. 2009;118:87-93.
51. Disner SG, Beevers CG, Haigh EA, Beck AT. Neural mechanisms of the cognitive model of depression. *Nat Rev Neurosci*. 2011;12:467-477.

Figure Captions

Figure 1: Design of the Backward Masking Task

- (a) Two neutral target faces were shown prior to the start of each run in which participants were instructed to remember and determine whether subsequent face presentations matched the identity. (b) Two faces were shown as part of each trial presentation. Examples of the masked faces

event types(SN, HN, NN) are shown with “N” placeholders to indicate the presentation of a neutral face.

HN = Happy/Neutral face presentation; ISI = interstimulus interval; NN = Neutral/Neutral face presentation; SN = Sad/Neutral face presentations

Figure 2: Amygdala hemodynamic signal during the Backward Masking Task for each group and visit.

Amygdala hemodynamic response was assessed using fMRI during exposure to:(a) masked sad face presentations(SN-NN condition) and(b) masked happy face presentations(HN-NN condition). Error bars indicate +/- one standard error of the mean. * indicates a significant difference from the corresponding baseline at $p_{\text{corrected}} < 0.05$. # indicates a significant difference from the experimental group at $p_{\text{corrected}} < 0.05$.

HN = Happy/Neutral face presentation; NN = Neutral/Neutral face presentation; SN = Sad/Neutral face presentation

Figure 3: Vigilance scores for Faces Dot Probe Task for each group and visit.

Vigilance Scores for(a) negative relative to neutral facial expressions and(b) happy relative to neutral facial expressions.

Error bars indicate +/- one standard error of the mean. * indicates a significant difference from the corresponding baseline at $p_{\text{corrected}} < 0.05$. # indicates a significant difference from the experimental group at $p_{\text{corrected}} < 0.05$.

Figure 4: Correlations between clinical and neurofeedback performance and changes in performance on the BMT and ETB.

(a) Correlation between the percent change in MADRS scores from baseline to follow-up and the change in amygdala activity to masked happy face presentation(HN-NN condition) during the BMT.(b) Correlation between neurofeedback success(defined as the change in amygdala activity

from the baseline neurofeedback run to the final neurofeedback transfer run) and the change in amygdala activity to masked happy face presentation(HN-NN condition) while performing the BMT.(c) Correlation between the percent change in MADRS scores from baseline to follow-up and the change in reaction time to positive faces during the ECAT.(d) Correlation between neurofeedback success and the change in reaction time to positive faces during the ECAT. e) Correlation between the percent change in MADRS scores from baseline to follow-up and the change in vigilance score to positive faces during the FDOT.

BMT= backward masking task; ECAT = emotional categorization task; ETB = emotional test battery; FDOT = faces dot probe task; MADRS = Montgomery-Asberg depression rating scale

Table 1: Performance on the Emotional Test Battery

	Experimental		Control	
	<i>Baseline</i>	<i>Follow-up</i>	<i>Baseline</i>	<i>Follow-up</i>
FERT				
Accuracy				
Positive	71.3(16.2)	75.6(4.29)	75.4(5.14)	75.6(5.25)
Negative	60.3(9.81)	61.8(10.7)	61.8(8.35)	63.6(7.96)
Neutral	84.2(18.6)	90.4(15.9)	86.3(12.1)	91.9(6.55)
RT(in ms)				
Positive	1660(254)	1423(254)*	1674(188)	1623(226)#
Negative	1400(305)	1356(305)	1465(162)	1440(203)
Neutral	1462(434)	1379(292)	1458(311)	1347(309)
ECAT				
Positive	1060(164)	928(118)*	1032(124)	1060(178)#
Negative	1069(222)	1087(229)	1116(166)	1136(201)
EREC				
Accuracy				

Positive	2.39(1.65)	3.28(1.81)	3.19(2.04)	3.25(1.23)
Negative	3.33(2.14)	3.78(2.21)	3.43(2.16)	3.81(2.04)
Intrusions				
Positive	1.17(1.34)	0.94(0.87)	1.63(1.89)	1.13(1.46)
Negative	2.83(1.95)	2.22(1.73)	3.00(1.86)	1.62(2.5)
FDOT				
Masked				
Positive	-5.28(29.6)	11.6(18.2)*	-6.56(25.2)	-7.81(32.8)#
Negative	5.94(29.2)	-16.0(24.2)*	9.00(35.6)	6.00(37.4)#
Unmasked				
Positive	-5.06(25.4)	35.3(22.6)*	-6.94(43.8)	-4.75(29.5)#
Negative	14.4(27.1)	-6.44(23.1)*	11.6(44.1)	8.88(29.6)#

Numbers in parentheses indicate one standard deviation of the mean. * indicates a significant difference from the corresponding baseline at $p_{\text{corrected}} < 0.05$. # indicates a significant difference from the experimental group at $p_{\text{corrected}} < 0.05$.

ECAT = emotional categorization task; EREC = emotional recognition task; FDOT = faces dot probe task; FERT = facial emotion recognition task; RT = reaction time.