Data supplement for Norman et al., Treatment-Specific Associations Between Brain Activation and Symptom Reduction Following Cognitive Behavioral Therapy in Obsessive-Compulsive Disorder: A Randomized fMRI Trial. Am J Psychiatry (doi: 10.1176/appi.ajp.2020.19080886)

## Supplementary Methods

### Study Design

Adolescent and adult subjects with OCD were randomized 1:1 to 12 weekly sessions of CBT or SMT to take place at the University of Michigan Health System, Michigan USA. This study was designed to detect neural correlates of CBT effect relative to an active control therapy. It was not intended to test the superiority of CBT to SMT. SMT was included to control for potential non-specific effects of time and weekly meetings with a therapist on neural correlates of symptom change.

Participants were consented and screened by a trained study coordinator, and prospective patient participants were reviewed at weekly team meetings, attended by both principal investigators, to confirm a diagnosis of obsessive-compulsive disorder. An in-house computer program was used to generate allocation sequences, in randomized blocks of 4, stratified by gender, age and medication status. While it was not possible to conceal the treatment assignment from patients and all of the study staff, steps were taken to ensure that the symptom raters were blinded to treatment assignment, such as blocking file access that would reveal treatment assignment status to blinded raters, segregating team meetings into blinded and unblinded sessions and instructing participants to not reveal clues as to their treatment

assignment when undergoing ratings. The blinded raters were also not apprised of the size of the randomization blocks. Weekly supervision by an expert clinician (Dr Himle) ensured fidelity to structured CBT and SMT manuals on a weekly basis, supplemented by reviews of session audio recordings. OCD severity was rated at sessions 1, 6 and 12 of therapy by an independent evaluator, blinded to condition, using the Yale-Brown Obsessive Compulsive Scale (Child version, as appropriate; (C)Y-BOCS) (1, 2). Patients were also assesssed using the Hamilton Anxiety Rating Scale (HAM-A), the Quick Inventory of Depressive Symptomatology (QIDS) and the Clinical Global Impression – Severity scale (CGI-S) (3–5). Therapy providers and independent evaluators were masters-level clinicians supervised by expert clinicians (Dr Himle for therapy; Dr Fitzgerald for (C)Y-BOCS). Therapy providers were trained to deliver both CBT and SMT. Both conditions were standardized using written manuals on which therapists were trained and supervised (available on request).

Importantly, the instruments used to measure OCD severity in adults and adolescents, the Y-BOCS and CY-BOCS respectively, are highly comparable. The 10 items on these instruments are identical in wording, with variation only on the question about OCD-related interference referencing school (rather than work) on the CY-BOCS. The group of investigators that developed the Y-BOCS and CY-BOCS deliberately employed identical format/wording to enable use of these instruments for the study of OCD over the lifespan (1, 2). Consistent with this intent, the Y-BOCS and CY-BOCS have similarly high validity and reliability for adults compared to adolescents, as demonstrated by internal consistency (Cronbach's alpha coefficient: Y-BOCS .89, CY-BOCS .87) and intraclass correlation coefficients across raters (Y-BOCS .98, CY-BOCS .92) (1, 2). Independent evaluators generally interviewed adolescent patients and one parent using

the CY-BOCS. Adult patients were typically interviewed alone but were encouraged to invite a partner and/or close adult friend if collateral information was deemed to be needed for accurate ratings by the patient and/or evaluator.

The current report is a planned interim analysis. An examination of the reward processing contrast was added as a primary outcome of interest based on our recent work showing altered reward network functioning in OCD (6). Hypotheses regarding cingulo-opercular regions including the rostral anterior cingulate cortex for the cognitive control and error processing contrasts were based on our recent meta-analysis (7). This report includes patients enrolled between March, 2015 and October, 2018.

**Exclusion criteria**

Exclusionary criteria included lifetime diagnoses of bipolar disorder, psychosis or mental retardation, a serious medical, neurological illness, a closed head injury, fMRI contraindications, substance abuse disorder in the past 6 months or a history of substance dependence in the past 24 months, suicidal intentions or behaviors in the previous 6 months, or having hoarding as a primary and/or only manifestation of an obsessive-compulsive related disorder. Patients were also excluded for taking antipsychotic medications, anticonvulsants, lithium or stimulants, but were permitted to take selective serotonin reuptake inhibitors if on a stable dosage for more than 4 weeks. Patients were excluded if they had experienced a failed course of ERP for OCD.

**Cognitive behavioral therapy.**

The protocol for CBT followed methods previously developed by co-author Dr. Himle and is consistent with established procedures in the field for treating OCD in adolescents (8) and adults

(9). It included 12 sessions, on a weekly basis over 3 months. In vivo and imaginal exposures were conducted, during which patients face their fears for a prolonged period of time without ritualizing. Patients were asked to stop ritualizing after the first exposure session. The rationale provided to patients was that by experiencing exposure without rituals, anxiety decreases with time alone ("habituation") and through the realization that feared consequences do not occur. Although formal cognitive therapy procedures were not used, dysfunctional cognitions were discussed within the context of exposure (e.g., asking the patient, "Did you notice that your anxiety decreased without your ritualizing and nothing bad happened?"). As homework, patients were asked to record any rituals and spend at least 1 hour per day conducting self-guided exposures.

**Stress management therapy.**

Stress management therapy followed procedures used by Lindsay et al. (10), including 12 weekly sessions over 3 months in which patients were taught stress management skills such as deep breathing, progressive muscle relaxation, positive imagery, assertiveness training, and problem solving. The rationale provided to patients was that life stressors can trigger OCD symptoms and that these stress management skills reduce stress and thereby reduce OCD symptoms. As homework, patients were asked to monitor daily stressors and practice the stress management skills for at least 1 hour each day.

In both conditions, family members were included in sessions 1, 6 and 12 to help them understand the components of treatment and to support patients in treatment participation.

**MRI Image Acquisition**

Imaging data were collected on a 3.0 T General Electric (GE) 750 scanner at the fMRI Laboratory, University of Michigan. For functional data acquisition, a T2* weighted image with gradient echo reverse spiral acquisition was acquired for each of four runs of the incentive flanker task (FOV=22cm, 43 slices, slice thickness=3.0 mm, TRs=184, TR length=2000 msec, TE=29 msec, Matrix=64x64). A T1-weighted image (FOV=22cm, 43 slices, slice thickness=3.0 mm, Matrix=256x256) was acquired in the same prescription as the functional images to facilitate co-registration. A high-resolution T1 spoiled gradient recalled echo (SPGR) scan was obtained for anatomic normalization (FOV=25.6 cm, 156 slices, slice thickness=1.0 mm, Matrix=256x256).

**fMRI preprocessing**

All functional and structural scans were inspected visually and included runs were free of excessive motion, defined as ≥ 1TR exceeding 3 mm or degrees translation or rotation. Standard preprocessing steps were performed in SPM 12 (http://www.fil.ion.ucl.ac.uk/spm; Wellcome Trust Centre for Neuroimaging, University College London, UK) and the Computational Anatomy Toolbox 12 (http://www.neuro.uni-jena.de/cat/index.html, Jena University Hospital, Jena, Germany). The raw data were slice-time corrected and rigidly realigned to the first image acquired. After realignment, the low-resolution T1 structural image was co-registered with the mean functional image. Following this, the high-resolution SPGR was co-registered to the (now co-registered to the functional image) low-resolution T1. The co-registered high-resolution SPGR was then segmented using the Computational Anatomy

Toolbox 12 and normalized to Montreal Neurological Institute (MNI) space using high dimensional Diffeomorphic Anatomical Registration through Exponentiated Lie algebra. This warp field was then applied to the functional data to bring all subjects into common stereotactic space. Normalized functional data were smoothed with a 6mm Gaussian kernel.

**First Level Model**

Task fMRI data were analyzed within the framework of the modified general linear model in SPM 12. First level regressors included high interference non-incentivised correct trials, low interference non-incentivised correct trials, high interference incentivised correct trials, low interference incentivised correct trials, high interference non-incentivised error trials, low interference non-incentivised error trials, high interference incentivised error trials, low interference incentivised correct trials. Each of these trial types was modeled from the point when patients responded during the blank screen until the end of the feedback screen. Response and feedback were not separable in our design, and a single model was used to examine the cognitive control, interference errors and reward processing contrasts. Additional regressors of non-interest included omission error trials, premature response trials, neutral cues and incentive cues. Six motion parameters, their first derivatives, and quadratic terms of original and derivatives were also included to remove signal related to spin history motion artifacts. First-level contrasts compared activation during correct high interference relative to correct low interference trials regardless of incentive level (cognitive control contrast), during high interference error trials relative to during correct high interference trials regardless of

incentive level (interference errors contrast) and during incentivised correct trials versus non-incentivised correct trials regardless of interference level (reward processing contrast).

**Statistical analysis**

Assumptions for linear mixed effect models were checked using the car and the predictmeans packages as well as custom code in R (11, 12). No clinical or brain data were outliers as defined using the criterion of 2.2 multiplied by the Interquartile Range (IQR) recommended by Hoaglin and Iglewicz (13). No datapoints had a large Cook's distance (e.g., larger than the 50th percentile of an F distribution) (14). Regardless, we repeated each linear mixed effects analysis using extracted cluster data after removing the three most influential datapoints as identified by the CookD function in the predictmeans package (11). There were no resultant qualitative changes to the findings. No clinical or brain data had a variance inflation factor meeting the criterion of >4, suggesting minimal multicollinearity (15). Levene's tests were non-significant and therefore the assumption of homoscedasticity was met (all $p>0.05$) (16). Normal QQ plots of residuals were examined, and revealed no signs of marked, problematic deviations from the normal distribution. Plotting model residuals against continuous predictor and fitted values revealed no deviations from the linearity assumption.

**Clinical and behavioral analyses**

Treatment response was compared between groups using a linear mixed effects model, performed using the following R syntax in the nlme package (17) for R (http://www.r-project.org).

(C)Y-BOCS$_{ij}$ ~ **Tx$_i$*week$_j$**+age-group$_i$+medication$_i$+Tx$_i$+week$_j$, (1 |subject)

The i subscript denotes subject, the j subscript denotes time-point, Tx denotes treatment group and models were adjusted for age group (adolescents, adults) and medication status (on versus off medication). Subscripts were added here to aid interpretation of the reader. Week was included as a continuous variable. The "1 | subject" term denotes the random intercept for subject. The interaction of interest is in **bold**.

To examine treatment response within each group, the following nlme model was used in CBT and SMT sub-group analyses where the focus was on week as a predictor of symptoms.

(C)Y-BOCS$_{ij}$ ~ **week$_j$**+age-group$_i$+medication$_i$, (1 |subject)

To examine the potential associations between task performance and symptom changes over the course of treatment, separate models were examined for each primary performance measure of interest (interference reaction time, incentive reaction time, interference errors). These took the form of the following.

(C)Y-BOCS$_{ij}$ ~ **task performance$_i$*Tx$_i$*week$_j$**+age-group$_i$+medication$_i$+task performance$_i$+Tx$_i$+week$_j$, (1 |subject)

Task performance is interference reaction time, incentive reaction time or interference errors. Interference RT was calculated by subtracting correct low interference reaction time from correct high interference reaction time for each subject. Incentive RT was calculated by subtracting reaction times for correct non-incentivised trials from reaction times to correct

incentivised trials (collapsing across high and low interference trials). The interaction of interest is in **bold**.

**fMRI analyses**

For the primary fMRI analysis, the following nlme model was used.

(C)Y-BOCS$_{ij}$ ~ **voxel-activation$_i$*Tx$_i$*week$_j$**+age-group$_i$+medication$_i$+voxel-activation$_i$+Tx$_i$+week$_j$, (1 |subject)

The i subscript denotes subject, the j subscript denotes time-point, Tx denotes treatment group (CBT, SMT) and models were adjusted for age group (adolescents/adults) and medication status (on/off medication). (C)Y-BOCS scores were treated as a repeated measures outcome, collected at the beginning, middle and end of treatment. Week was included as a continuous variable. The "1 | subject" term denotes the random intercept for subject. Analyses were repeated treating age as a linear term, with no qualitative changes to the reported findings (not shown).

To establish which groups were driving significant voxelwise **voxel-activation$_i$*Tx$_i$*week$_j$** interactions, mean BOLD data were extracted from significant clusters using MarsBar (http://marsbar.sourceforge.net/) and subjected to follow-up tests at each level of Tx (i.e. in CBT and SMT sub-groups).

(C)Y-BOCS$_{ij}$ ~ **cluster-mean$_i$*week$_j$**+age-group$_i$+medication$_i$+cluster-mean$_i$+week$_j$, (1 |subject)

In addition, plots were created using the sjplot, ggplot2 and sjmisc packages in R (18–20).

Follow-up analyses examining **voxel-activation$_i$*week$_j$** interactions were also performed within each level of Tx (i.e., CBT, SMT) and thresholded at a liberal, exploratory voxel threshold of $p<0.005$ and an FWE corrected cluster threshold of $p<0.05$, using the model below.

(C)Y-BOCS$_{ij}$ ~ **voxel-activation$_i$*week$_j$**+age-group$_i$+medication$_i$+voxel-activation$_i$+week$_j$, (1 |subject)

To assess potential relationships between baseline symptoms ((C)YBOCS at baseline) and task performance (interference RT and interference errors) on brain activation, we also ran separate linear multiple regression models in SPM 12 controlling for age group and medication status.

**Analyses of the effects of age**

We conducted follow-up analyses to check that findings were consistent across adolescent and adult age groups. First, within each age group, we examined the interaction between week and treatment group on (C)Y-BOCS scores controlling for medication status.

(C)Y-BOCS$_{ij}$ ~ **Tx$_i$*week$_j$**+medication$_i$+Tx$_i$+week$_j$, (1 |subject)

Second, we used extracted cluster data from significant clusters in the primary imaging analysis to check whether similar patterns of findings were observable in each age group.

(C)Y-BOCS$_{ij}$ ~ **cluster-mean$_i$*Tx$_i$*week$_j$**+medication$_i$+cluster-mean$_i$+Tx$_i$+week$_j$, (1 |subject)

Finally, we also checked for **cluster-mean$_i$*Tx$_i$*week$_j$*age group$_j$** interactions within these clusters using the R syntax below.

(C)Y-BOCS$_{ij}$~**cluster-mean$_i$*Tx$_i$*week$_j$*age group$_i$**+medication$_i$+cluster-mean$_i$+Tx$_i$+week$_j$, (1 |subject)

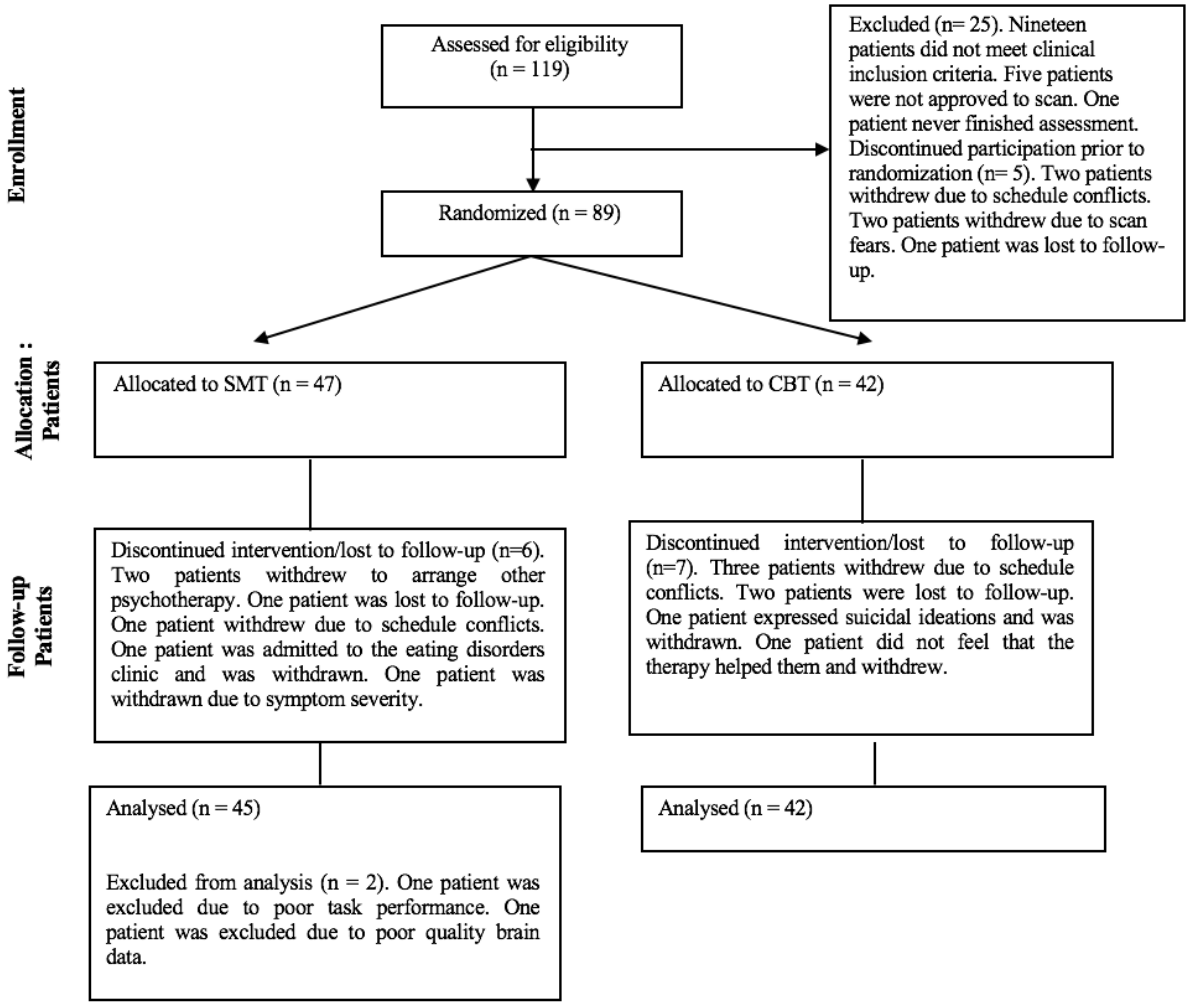


**FIGURE S1. CONSORT flow chart displaying the progress of all participants through the trial.**

Instructions:

Target letter:

K, S → Press right button

C, H → Press left button

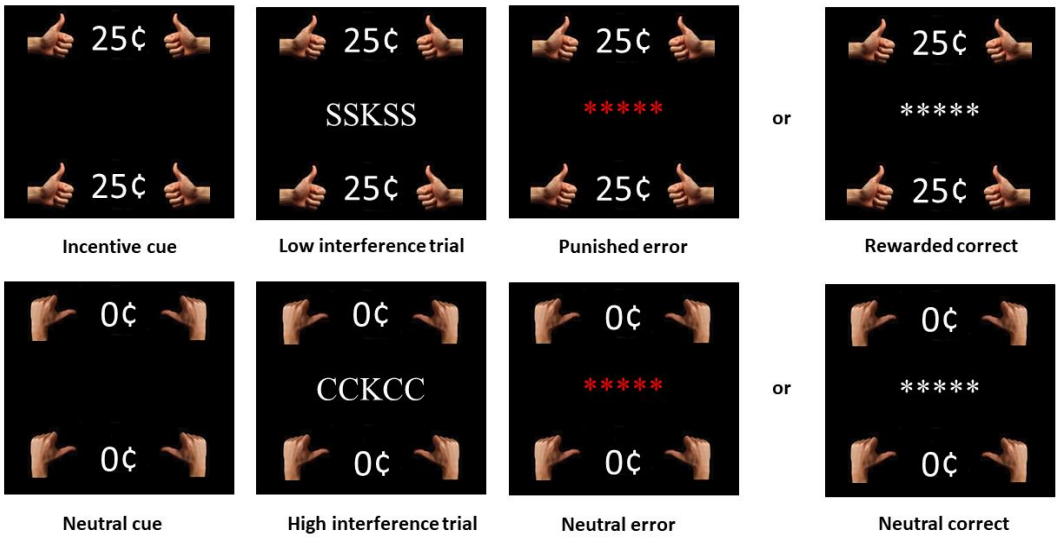


**FIGURE S2. Diagram of the Incentive Flanker Task** (21). Target and flanker stimuli were preceded by cues (1.5 – 10 s) indicating how much money patients stood to lose (for an error) or gain (for a correct response) on the upcoming trial (0¢ -- 50% of trials, 10¢, -- 25% of trials, or 25¢-- 25% of trials). During flanker trials, patients pressed one of two buttons to identify a target letter (S, K, H, and C) surrounded by four flankers which appeared for 300 ms. Two letters were associated with a left button press (e.g., "S" and "K") and two different letters with a right button press (e.g., "H" and "C"). Patients pressed the left or right button based on the identity of a target letter placed in the second, third, or fourth position in a string of five letters. The target letter was always different from the flanking letters. On low interference trials (50% of trials), both target and flankers indicated the same button press, while on high interference trials (50% of trials), target and flankers designated opposing responses. This led to a feedback signal – asterisks in place of the target/flanker stimulus as white (correct) or red (incorrect). In total, patients completed 4 runs, each consisting of 48 trials (scan duration ~25 min). Interference reaction time was calculated by subtracting correct low interference reaction time from correct high interference reaction time. Incentive reaction time was calculated by subtracting reaction times for correct non-incentivised trials from reaction times to correct incentivised trials (collapsing across high and low interference trials). Interference errors are errors from the high interference condition, collapsed across incentivised and non-incentivised trials.

**Supplementary Results**

Analyses included all N=87 randomized patients for whom there was usable baseline fMRI data. Two further patients took part in the study but were excluded from presently reported analyses due to a failure to comply with the task (N=1) or due to poor quality brain data (N=1). In the CBT group, 35 patients completed therapy and the week 12 assessment. One patient completed CBT but was removed from the study during the final assessment due to reported suicidal ideation. Forty-two patients completed therapy and the week 12 assessment in the SMT group. One patient was removed from the SMT group due to concerns regarding symptom severity at week six. Two patients from the CBT group and one from the SMT group withdrew from the study due to scheduling conflicts, while two patients in the CBT group and one patient in the SMT group were lost to follow-up. One patient in the CBT group withdrew from the study due to scheduling conflicts. Two patients in the SMT group withdrew prior to arrange other psychotherapy. One patient in the CBT group felt that treatment was not helping. The Ns for the CBT and SMT groups, as well as for the two groups by age subgroup (adolescents, adults) are given in Table S1. Details on comorbid disorders are given in Table S2.

There were no significant performance by week by treatment group interactions on (C)Y-BOCS for interfence reation time ($B$=-0.1, $t$=-1.68, $p$=0.095, 95% CI (-0.22, 0.018)), interference errors ($B$=-0.05, $t$=-1, $p$=0.32, 95% CI (-0.16, 0.05)) or incentive reaction time ($B$=-0.01, $t$=-0.47, $p$=0.64, 95% CI (-0.05, 0.03)). There were no significant voxelwise relationships between any task performance measures or baseline (C)Y-BOCS and brain activation for any of the contrasts.

**TABLE S1. Ns for the treatment groups and age groups, by week.**

| | Adolescent SMT | Adult SMT | All SMT | Adolescent CBT | Adult CBT | All CBT |
|---|---|---|---|---|---|---|
| | N | N | N | N | N | N |
| Week 1 | 20 | 25 | 45 | 19 | 23 | 42 |
| Week 6 | 20 | 21 | 41 | 16 | 21 | 37 |
| Week 12 | 19 | 21 | 40 | 16 | 20 | 36 |

**TABLE S2. Comorbid diagnoses.**

| | SMT adolescents (n = 20) | SMT adults (n =25) | CBT adolescents (n =19) | CBT adults (n =23) |
|---|---|---|---|---|
| **Anxiety Disorders, besides OCD (any)** | 11 (55%) | 13 (52%) | 10 (53%) | 7 (30%) |
| Generalized Anxiety Disorder | 6 (30%) | 11 (44%) | 6 (32%) | 4 (17%) |
| Separation Anxiety Disorder | 1 (5%) | 0 | 0 | 0 |
| Social Anxiety | 7 (35%) | 6 (24%) | 5 (26%) | 4 (17%) |
| Specific Phobia | 4 (20%) | 0 | 2 (11%) | 1 (4%) |
| Panic Disorder | 1 (5%) | 0 | 3 (16%) | 1 (4%) |
| Anxiety Disorder NOS | 0 | 1 (4%) | 1(5%) | 0 |
| | | | | |
| **Tic Disorders (any)** | 1 (5%) | 0 | 2 (11%) | 3 (13%) |
| Tics | 0 | 0 | 2 (11%) | 2 (9%) |
| Tourettes | 1 (5%) | 0 | 0 | 1 (4%) |
| | | | | |
| **Subclinical Depression (any)** | 2(5%) | 0 | 2 (11%) | 0 |
| Depression NOS | 2(5%) | 0 | 2 (11%) | 0 |
| | | | | |
| **Other Disorders (any)** | 3 (15%) | 3 (12%) | 1 (5%) | 2 (9%) |
| Impulse Control Disorder NOS | 3 (15%) | 2 (8%) | 1 (5%) | 1 |
| Trichotillomania | 0 | 2 (8%) | 0 | 1 |
| | | | | |
| **None** | 6(30%) | 10 (40%) | 7 (37%) | 12 (52%) |

*Abbreviations: CBT, cognitive behavioral therapy; NOS, not otherwise specified; SMT, stress management therapy.*

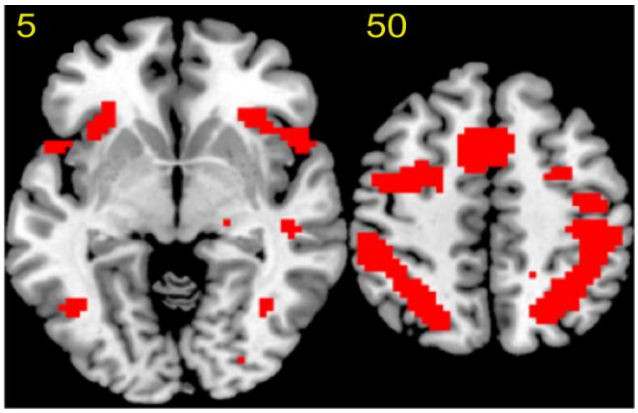


**FIGURE S3. Within-group activation map for the cognitive control contrast collapsed across all N=87 patients.** Axial slices showing within-group brain activation for the contrast comparing correct high interference trials > correct low interference trials. Data presented at a cluster forming threshold of $p<0.001$ (uncorrected) and a familywise error cluster corrected cluster threshold of $p<0.05$.

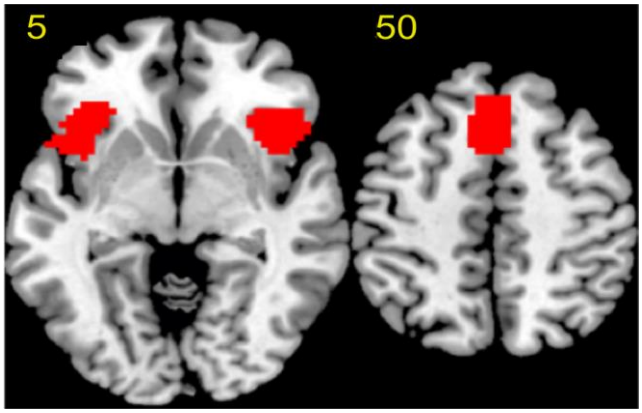


**FIGURE S4. Within-group activation map for the error processing contrast collapsed across all N=87 patients.** Axial slices showing within-group brain activation for the contrast comparing incorrect high interference trials > correct high interference trials. Data presented at a cluster forming threshold of $p<0.001$ (uncorrected) and a familywise error cluster corrected cluster threshold of $p<0.05$.

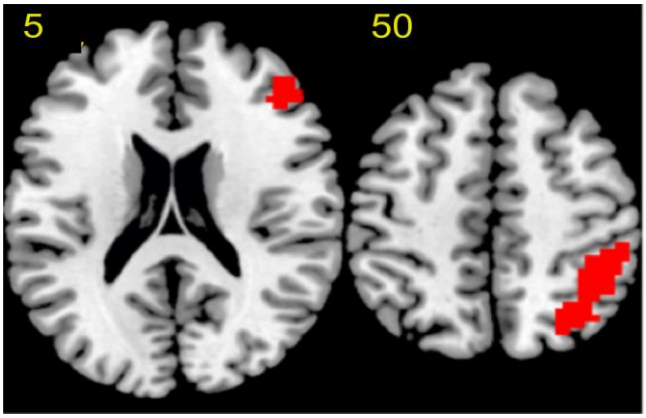


**FIGURE S5. Within-group activation map for the reward processing contrast collapsed across all N=87 patients.** Axial slices showing within-group brain activation for the contrast correct rewarded trials>correct neutral trials. Data presented at a cluster forming threshold of p<0.001 (uncorrected) and a familywise error cluster corrected cluster threshold p<0.05.

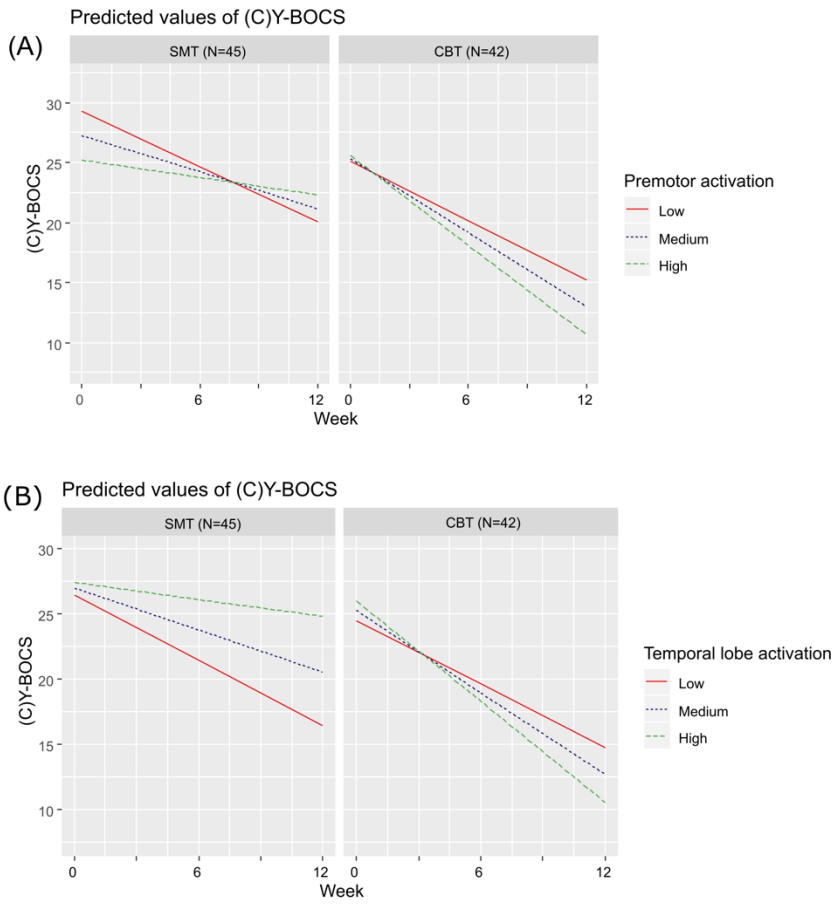


**FIGURE S6.** Graphs showing predicted model estimates for cognitive behavioral therapy (CBT) and stress management therapy (SMT) groups. The y-axis represents the predicted (C)Y-BOCS based on model estimates, and separate lines indicate level of activation within (A) premotor cortex and (B) temporal lobe clusters during cognitive control ("Low" = one standard deviation below mean, "Medium" = mean, "High" = one standard deviation above the mean).

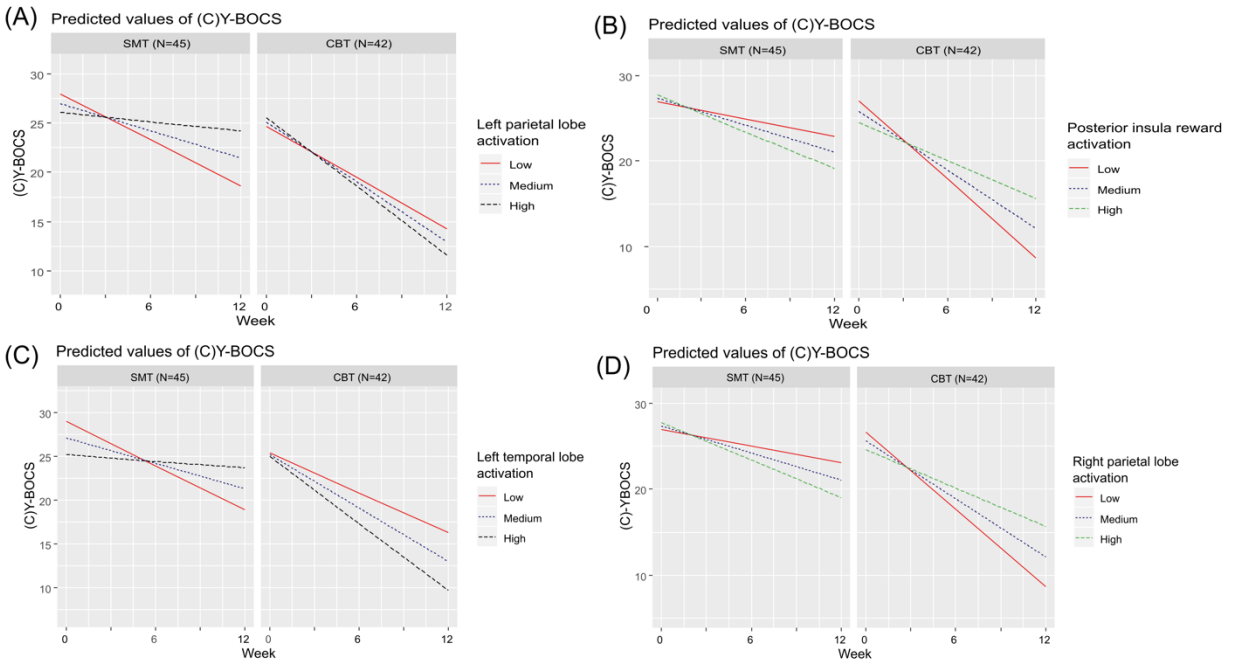


**FIGURE S7.** Graphs showing predicted model estimates for cognitive behavioral therapy (CBT) and stress management therapy (SMT) groups. The y-axis represents the predicted (C)Y-BOCS based on model estimates, and separate lines indicate level of activation within (A) left parietal lobe, (B) posterior insular, (c) left temporal lobe and (d) right parietal lobe clusters during reward processing ("Low" = one standard deviation below mean, "Medium" = mean, "High" = one standard deviation above the mean).

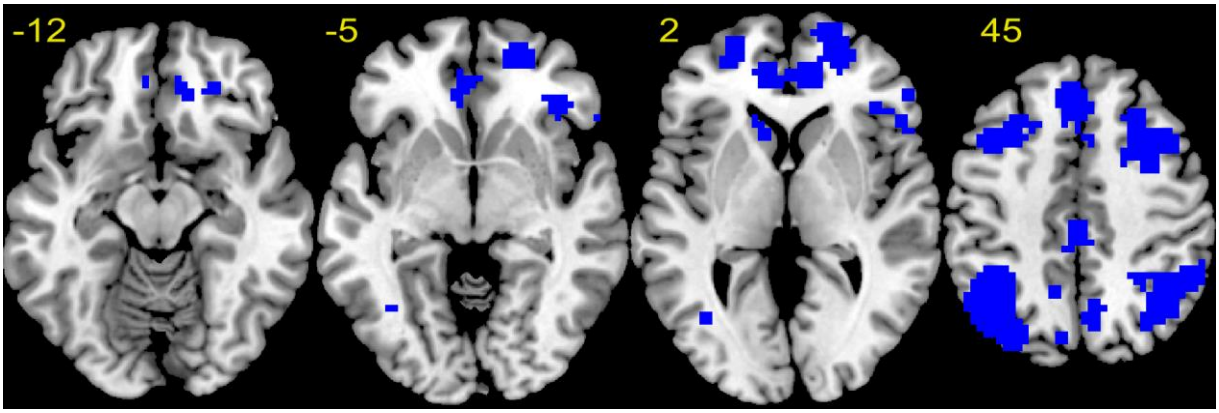


**FIGURE S8. Brain regions associated with treatment response to CBT for the cognitive control contrast.** Axial slices showing regions of activation that were associated with treatment response for the contrast comparing correct high interference trials > correct low interference trials. Data presented at an initial cluster forming threshold of p<0.005 (uncorrected) and cluster threshold p<0.05 (familywise error corrected). Blue indicates regions where more pre-treatment activation was associated with a better response to treatment. Regions include bilateral parietal lobe, precuneus and posterior cingulate cortex (MNI x,y,z=-30,-73,44, *k*=2559, Max-t=-6), bilateral inferior frontal gyrus, anterior insula, caudate, rostral anterior cingulate and dorsomedial prefrontal cortex and left dorsolateral prefrontal cortex (MNI x,y,z=,-3,32,44, *k*=1835, Max-t=-5.28) and right dorsolateral prefrontal cortex (MNI x,y,z=27 ,17,47, *k*=236, Max-t=-4.83).

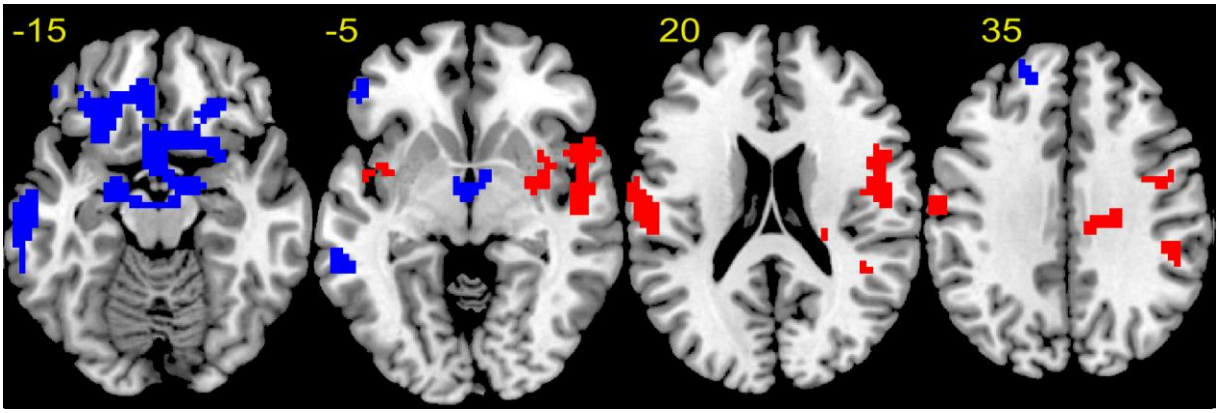


**FIGURE S9. Brain regions associated with treatment response to CBT for the reward processing contrast.** Axial slices showing regions of activation that were associated with treatment response for the contrast comparing correct incentivised trials>correct non-incentivised trials. Data presented at an initial height threshold p<.005 (uncorrected) and cluster threshold p<0.05 (familywise error corrected). Blue indicates regions where more pre-treatment activation was associated with a better response to treatment. Regions include bilateral orbitofrontal cortex, ventromedial prefrontal cortex, amygdala, thalamus and left inferior frontal gyrus (MNI x,y,z=3,-4,-10, *k*=653, Max-t=-5.55), left temporal lobe (MNI x,y,z=-63 -13 -19, *k*=190, Max-t=-4.92) and left dorsolateral prefrontal cortex (MNI x,y,z=-9,26,62, *k*=197, Max-t=-4.94). Red indicates regions where more pre-treatment activation was associated with a worse response to treatment. Regions include left postcentral gyrus, temporal lobe and posterior insula (MNI x,y,z=-63,-13,29, *k*=201, Max-t=4.46), right postcentral gyrus, temporal lobe and posterior insula (MNI x,y,z=51 ,-4,2, *k*=527, Max-t=5.27), right supplementary area (MNI x,y,z=6,2,56, *k*=141, Max-t=4.79) and right parietal lobe (MNI x,y,z=48,-34, 32, *k*=209, Max-t=5.7).

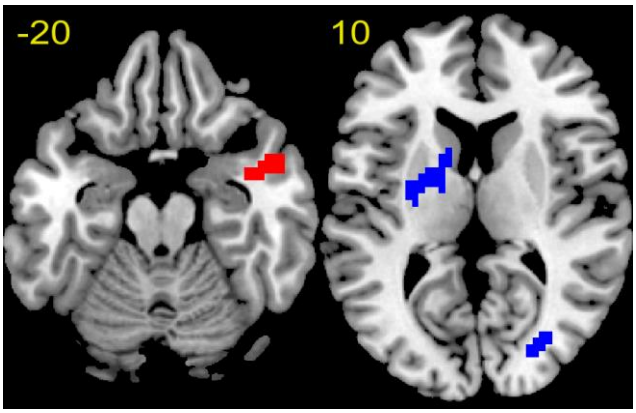


**FIGURE S10. Brain regions associated with treatment response to SMT for the cognitive control contrast.** Axial slices showing regions of activation that were associated with treatment response for the contrast comparing correct high interference trials > correct low interference trials. Data presented at an initial height threshold p<.005 (uncorrected) and cluster threshold p<0.05 (familywise error corrected). Blue indicates regions where more pre-treatment activation was associated with a better response to treatment. Regions include left putamen (MNI x,y,z=-27,-7,11, *k*=102, Max-t=-4.34), right occipital lobe (MNI x,y,z=36,-64,20, *k*=103, Max-t=4.38). Red indicates regions where more pre-treatment activation was associated with a worse response to treatment. Regions include right temporal lobe lobe (MNI x,y,z=51,8, -22 *k*=106, Max-t=5.35).

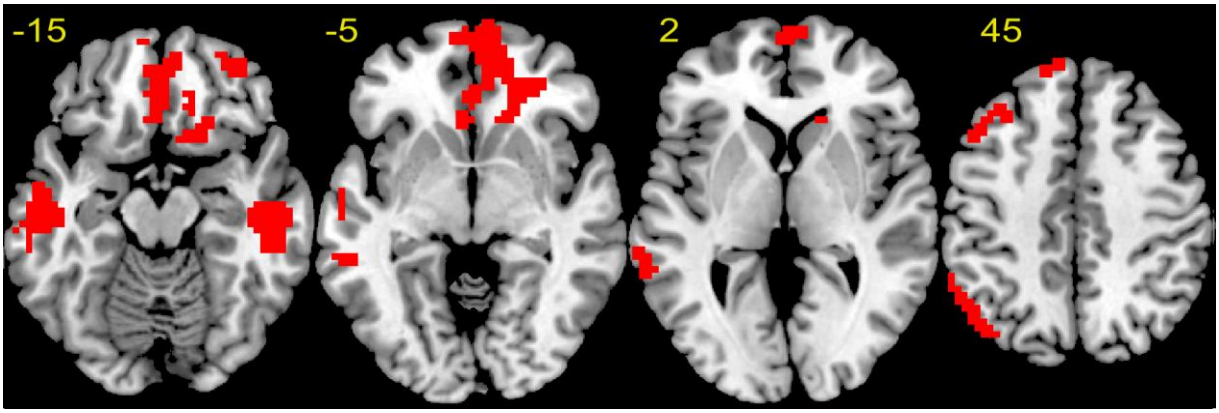


**FIGURE S11. Brain regions associated with treatment response to SMT for the reward processing contrast.** Axial slices showing regions of activation that were associated with treatment response for the contrast comparing correct incentivized trials>correct non-incentivized trials. Data presented at an initial height threshold p<.005 (uncorrected) and cluster threshold p<0.05 (familywise error corrected). Red indicates regions where more pre-treatment activation was associated with a worse response to treatment. Regions include bilateral orbitofrontal cortex, ventromedial prefrontal cortex, left inferior frontal gyrus, left dorsolateral prefrontal cortex and right caudate (MNI x,y,z=-6,65,-13, *k*=879, Max-t=5.84), left parietal lobe (MNI x,y,z=-39,-61,29, *k*=210, Max-t=3.96), left temporal lobe (MNI x,y,z=-54,-16,-13, *k*=282, Max-t=5.42) and right temporal lobe (MNI x,y,z=51,-19,-25, *k*=184, Max-t=5.34).

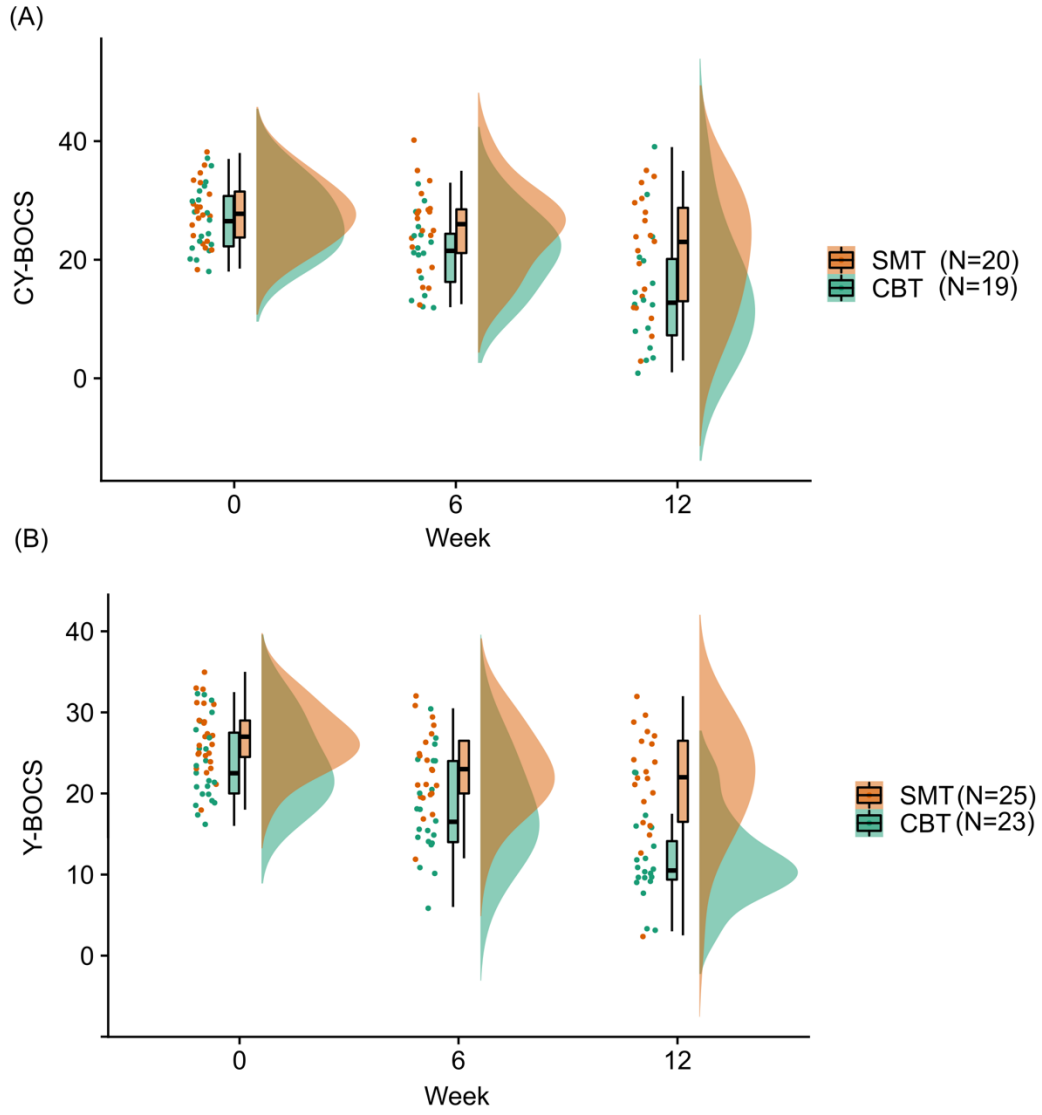


**FIGURE S12. Raincloud plot of the change in OCD symptoms over the course of CBT and SMT treatment in patients with OCD. (A) Shows symptom change in adolescents. (B) Shows symptom change in adults.** Both sub-groups showed similar symptom reductions during CBT (adolescent:*B*=-6.13,p<0.001; adult:*B*=-6.12, p<0.001) and SMT (adolescent:*B*=-3.21, p<0.001; adult:*B*=-2.74,p<0.001), as well as a group by time interaction, indicating greater efficacy of CBT relative to SMT (adolescent:*B*=-2.93,p=0.01; adult:*B*=-3.43,p< 0.001).

**TABLE S3. Linear mixed-effects models examining the relationship between activation, treatment group and week on YBOCS performed in adolescent and adult sub-groups**

| Contrast | Region | Age group | Est | SE | T | *p* |
|---|---|---|---|---|---|---|
| **Interference inhibition** | | | | | | |
| | L & R rACC | *Adolescents* | -5.35 | 2.15 | -2.49 | 0.015 |
| | | *Adults* | -6.43 | 2.12 | -3.04 | 0.003 |
| | L premotor cortex | *Adolescents* | -7.44 | 2.1 | -3.55 | 0.0007 |
| | | *Adults* | -5.82 | 2.5 | -2.33 | 0.02 |
| | R temporal lobe | *Adolescents* | -7.83 | 2.54 | -3.08 | 0.003 |
| | | *Adults* | -10.23 | 3.54 | -2.89 | 0.005 |
| **Reward processing** | | | | | | |
| | L & R vmPFC/OFC/amygdala/ IFG/DLPFC | *Adolescents* | -8.47 | 2.07 | -4.09 | 0.0001 |
| | | *Adults* | -11.38 | 1.95 | -5.79 | < 0.0001 |
| | L temporal lobe | *Adolescents* | -5.57 | 1.91 | -2.91 | 0.005 |
| | | *Adults* | -11.06 | 1.92 | -5.75 | < 0.0001 |
| | L inferior parietal lobe | *Adolescents* | -3.88 | 1.43 | -2.71 | 0.009 |
| | | *Adults* | -4.58 | 1.29 | -3.55 | 0.0007 |
| | R premotor cortex/ posterior insula | *Adolescents* | 8.37 | 1.89 | 4.41 | < 0.0001 |
| | | *Adults* | 9.13 | 3.28 | 2.78 | 0.007 |
| | R inferior parietal lobe | *Adolescents* | 11.03 | 2.73 | 4.04 | 0.0001 |
| | | *Adults* | 6.41 | 2.23 | 2.86 | 0.005 |

*Abbreviations: DLPFC, dorsolateral prefrontal cortex; Est, estimate; IFG, inferior frontal gyrus; OFC, orbitofrontal cortex; SE, standard error; OFC, orbitofrontal cortex; vmPFC, ventromedial prefrontal cortex.* Model: $(C)Y\text{-}BOCS_{ij} \sim \mathbf{cluster\text{-}mean_i * Tx_i * week_j} + medication_i + cluster\text{-}mean_{ii} + Tx_i + week_j$, (1 |subject)

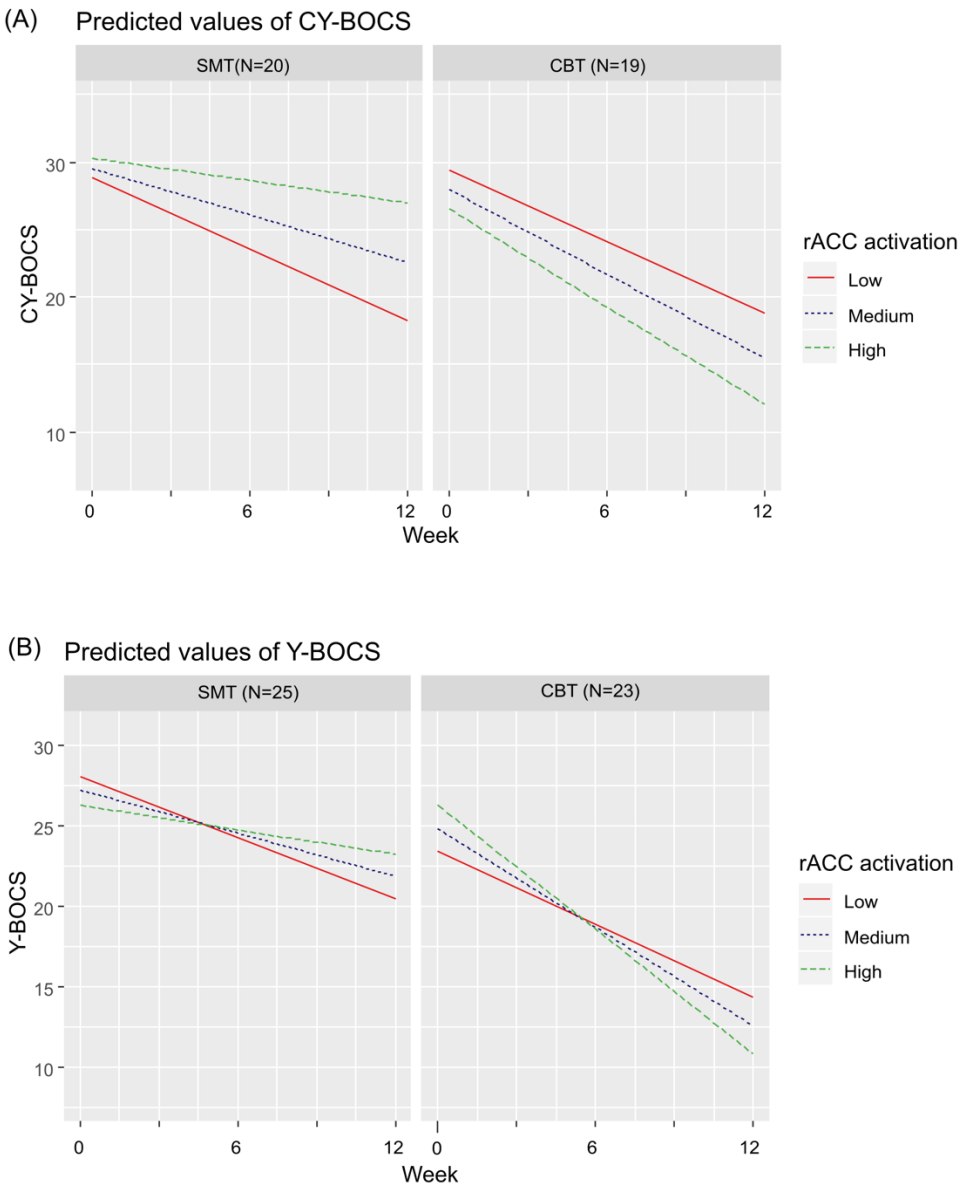


**FIGURE S13.** Graphs showing predicted model estimates for cognitive behavioral therapy (CBT) and stress management therapy (SMT) groups. The y-axis represents the predicted (C)Y-BOCS based on model estimates, and separate lines indicate level of activation within the rostral anterior cingulate cortex (rACC) ("Low" = one standard deviation below mean, "Medium" = mean, "High" = one standard deviation above the mean). (A) Shows findings in adolescents. (B) Shows findings in adults.

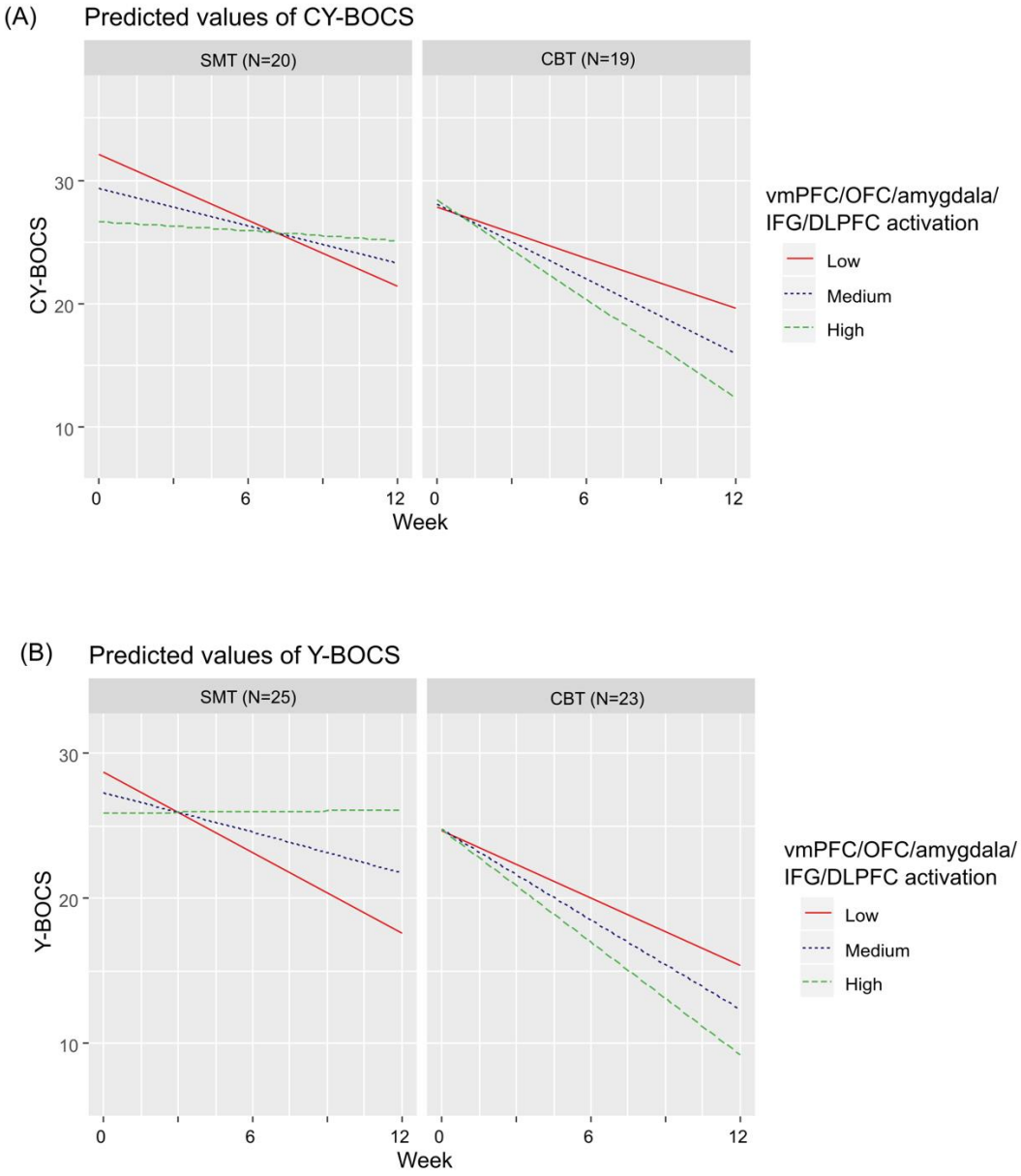


**FIGURE S14.** Graphs showing predicted model estimates for cognitive behavioral therapy (CBT) and stress management therapy (SMT) groups. The y-axis represents the predicted (C)Y-BOCS based on model estimates, and separate lines indicate level of activation within vmPFC/OFC/amygdala/IFG/DLPFC ("Low" = one standard deviation below mean, "Medium" = mean, "High" = one standard deviation above the mean). (A) Shows findings in adolescents. (B) Shows findings in adults.

## Supplementary References


1. Goodman WK, Price LH, Rasmussen SA, et al.: The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Arch Gen Psychiatry 1989; 46:1006–1011

2. Scahill L, Riddle MA, McSwiggin-Hardin M, et al.: Children's Yale-Brown Obsessive Compulsive Scale: reliability and validity. J Am Acad Child Adolesc Psychiatry 1997; 36:844–852

3. Hamilton M: The assessment of anxiety states by rating. Br J Med Psychol 1959; 32:50–55

4. Trivedi MH, Rush AJ, Ibrahim HM, et al.: The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol Med 2004; 34:73–82

5. Guy W: ECDEU assessment manual for psychopharmacology. US Department of Health, Education, and Welfare, Public Health Service …, 1976

6. Norman LJ, Carlisi CO, Christakou A, et al.: Frontostriatal Dysfunction During Decision Making in Attention-Deficit/Hyperactivity Disorder and Obsessive-Compulsive Disorder. Biol Psychiatry Cogn Neurosci Neuroimaging 2018; 3:694–703

7. Norman LJ, Taylor SF, Liu Y, et al.: Error Processing and Inhibitory Control in Obsessive-Compulsive Disorder: A Meta-analysis Using Statistical Parametric Maps. Biol Psychiatry 2019; 85:713–725

8. March JS, Mulle K: OCD in children and adolescents: A cognitive-behavioral treatment manual. Guilford Press, 1998

9. Kozak M, Foa E: Mastery of obsessive-compulsive disorder: A cognitive-behavioral approach: Therapist guide. Oxford, England, Oxford University Press, 1997

10. Lindsay M, Crino R, Andrews G: Controlled trial of exposure and response prevention in obsessive-compulsive disorder. Br J Psychiatry J Ment Sci 1997; 171:135–139

11. Luo D, Ganesh S, Koolaard J, et al.: Package 'predictmeans'2018;

12. Fox J, Weisberg S: An R Companion to Applied Regression [Internet]. Third. Thousand Oaks CA, Sage, 2019Available from: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

13. Hoaglin DC, Iglewicz B: Fine-tuning some resistant rules for outlier labeling. J Am Stat Assoc 1987; 82:1147–1149

14. Bollen KA, Jackman RW: Regression diagnostics: An expository treatment of outliers and influential cases. Sociol Methods Res 1985; 13:510–542

15. Belsley DA, Kuh K, Welsch RE: Regression diagnostics: Identifying influential data and sources of collinearity. New York, John Wiley & Sons, 1980

16. Levene H: Robust tests for equality of variances, inContributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press, 1960, pp 278–292.

17. Pinheiro J, Bates D, DebRoy S, et al.: nlme: Linear and nonlinear mixed effects models. R Package Version 2012; 3

18. Wickham H: ggplot2: Elegant Graphics for Data Analysis [Internet]. Springer-Verlag New York, 2016Available from: https://ggplot2.tidyverse.org

19. Lüdecke D: sjPlot: Data Visualization for Statistics in Social Science [Internet]. 2020Available from: https://CRAN.R-project.org/package=sjPlot

20. Lüdecke D: sjmisc: Data and Variable Transformation Functions. J Open Source Softw 2018; 3:754

21. Stern ER, Welsh RC, Fitzgerald KD, et al.: Hyperactive error responses and altered connectivity in ventromedial and frontoinsular cortices in obsessive-compulsive disorder. Biol Psychiatry 2011; 69:583–591

22. Allen M, Poggiali D, Whitaker K, et al.: Raincloud plots: a multi-platform tool for robust data visualization. Wellcome Open Res 2019; 4:63