**Supplementary Materials**

**Materials and Methods**

*Participants*

Exclusion criteria for patients were major active medical or neurological problems; past or current major psychiatric illness (e.g., bipolar disorder, schizophrenia); active suicidal ideation or non-suicidal self-injury in the last 6 months; use of psychotropic medication in the 6 weeks prior to study entry or during the study; cognitive dysfunction (e.g., traumatic brain injury, dementia, intellectual disability); pervasive developmental disorder (e.g., autism, learning disability); substance abuse or dependence in the last 6 months; contraindications for functional magnetic resonance imaging (fMRI) (e.g., ferrous metal, pregnancy, claustrophobia); and a positive toxicology test before the MRI scan.

*Treatment Procedures*

Patients randomized to CBT received manualized treatment for SAD or MDD that was consistent with their principal diagnosis (e.g., Beck et al., 1979; Hope et al., 2006; Martell et al., 2010). CBT treatment included psychoeducation, cognitive restructuring, behavioral activation (for MDD), in vivo exposures (for SAD), and relapse prevention. ST resembled client-centered therapy (e.g., Markowitz et al., 2008; Rogers, 1946) and emphasized reflective listening and elicitation of affect as appropriate.

Psychotherapy sessions were taped to assess treatment fidelity, and 15% of taped sessions were randomly selected and rated by assessors blinded to treatment arm using the Cognitive Therapy Rating Scale (CTRS; Young & Beck, 1980). The CTRS is an observer-rated measure of CBT fidelity with excellent psychometric properties (Dobson, Shaw, & Vallis, 1985; Vallis, Shaw, & Dobson, 1986) that assesses clinicians’ general psychotherapy skills (e.g., empathy) as well as the extent to which they utilize CBT-specific strategies (e.g., methods to alter maladaptive cognitions and behaviors) during treatment sessions. Given that CBT strategies are a core component of CBT but are explicitly not used in ST, clinician CTRS scores should be significantly higher when providing CBT versus ST. An independent t-test revealed average CTRS total score was higher in the CBT arm (*Mean* = 43.03, *SD* = 7.55) than the ST arm (*Mean* = 16.58, *SD* = 6.34), *t*(125) = 21.43, p < 0.001, *reffect size* = .89) indicting fidelity was acceptable.

*Diagnosis-Specific Symptom Outcome Measure*

Consistent with the decision to examine diagnosis-specific symptoms as a marker of treatment response, patient HAMD and LSAS scores were negatively correlated at baseline, *r =* -.37, *p =* .001, which is consistent with the fact that patients with SAD were not permitted to have clinically elevated levels of MDD symptoms, and vice versa. Patient HAMD and LSAS scores were not correlated post-treatment, *r =* .21, *p =* .07.

Within the specific patient groups, HAMD and LSAS scores were not correlated for SAD patients at baseline, *r =* .04, *p =* .84, or post-treatment, *r =* .23, *p =* .20. Additionally, HAMD and LSAS scores were also not correlated for MDD patients at baseline, *r =* .05, *p =* .77, though they were correlated post-treatment, *r =* .41, *p =* .007.

*fMRI Data Collection and Preprocessing*

Blood oxygen-level dependent (BOLD) functional images were acquired using a gradient-echo echo-planar imaging sequence with the following parameters: repetition time (TR)=2 s, echo time (TE)=25 ms, flip angle=90⁰, field of view=22 × 22 cm, acquisition matrix 64 × 64; 44 axial, 3-mm-thick slices with no gap. The first 4 volumes from each run were discarded to allow for T1 equilibration effects. For anatomical localization, a high-resolution, T1-weighted volumetric anatomical scan was acquired.

Conventional preprocessing steps were used in Statistical Parametric Mapping (SPM12) software package (Wellcome Trust Centre for Neuroimaging, London www.fil.ion.ucl.ac.uk/spm). Images were temporally corrected to account for differences in slice time collection, spatially realigned to the first image of the first run, coregistered to the anatomical, normalized to a Montreal Neurological Institute (MNI) space using warping based on the anatomical, resampled to 2 × 2 × 2 mm voxels, and smoothed with an 8 mm isotropic Gaussian kernel. All participants were required to have no movement greater than 3-mm in any direction for analysis.

A general linear model identifying the Look Neutral, Look Negative, and Reappraise Negative blocks was applied to the time series, convolved with the canonical hemodynamic response function and with a 128 s high-pass filter. First level models included 6 motion parameters as nuisance regressors.

*Brain Mask Construction*

 An anatomy-based fronto-temporal mask was created and applied to all second-level models in SPM12. Specifically, the mask comprised brain regions shown to consistently engage during situation-focused reappraisal of negative stimuli (Messina, Bianco, Sambin, & Viviani, 2015). Therefore, the following bilateral regions based on AAL 3 (Rolls, Huang, Lin, Feng, & Joliot, 2020) were combined to create the mask (volume = 383,816 mm3): superior frontal gyrus (medial and dorsolateral portions), middle frontal gyrus (i.e., dorsolateral prefrontal cortex), inferior frontal gyrus comprising opercular, triangular, and pars orbitalis portions (i.e., ventrolateral prefrontal cortex), supracallosal anterior cingulate cortex, middle cingulate (anterior to y = 0), angular gyrus, and middle temporal gyrus.

**Results**

*Behavioral Performance*

A 2 (Treatment Arm: CBT, ST) x 3 (Condition: Look Neutral, Look Negative, Reappraise Negative) mixed ANOVA with repeated measures for the last factor was conducted to verify participants followed instructions and to ensure there were no differences in reappraisal facility in participants before randomization to CBT or ST. Results showed a main effect of Condition, *F*(2, 72) = 214.93, *p* < .001. Pairwise comparisons with a Bonferroni correction indicated affective state significantly differed between each of the conditions (all *p*s < .001), such that patients reported the lowest levels of negative emotion in the Look Neutral condition (*Mean* = 1.17, *SE* = 0.36), followed by the Reappraise Negative (*Mean* = 2.42, *SE* = .090), and the Look Negative (*Mean* = 2.91, *SE* = .093) conditions. As expected, there was no main effect of Arm or a Condition × Arm interaction (all *p*s > .81).

*Overlap of Significant Clusters with Voxels Sensitive to Task Effects*

To confirm that significant clusters overlapped with voxels that were sensitive to the primary contrast of interest (Reappraise Negative > Look Negative), models in SPM were reconducted such that evaluation of significant clusters of brain activity identified in our primary analyses were further constrained to voxels that were sensitive to task effects using a voxel threshold of *p* < .005.

First, we examined the bilateral clusters of vlPFC activation that significantly corresponded with the Slope × Arm interaction for the model examining trajectories of diagnosis-specific symptoms. Results indicated that the majority of voxels within both vlPFC clusters were also sensitive to the Reappraise Negative > Look Negative contrast of interest (right peak [40, 12, 14], k=435 voxels, z=3.63, p<.001, *reffect size* = .42; left peak [-34, 22, 12], k=301 voxels, z=3.44, p<.001, *reffect size* = .39).

Second, we confirmed the extent to which the left prefrontal cluster that was significantly predicted by the main effect of reappraisal change trajectory overlapped with task-sensitive voxels. Results indicated that 83 voxels from this significant cluster overlapped with task-sensitive voxels (first peak [-30, 52, 24], k=79 voxels, z=2.90, p=.002; second peak [-44, 32, 22], k=4 voxels, z=2.72, p=.003), the majority of which were located within the dorsolateral superior frontal gyrus (k=79).

*Specificity of Findings to Reappraisal versus Affective Processing*

 Given that significant clusters identified within the context of the Reappraise Negative > Look Negative contrast may have been driven by individual differences in neural activation during either reappraisal (i.e., Reappraise Negative) or basic affective processing (i.e., Look Negative), additional analyses were conducted to aid in interpretation of findings. Specifically, spherical ROIs for significant clusters in primary analyses were extracted for the Reappraise Negative > Look Neutral and Look Negative > Look Neutral contrasts and were submitted to partial correlations in SPSS to confirm main findings.

 Reappraise Negative > Look Neutral. We first examined the clusters that were significant for the Slope × Arm interaction for the model examining trajectories of diagnosis-specific symptoms. Follow-up partial correlations indicated that for patients assigned to CBT, the association between trajectories of diagnosis-specific symptom reduction and baseline neural activation was approaching significance for the right vlPFC, *r*=.30, *p*=.068, but was not significant for left vlPFC, *r*=.25, *p*=.135. For patients randomized to ST, steeper trajectories of diagnosis-specific symptom reduction were associated with greater baseline activation of the right vlPFC, *r*=-.35, *p*=.049, but not left vlPFC, *r*=-.26, *p*=.150.

 Regarding the main effect of trajectories of reappraisal improvement and neural activation, partial correlation indicated that greater trajectories of self-reported reappraisal improvement were associated with less baseline activation of left dlPFC, *r*=-.27, *p*=.024.

Look Negative > Look Neutral. Follow-up partial correlations conducted separately for patients assigned to either CBT or ST revealed no significant associations between trajectories of diagnosis-specific treatment improvement and baseline activation of either right or left vlPFC (|*r|*s = .02 – .14, *p*s = .412 – .925).

Similarly, the partial correlation examining the association between trajectories of self-reported reappraisal improvement and baseline activation of the left dlPFC was not significant, *r*=.13, *p*=.281.

*Exploratory Analyses.*

 This study was significantly underpowered to test further moderation by participant diagnosis, particularly within the context of treatment arm differences (MDD patients randomized to CBT: *n* = 22; MDD and ST patients: *n* = 19; SAD and CBT patients: *n* = 18; SAD and ST: *n* = 16). However, exploratory analyses were conducted to examine whether primary findings regarding associations between baseline neural activation and trajectories of clinical improvement were maintained when examining the MDD and SAD groups separately.

 MDD Patient Group. Focusing first on findings for diagnosis-specific symptom improvement, exploratory partial correlations indicated that for MDD patients assigned to CBT, trajectories of diagnosis-specific symptom reduction were not associated with baseline activation of the left, *r*=-.12, *p*=.638, or right vlPFC, *r*=-.08, *p*=.748. Similarly, trajectories of diagnosis-specific symptom reduction were not associated with left, *r*=-.40, *p*=.124, or right vlPFC, *r*=-.26, *p*=.330, for patients assigned to ST.

 However, findings regarding the prediction of reappraisal trajectories were maintained for patients in the MDD group, such that greater trajectories of reappraisal improvement were associated with less baseline activation of the left dlPFC, *r*=-.45, *p*=.004.

SAD Patient Group. Exploratory partial correlations indicated that for SAD patients assigned to CBT, steeper trajectories of diagnosis-specific symptom reduction were associated with less baseline activation of the left vlPFC, *r*=.60, *p*=.019, but not the right vlPFC, *r*=.41, *p*=.128. However, for SAD patients assigned to ST, steeper trajectories of diagnosis-specific symptom reduction were associated with greater baseline activation of the right vlPFC, *r*=-.58, *p*=.039, but not the left vlPFC, *r*=-.28, *p*=.364.

 Regarding the finding for neural predictors of reappraisal improvement, greater trajectories of reappraisal improvement were associated with less baseline activation of the left dlPFC for SAD patients, *r*=-.38, *p*=.036.

**References**

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press.

Dobson, K. S., Shaw, B. F., & Vallis, T. M. (1985). Reliability of a measure of the quality of cognitive therapy. *British Journal of Clinical Psychology*, *24*(4), 295–300. https://doi.org/10.1111/j.2044-8260.1985.tb00662.x

Hope, D. A., Heimberg, R. G., & Turk, C. L. (2006). *Managing social anxiety: A cognitive-behavioral therapy approach*. Oxford University Press.

Markowitz, J. C., Manber, R., & Rosen, P. (2008). Therapists’ responses to training in brief supportive psychotherapy. *American Journal of Psychotherapy*, *62*(1), 67–81. https://doi.org/10.1176/appi.psychotherapy.2008.62.1.67

Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2010). *Behavioral Activation for Depression: A Clinician’s Guide*. Guilford Press.

Messina, I., Bianco, S., Sambin, M., & Viviani, R. (2015). Executive and semantic processes in reappraisal of negative stimuli: insights from a meta-analysis of neuroimaging studies. *Frontiers in Psychology*, *6*.

Rogers, C. R. (1946). Significant aspects of client-centered therapy. *The American Psychologist*, *1*(10), 415–422. https://doi.org/10.1037/h0060866

Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., & Joliot, M. (2020). Automated anatomical labelling atlas 3. *Neuroimage*, *206*, 116189. https://doi.org/10.1016/j.neuroimage.2019.116189

Vallis, T. M., Shaw, B. F., & Dobson, K. S. (1986). The cognitive therapy scale: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *54*(3), 381–385. https://doi.org/10.1037/0022-006X.54.3.381

Young, J., & Beck, A. T. (1980). *Cognitive Therapy Rating Scale*. Philadelphia.

Supplementary Table 1 *Intent to Treat Analysis*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Treatment Completers (*n* = 95) | Treatment Non-Completers(*n* = 22) | *t/χ2* |
| Demographics |  |  |  |
| Age | 28.28 (9.54) | 25.27 (6.39) | t = -1.41, p = .16 |
| Sex |  |  | χ2 = 1.06, p = .59 |
| Female | 68.4% | 59.1% |  |
| Male | 30.5% | 40.9% |  |
| Not Reported | 1.1% | 0.0% |  |
| Ethnicity (Hispanic or Latino) | 29.5% | 27.3% | χ2 = 0.04, p = .84 |
| Racial Identity |  |  | χ2 = 11.38, p = .08 |
| White | 50.5% | 27.3% |  |
| Black | 11.6% | 18.2% |  |
| Asian | 14.7% | 40.9% |  |
| Native American or Alaskan Native | 1.1% | 0.0% |  |
| Multi-Racial/ Another Identity | 22.1% | 13.6% |  |
| Clinical Characteristics |  |  |  |
| Assigned Treatment |  |  | χ2 = 0.49, p = .49 |
| CBT | 53.7% | 45.5% |  |
| ST | 46.3% | 54.5% |  |
| Diagnosis |  |  | χ2 = 1.89, p = .17 |
| SAD | 47.4% | 63.6% |  |
| MDD | 52.6% | 36.4% |  |
| Baseline Clinical Measures |  |  |  |
| HAMD | 11.74 (4.98) | 9.86 (5.97) | t = -1.53, p = .13 |
| LSAS | 57.93 (29.61) | 71.64 (26.81) | t = 1.99, p = .049 |
| Composite Symptoms | 0.91 (0.26) | 0.94 (0.20) | t = 0.50, p = .62 |
| Diagnosis-Specific Symptoms | 0.59 (0.17) | 0.63 (0.18) | t = 1.15, p = .26 |
| Reappraisal | 24.75 (6.32) | 25.33 (8.82) | t = 0.36, p = .72 |
| *Note.* Intent to treat analyses include all patients randomized to treatment. CBT = Cognitive Behavioral Therapy. ST = Supportive Therapy. SAD = Social Anxiety Disorder. MDD = Major Depressive Disorder. LSAS = Liebowitz Social Anxiety Scale. HAMD = Hamilton Depression Rating Scale. Composite Symptoms = Summation of Liebowitz Social Anxiety Scale (LSAS) and Hamilton Depression Rating Scale (HAMD) proportion of maximum scaling (POMS) scores. Diagnosis-Specific Symptoms = LSAS POMS scores for patients with SAD, HAMD POMS scores for patients with MDD.  |

Supplementary Table 2
*Results from Full Factorial Analyses*

|  |
| --- |
| **Composite Symptom Model** |
|  |  |  | Peak Coordinates |  |
|  | Anatomical Regions | k | x | y | z | z-score |
| Symptom Slope |  |  |  |  |  |  |
|  | *ns* |  |  |  |  |  |
| Treatment Arm |  |  |  |  |  |  |
|  | *ns* |  |  |  |  |  |
| Slope × Arm |  |  |  |  |  |  |
|  | *ns* |  |  |  |  |  |
| **Diagnosis-Specific Symptom Model** |
|  |  |  | Peak Coordinates |  |
|  | Anatomical Regions | k | x | y | z | z-score |
| Symptom Slope |  |  |  |  |  |  |
|  | *ns* |  |  |  |  |  |
| Treatment Arm |  |  |  |  |  |  |
|  | *ns* |  |  |  |  |  |
| Slope × Arm |  |  |  |  |  |  |
|  | R IFG (opercular and triangular parts) | 602 | 40 | 12 | 14 | 3.64 |
|  | L IFG (opercular and triangular parts) | 406 | -32 | 20 | 10 | 3.67 |
| **Reappraisal Model** |
|  |  |  | Peak Coordinates |  |
|  | Anatomical Regions | k | x | y | z | z-score |
| Reappraisal Slope(negative association) |  |  |  |  |  |  |
|  | L MFG, SFG (dorsolateral part) and IFG (triangular part) | 668 | -42 | 40 | 22 | 3.92 |
| Treatment Arm |  |  |  |  |  |  |
|  | *ns* |  |  |  |  |  |
| Slope × Arm |  |  |  |  |  |  |
|  | *ns* |  |  |  |  |  |
| *Note*. All models included patient age as a covariate of no interest. Additionally, each model included baseline symptom or reappraisal severity as a covariate of no interest, depending on the model being specified (i.e., baseline composite symptom scores were included as a covariate in the composite symptom model. L = left. R = right. IFG = inferior frontal gyrus. MFG = middle frontal gyrus. SFG = superior frontal gyrus.  |