Frontoinsular Network Markers of Current and Future Adolescent Mood Health

Supplementary Information

Supplementary Methods

Participants

Power analyses were performed in R (pwr package) to estimate the sample required to detect an association between frontoinsular functioning and baseline mood symptoms, based on effects ranging from r=0.28-0.52 (average r=0.4275) in pilot testing and in (1,2). For an 80% likelihood of detecting a significant effect of a comparable effect size, the target sample was computed at n=39.81. Therefore, we recruited a final sample of 40 adolescents (main text **Table 1**); an additional n=5 were enrolled in the study, but did not yield eligible data (one due to button box malfunction during task administration, two who elected to terminate scanning before or during task administration, and two who reported failure to understand task instructions) and therefore were excluded from the final sample. Power analyses were performed before analyses, but after recruitment was initiated for the study and after additional funding had been secured.

Participants were adolescents with varying severity of depressive symptoms, and included teens either with no history of depression or other psychiatric diagnoses (n=21) or with a primary diagnosis of major depression (n=19). The main experimental analyses followed the conceptual perspective that mood disorders are continuous phenomena (3, 4), but analyses that consider categorical diagnosis of depression are also included below. As noted in the main text and **Tables 1** and **S1**, recent/current use of stimulant medications or benzodiazepines were exclusionary, but participants taking other psychoactive medications were eligible. Medication class did not

moderate experimental effects and was not significantly related to experimental variables, and the pattern of experimental effects was maintained when omitting adolescents using lithium or anticonvulsants/antipsychotics from analyses (although effect sizes were reduced, which may stem from reduced statistical power for analyses in smaller samples). See main text for further discussion of limitations related to medication use. To evaluate inclusion/exclusion criteria, information on psychiatric history was drawn from patient records or assessed by a member of the research team using the MINI International Neuropsychiatric Interview (5). Information drawn from patient records was provided by co-authors Drs. Aguirre (Director of the McLean Hospital 3 East Program), Van der Feen (Director of the McLean Hospital Adolescent Partial Program), or Auerbach (at the time of recruitment, Director of the McLean Hospital Child and Adolescent Mood Disorders Laboratory); diagnostic information in patient records provided by these collaborators was based on clinical interview, using the MINI or other standardized diagnostic interviews.

Procedures

As noted in the main text, the study included a neurocognitive testing session, in which participants completed a clinical interview, self-report measures, and a magnetic resonance imaging (MRI) scan, followed by a two-week daily diary procedure in which they reported on daily functioning. The MRI scan included anatomical scanning, a resting-state functional scan, and a second functional scan during administration of a spatial working memory task with emotional face images (reported in the main text). After the conclusion of the research session, participants completed a two-week daily diary procedure, a member of the research staff contacted each participant via telephone at the end of the follow-up period for final assessment and debriefing. Analyses focusing on resting-state procedures are reported in (6); analyses examining magnitude of activation in response to specific working memory and emotional conditions will be reported

elsewhere. Participants were reimbursed for their time, and were debriefed and provided (upon request) with referral information for psychological services in the area.

Measures

Positive and negative affect. To evaluate current mood state, participants completed a subset of items from the Positive and Negative Affect Schedule (PANAS; (7), see (8) for prior research using the same truncated scale in similar procedures). In the PANAS, the individual rates each of a series of words according to how strongly s/he feels that way in the current moment on a scale of 1 (*very slightly or not at all*) to 5 (*extremely*). For the present study, positive affect (truncated PANAS-P, possible score range of 4 to 20) was evaluated with the items: "happy", "cheerful", "interested", and "excited"; and negative affect (truncated PANAS-N, possible score range of 5 to 25) was evaluated with the items: "sad", "nervous", "upset", "angry" and "bored". The PANAS was administered at the time of the experimental session, and electronically once per day, every day, for two weeks following the experimental session. At the baseline session, participants chose the time of day for the daily diary to be sent to their electronic device, and they had 24 hours to respond to each diary survey. The time selected to receive daily diary surveys tended to be in the afternoon (n=31 between 12pm-5pm) or evening (n=9 between 5pm-10pm). On average, participants responded to each survey within 4 hours of receipt. Time of day of survey response was not related to affective ratings (ps<0.05) and was not related to experimental variables (depression, task-related functional connectivity). Of note, this daily diary approach is distinct from ecological momentary assessment, in which the goal is to perform an evaluation at specific moments in time; future research that integrates both daily diary and EMA approaches may provide additional insight into fluctuations in mood over the day and at specific times. Subscales were scored for each date of collection, and the primary experimental measures were

maximum PANAS-N (for a measure of peak negative emotional intensity over the two-week follow-up) and standard deviation in PANAS-N (for a measure of the magnitude of negative emotional fluctuations over the two-week follow-up). The timing of maximum PANAS-N was, on average, 6.09 (SD=3.97) days after the date of scanning and was not related to experimental variables (*p*s<0.05), and the inclusion of time-to-maximum as a covariate did not influence the pattern or significance of any experimental effects. Of note, to be included in prospective analyses, participants had to complete at least half (≥7 days) of the daily diary assessments, yielding a sample of *n*=28 for those analyses. The threshold of at least seven days of assessment was selected to ensure that the daily diary estimates captured an adequate window of daily functioning to calculate lability over time. Of the *n* = 28 subjects who were compliant, average days of daily diary completed=12.25 (SD=2.19). Of the *n*=12 subjects who failed to complete at least seven days of daily diary assessments, the majority (*n*=8) completed zero assessments; an additional *n*=1 completed one; *n*=1 completed three; *n*=1 completed four; *n*=1 completed six. Changing the threshold for minimum completed surveys did not alter the pattern or significance of findings. Specifically, irrespective of whether we included subjects who completed 3+, 4+, 6+ or 7+ days of assessments, insula-FN hypoconnectivity remained related to higher maximum negative affect (all *p*s<0.05), and higher lability of negative affect (*p*s<0.05 except when including 3+ days, *p*=0.051), but remained non-significantly related to higher depression at two-week follow-up (*p*s=0.080 to 0.104); whereas insula-DN hyperconnectivity remained related to higher maximum negative affect (*p*s<0.05), higher lability of negative affect (*p*s<0.05), and higher depression at two-week follow-up (*p*s<0.05). There were no differences in demographic or clinical characteristics as a function of follow-up adherence (main text, **Tables 1** and **S1**); 52% of participants with a current diagnosis of major depression were adherent to the daily diary,

compared with 85% of participants with no psychiatric diagnosis, but there was not a significant difference in compliance as a function of diagnosis (Fisher's exact = 0.46, $p>0.05$). Therefore, although it is unknown what contributed to attrition, there did not appear to be systematic differences between subjects who completed daily diaries versus those who did not.

Emotional working memory task. The emotion regulation task was the Emotion Face Sorting (EFS) task (main text **Figure 1**) presented using EPrime 2.6 (Psychology Software Tools, Pittsburgh, PA). Participants were instructed that responses should be given as quickly and accurately as possible. For each trial, the participant was first presented with a set of either two fearful faces ("negative" condition) or two happy faces ("positive" condition). Next, the participant was presented with a cue to either maintain the same spatial position of the faces ("stay" condition), or to mentally reverse the spatial position of the faces (moving the left-side image to the right, and the right-side image to the left; "switch" condition). Finally, the participant was presented with a set of face images (both original faces sorted correctly, both original faces sorted incorrectly, or one original face paired with a new face image of the same emotional valence) and responded to indicate whether the content and spatial organization of the images matched the images held in working memory. All faces presented within an individual trial were of the same emotional valence. Face stimuli consisted of negative or positive emotional face images from 12 individuals (6 female) from the NimStim set available at http://www.macbrain.org (9). Prior to scanning, participants completed 10 practice trials. During functional scanning, participants completed four blocks of 20 trials (9.36 seconds/trial), separated by four fixation blocks (26 seconds/block). Individual trials were separated by a jittered inter-trial interval (0.72-8.00 seconds following an exponential function with progressively higher representation of shorter intervals). Thus, this task design permits either blocked or trial-by-trial analyses. Because the focus of this study was on

large-scale network response to emotion regulation, overall, we adopted a blocked functional connectivity analytic approach (see main text).

Analyses

Behavioral analyses. The outcome variables of interest from the EFS task were reaction time (RT) and accuracy. For RT calculations, incorrect trials and trials on which RTs were less than 200ms or more than 3 standard deviations above the within-subject mean were excluded (consistent with (10-12)). RTs were natural log transformed to reduce the skew common to RT and which violates the statistical assumption of normal distribution. Reaction time was calculated as an average, and accuracy was calculated as the proportion correct, for each trial type. All RT and accuracy distributions met normality requirements (**Table S2**).

Behavioral analyses were conducted using SPSS (IBM; New York, NY), and relied on mixed-effects analysis of variance ((M)ANOVA), in which within-subject variables included working memory load (switch vs. stay), image type (negative vs. positive valence), and their interaction. Between-subject variables added to the (M)ANOVA were individual differences in task-related functional connectivity (a) between insula and FN, (b) between insula and DN, (c) within the FN, and (d) within the DN, (see below). As noted in the main text, age and gender were included as group-level covariates in all analyses.

General image preprocessing. We preprocessed functional data in SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) using the standard steps of slice-time correction, realignment, segmentation, normalization in Montreal Neurological Institute (MNI) space, and smoothing with a 6-mm kernel.

Head motion and artifact detection. We used SPM12 to evaluate head motion by translation and rotation in *x, y, z* directions, and Artifact Detection Tools (ART,

www.nitrc.org/projects/artifact_detect/) to calculate time points of significant head motion or fluctuations in the magnetic field (>0.5 mm motion from previous frame, global mean intensity >3 standard deviations from mean intensity across functional scans) for each participant. These procedures provided estimates of volume-by-volume displacement and outlier volumes for each participant. Then, motion parameters and outlier images were modeled in each participant's first-level general linear model (with outliers modeled as a vector the length of the timeseries, with 1 for outlier time points and 0 for non-outlier time points, to censor the influence of outlier time points on estimates of functional connectivity while maintaining the temporal structure of the data). Therefore, together, motion correction included the regressing out of both residual head motion parameters (three translation and three rotation parameters, plus one composite motion parameter reflecting the maximum scan-to-scan movement), and outlier volumes. Of note, motion and outlier volume frequency were not significantly associated with depressive symptoms at baseline, r=0.03, p=0.854, or at follow-up, r=-0.01, p=0.954, and were not significantly associated with measures of negative mood intensity, r=-0.07, p=0.651, or lability, r=-0.14, p=0.436, over the follow-up period.

Denoising. Denoising the timeseries to correct for physiological noise and average activation associated with task conditions was performed in the CONN toolbox (https://www.nitrc.org/projects/conn/; (13)). The CONN toolbox used CompCor (14) to estimate physiological noise from white matter and cerebrospinal fluid for each subject using principal component analysis. The first five components were then regressed out of each subject's functional data on the first level of analysis, as were the main task effects (average activation associated with each task condition), after which a band-pass filter of 0.008–0.09 Hz was applied to the timeseries. In sum, the denoising corrections performed on the timeseries included: detrending, outlier

correction, motion regression, regressing out the average effects of task, and CompCor correction (performed together in a single first-level regression model), followed by band-pass filtering. These corrections produced a residual BOLD time course at each voxel that was used for subsequent analyses.

Neuroimaging first-level analysis. For regions of interest (ROIs) in frontoparietal network (FN), default network (DN), and insula, see (**Figure S1**). Regions of interest were defined on the basis of an *a priori* resting-state functional network parcellation (15) and included 4 insula ROIs, 26 ROIs in the FN, and 24 ROIs in the DN (**Figure S1**). For the purposes of the present study, ROI-to-ROI associations were calculated and compared between task and rest for: (a) insula and FN ROIs [insula-FN], (b) insula and DN ROIs [insula-DN], (c) FN ROIs [within FN], and (d) DN ROIs [within DN]. These *a priori* networks can be further decomposed into subnetworks (15) representing functional circuits that are especially highly coordinated and have been differentially implicated in psychiatric illness (e.g., (16)). Follow-up analyses examined ROI-to-ROI associations within subnetworks: (a) insula with FN subnetworks [insula-FN_A, insula-FN_B, insula-FN_C], (b) insula with DN subnetworks [insula-DN_A, insula-DN_B, insula-DN_C], (c) FN subnetworks [within FN_A, within FN_B, within FN_C], and (d) DN subnetworks [within DN_A, within DN_B, within DN_C]. See below for discussion of post-hoc subnetwork analyses.

Neuroimaging group-level analyses. The main group-level analyses were focused on associations between brain network functioning, current depressive symptoms, future depressive symptoms, or future negative affect. Conceiving measures of brain network functioning as the predictor variables, and measures of mood health as the outcome variables, is consistent with cognitive neuroscience approaches and the idea that biomarkers may be used to predict emotional wellness (17).

Multiple regression was selected for hypothesis testing because this approach allows for examination of the predictive effects of functional connectivity in one network, over and above task response in other networks. For example, here, subjects characterized by higher task-related functional connectivity in insula-FN also showed higher task-related functional connectivity within FN, r=0.64, p<0.001, (although not within DN, r=0.18, p=0.248). The finding that task responsiveness across networks was positively correlated (but not isomorphic, i.e., rs well below 1) is not surprising, but highlights the importance of including network measures together in a single model to examine specificity of effects.

All group-level prospective analyses controlled for baseline mood measures (along with age and gender covariates). As noted in the main text, this approach is important in order to disentangle predictive biomarker effects from tertiary effects explained by within-subject stability in symptom severity over time (i.e., people who report high symptom severity at baseline are also likely to report high symptom severity in the future). However, the omission of baseline covariates did not alter the pattern or significance of prospective effects (most effects becoming stronger; average change in η^2_p =0.067).

Post-hoc subnetwork analyses. To gain a more precise understanding of the frontoinsular regions that contributed to significant experimental effects, we performed a series of post-hoc partial correlation analyses (controlling for demographic covariates and, in prospective analyses, for baseline mood measures). Specifically, following significant regression analyses indicating a relationship between task-related functional connectivity in a given network and mood measures, we performed partial correlations to test the associations between frontoinsular subnetworks and the mood measure of interest.

Diagnostic status. The analytic focus on depressive symptoms, rather than diagnoses, was motivated by the perspective that depression may be better conceived as a continuum than a categorical phenomenon (3, 4). Participants were recruited with the goal of including non-depressed and depressed adolescents reporting a wide range of depressive symptom severity (measured with the Center for Epidemiological Studies-Depression Scale, CESD). Unsurprisingly, teens who met criteria for current major depressive disorder (MDD) at baseline reported higher severity of current depressive symptoms than teens who did not meet criteria for MDD (CESD for participants with MDD M=33.89, SD=8.49; CESD for participants without MDD M=4.43, SD=4.29). However, there was insufficient evidence for bimodality when considering the full distribution of CESD scores: mean CESD scores for participants with versus without MDD were within two pooled standard deviations of one another, SD_p=16.27, and Sarle's bimodality coefficient <0.6 (a calculation of biomodality based on skewness, kurtosis, and sample size of a distribution; values >0.6 are interpreted as evidence for bimodality (18)). Together, this supports the present dimensional models. However, for completeness, the baseline regression analysis in which task-related functional connectivity was used to predict baseline depressive symptoms was repeated as a binary logistic regression to predict baseline diagnostic status (MDD versus non-MDD). In this logistic regression, the predictor variables were task-related functional connectivity between insula-FN, between insula-DN, within FN, and within DN, regressed (together with demographic covariates) on log-odds of meeting criteria for a diagnosis of MDD.

Future negative affect. To supplement the main experimental analyses related to negative affect, in order to test the specificity of negative affect effects, we repeated all regression analyses addressing future negative affect when controlling for future positive affect. In the first analysis, brain connectivity variables were regressed together with demographic covariates, baseline

PANAS-N, baseline PANAS-P, and future PANAS-P maximum score, on the dependent variable: maximum PANAS-N. In the second analysis, the brain connectivity variables were regressed together with demographic covariates, baseline PANAS-N, baseline PANAS-P, and future standard deviation in PANAS-P over follow-up, on the dependent variable: standard deviation in PANAS-N over follow-up.

Supplementary Results

Task performance across the group

Accuracy was computed for each task condition, for each participant, to confirm above-chance (accuracy ≥ 55%) performance; a total of n=40 performed adequately on the task. Next, to evaluate emotional working memory ability across the group, (M)ANOVAs were performed on RT or accuracy, with working memory load (switch vs. stay), image type (negative vs. positive valence), and their interaction entered as within-subject variables (and demographic covariates on the group level). On average across participants, there was a robust effect of working memory load on RT, $F(1,37)=25.57$, $p<0.001$, $\eta^2_p=0.41$, in which participants were slower to respond on "switch" compared with "stay" trials (**Figure S2**); however, there was no significant effect of image valence or interaction between working memory load and valence, ps>0.10. There was no significant main effect of working memory load on accuracy, $p>0.10$, but there was an effect of valence, $F(1,37)=6.19$, $p=0.018$, $\eta^2_p=0.14$, and an interaction between load and valence, $F(1,37)=4.52$, $p=0.040$, $\eta^2_p=0.11$, (**Figure S2**). Participants were more accurate for negative than positive trials, and this difference emerged specifically for the "stay" condition, indicating that valence effects on accuracy emerge more strongly at lower working memory load.

Association between functional network response to task and task performance

To test associations between network functioning and task performance, the above (M)ANOVAs on RT and accuracy were repeated, adding task-related functional connectivity between insula-FN, between insula-DN, within FN, and within DN, as between-subjects variables (together with demographic covariates).

The first (M)ANOVA showed that task-related connectivity between insula-FN, $F(1,31)=7.85$, $p=0.009$, $\eta^2_p=0.20$, and within FN, $F(1,31)=6.08$, $p=0.019$, $\eta^2_p=0.16$, moderated the effects of working memory load on response speed (but did not moderate valence effects, or valence-by-load effects, ps>0.10). Follow-up partial correlation analyses were performed to clarify these effects, and showed that teens who exhibited higher insula-FN functional connectivity in response to the task were especially faster to respond to "stay" trials, $r(31)=-0.35$, $p=0.049$, but showed no significant difference in response speed to "switch" trials, $r(31)=-0.08$, $p=0.653$, (**Figure S3**). In contrast, teens who exhibited higher within-FN functional connectivity in response to the task were especially faster to respond to "switch" trials, $r(31)=-0.41$, $p=0.019$, but showed no significant difference in response speed to "stay" trials, $r(31)=-0.22$, $p=0.212$. Together, these results suggest that insula-FN functional connectivity may be especially important for goal-directed attention during low working memory load trials (perhaps when it is more difficult to keep attention on the task and resist mind-wandering), whereas functional connectivity within FN is especially important for goal-directed attention during high working memory load trials (when executive control must be exerted to accomplish the working memory manipulation). Finally, there was a trend for a main effect of higher insula-DN functional connectivity on slower RT, overall, $F(1,31)=3.80$, $p=0.058$, $\eta^2_p=0.11$. Thus, decreases in task-related insula-DN functional connectivity may reflect the ability to regulate attention away from internal thoughts, regardless

of task demands. In the second (M)ANOVA, there were no significant effects of task-related network functioning on task accuracy, *p*s>0.10.

Association between mood measures and task performance

The focus of this study was on frontoinsular response to emotional working memory as a biomarker of prospective mood. However, it is also possible that behavioral measures of emotional working memory are associated with mood. To explore this possibility, we performed a (M)ANOVA on RT, with working memory load (switch vs. stay), image type (negative vs. positive valence), and their interaction entered as within-subject variables; and depression severity (z-transformed CESD scores) at baseline, together with demographic covariates, on the group level. Results showed a trending effect in which higher baseline depression moderated the effects of working memory load on response speed, $F(1,36)=3.48$, $p=0.070$, $\eta^2_p=0.09$, driven by relatively slower RT to "stay" than to "switch" trials. In light of these results, we next performed a set of partial correlations (controlling for age and gender) to examine the relationships between RT to "switch" or "stay" trials and prospective measures of mood. Results showed that people who responded more slowly to "stay" trials reported higher future intensity and lability of negative affect, and higher future depression severity (all *p*s < 0.05); however, response speed to "switch" trials was not significantly related to prospective mood measures (all *p*s > 0.05). Together, these supplementary analyses suggest that impairments in goal-directed attention during low-working memory-load trials may be reflected in weaker insula-FN task response, and may constitute important area of cognitive deficit in depression.

Post-hoc subnetwork analyses: *baseline depressive symptoms*

As reported in the main text, the first hypothesis-testing regression revealed that teens exhibiting lower task-related functional connectivity between insula-FN reported higher severity

of depression at baseline. Post-hoc partial correlation analyses aimed at localizing these effects did not indicate specific FN subnetworks in which hypoconnectivity with insula was particularly related to depression, although trending effects were observed when considering insula hypoconnectivity with dorsolateral and ventrolateral prefrontal systems (insula-FN_A, $r(31)=-0.33$, $p=0.060$, insula-FN_B, $r(31)=-0.30$, $p=0.094$).

Post-hoc subnetwork analyses: *prospective negative affect*

In the second set of experimental regressions, insula-FN hypoconnectivity and insula-DN hyperconnectivity were each associated with both higher intensity and higher lability of future negative affect (see main text). Post-hoc partial correlation analyses showed that, in particular, hypoconnectivity between insula and frontoparietal subnetworks including dorsolateral prefrontal systems (subnetwork FN_A) was associated with increased future negative affect intensity (insula-FN_A, $r(17)=-0.51$, $p=0.026$) and lability (insula-FN_A, $r(17)=-0.45$, $p=0.049$). Meanwhile, hyperconnectivity between insula and default subnetworks that include dorsomedial prefrontal systems (subnetwork DN_B) was associated with increased future negative affect intensity (insula-DN_B, $r(17)=0.56$, $p=0.012$) and lability (insula-DN_B, $r(17)=0.53$, $p=0.020$).

Post-hoc subnetwork analyses: *prospective depressive symptoms*

As reported in the main text, in the final regression, participants exhibiting insula-DN hyperconnectivity in response to the task at baseline reported more severe depression at follow-up. Follow-up partial correlation analyses designed to pinpoint these effects showed that hyperconnectivity between insula and ventromedial prefrontal or temporal regions of DN (subnetworks DN_A, DN_C) was predictive of future depression (insula-DN_A, $r(19)=0.44$, $p=0.045$, insula-DN_C, $r(19)=0.60$, $p=0.004$).

Association between functional network response to task and baseline diagnosis of MDD

A binary logistic regression was performed in which task-related functional connectivity between insula-FN, between insula-DN, within FN, and within DN, and demographic covariates, were together regressed on log-odds of having a diagnosis of MDD at baseline. Results showed a trend for teens with lower task-related functional connectivity between insula-FN to be more likely to report a diagnosis of depression at baseline, Wald χ^2 (1,31) = 3.00, p=0.083. Task-related functional connectivity in other networks (between insula-DN, within FN, within DN) was not associated with likelihood of a depression diagnosis at baseline, ps>0.10. These results considering diagnostic status are consistent with results of analyses that considered symptom severity, albeit yielding a weaker (non-significant but trending) association, which may stem from enhanced statistical power of a continuous measure of depression (19).

Association between functional network response to task and future negative affect, controlling for future positive affect

Regression analyses considering the associations between task-related functional connectivity and future negative affect were repeated controlling for future positive affect, to investigate the specificity of valence effects. Consistent with the original analyses, both insula-FN hypoconnectivity, F(1,16)=7.00, p=0.018, η^2_p=0.30, and insula-DN hyperconnectivity, F(1,16)=5.92, p=0.027, η^2_p=0.27, were associated with higher intensity of future negative affect. Also as in the original analyses, insula-FN hypoconnectivity was associated with higher future negative mood lability, F(1,16)=7.21, p=0.016, η^2_p=0.31. However, the association between insula-DN hyperconnectivity and negative mood lability dropped out of significance, F(1,16)=2.60, p=0.124, η^2_p=0.14. These supplementary results suggest that, overall, frontoinsular

task response was specifically related to future negative affect, and this association could not be explained by intensity or lability of affect in general.

Supplementary Tables

Table S1. Sample demographics and clinical characteristics by diagnostic group.

	Full Sample (*n*=40)			
	MDD (*n*=19)		No Psychiatric Diagnosis (*n*=21)	
	Mean (SD)		Mean (SD)	
Age (years)	16.26 (1.49)		15.86 (1.65)	
Depressive Symptoms (CESD Score) At Baseline	33.89 (8.49)		4.43 (4.29)	
	Follow-up (*n*=10)		Follow-up (*n*=18)	
Depressive Symptoms (CESD Score) At Follow-up (Day 15)	22.90 (10.01)		4.67 (4.41)	
Max Negative Affect (short PANAS-N) Over Follow-up	15.8 (3.12)		10.28 (2.45)	
SD in Negative Affect (short PANAS-N) Over Follow-up	2.99 (0.88)		1.58 (0.64)	
Max Positive Affect (short PANAS-P) Over Follow-up	13.10 (4.89)		14.94 (2.44)	
SD in Positive Affect (short PANAS-P) Over Follow-up	2.74 (1.24)		2.25 (0.93)	
	%		%	
Gender				
Female	63.2%		81.0%	
Male	26.3%		19.0%	
Non-binary	10.5%		0.0%	
Medication Use				
Norepinephrine-dopamine reuptake inhibitor	26.3%		0.0%	
Selective serotonin-norepinephrine reuptake inhibitor	15.8%		4.8%	
Selective serotonin reuptake inhibitor	73.7%		0.0%	
Tetracyclics	0.0%		0.0%	
Anticonvulsants/Antipsychotics	31.6%		0.0%	
Lithium	15.8%		0.0%	
Anxiolytics (non-Benzodiazepine)	15.8%		0.0%	
Any psychoactive medication use	78.9%		4.8%	
Race				
African American	5.3%		0.0%	
American Indian/Alaskan Native	0.0%		0.0%	
Asian	5.3%		14.3%	
Biracial or Other	26.3%		0.0%	
White	63.2%		85.7%	
Ethnicity				
Hispanic	0.0%		4.8%	
Not Hispanic or Other	100.0%		95.2%	
Education (Parent Highest)				
Without High School Diploma	0.0%		0.0%	
High School Graduate Without College Degree	0.0%		4.8%	
Some College Education	5.3%		4.8%	
Degree from Four-Year College (or more)	94.7%		90.5%	
	Current %	Life %	Current %	Life %
Major Depressive Disorder (MDD)	100.0%	100.0%	0.0%	0.0%
Anxiety Disorders secondary to MDD	31.6%	31.6%	0.0%	0.0%
Posttraumatic Stress Disorder	0.0%	0.0%	0.0%	0.0%
Generalized Anxiety Disorder	31.6%	31.6%	0.0%	0.0%
Panic Disorder	10.5%	10.5%	0.0%	0.0%
Agoraphobia	0.0%	0.0%	0.0%	0.0%
Social Phobia	5.3%	5.3%	0.0%	0.0%
Specific Phobia	0.0%	0.0%	0.0%	0.0%
(Mild) Substance Use Disorders	0.0%	0.0%	0.0%	0.0%

Note: Center for Epidemiological Studies Depression Scale (CESD), Maximum (Max), Positive and Negative Affect Schedule - Negative affect subscale, short version (PANAS-N, possible score range of 5 to 25), Positive and Negative Affect Schedule - Positive affect subscale, short version (PANAS-P, possible score range of 4 to 20), standard deviation (SD).

Table S2. Emotional working memory task: Descriptive statistics

	mean	SD	skew	kurtosis
Reaction Time	(ln ms)			
Positive Stay	7.132	0.182	-0.125	-0.673
Positive Switch	7.192	0.190	-0.357	0.103
Negative Stay	7.104	0.186	-0.341	0.346
Negative Switch	7.196	0.186	-0.139	0.676
Accuracy	(proportion)			
Positive Stay	0.843	0.096	-0.608	0.130
Positive Switch	0.870	0.100	-0.501	-0.619
Negative Stay	0.898	0.093	-0.675	-0.602
Negative Switch	0.874	0.096	-0.475	-0.800

Note: ln ms = natural log transformed millisecond reaction time, proportion = proportion accurate trials in each condition, SD = standard deviation.

Supplementary Figures

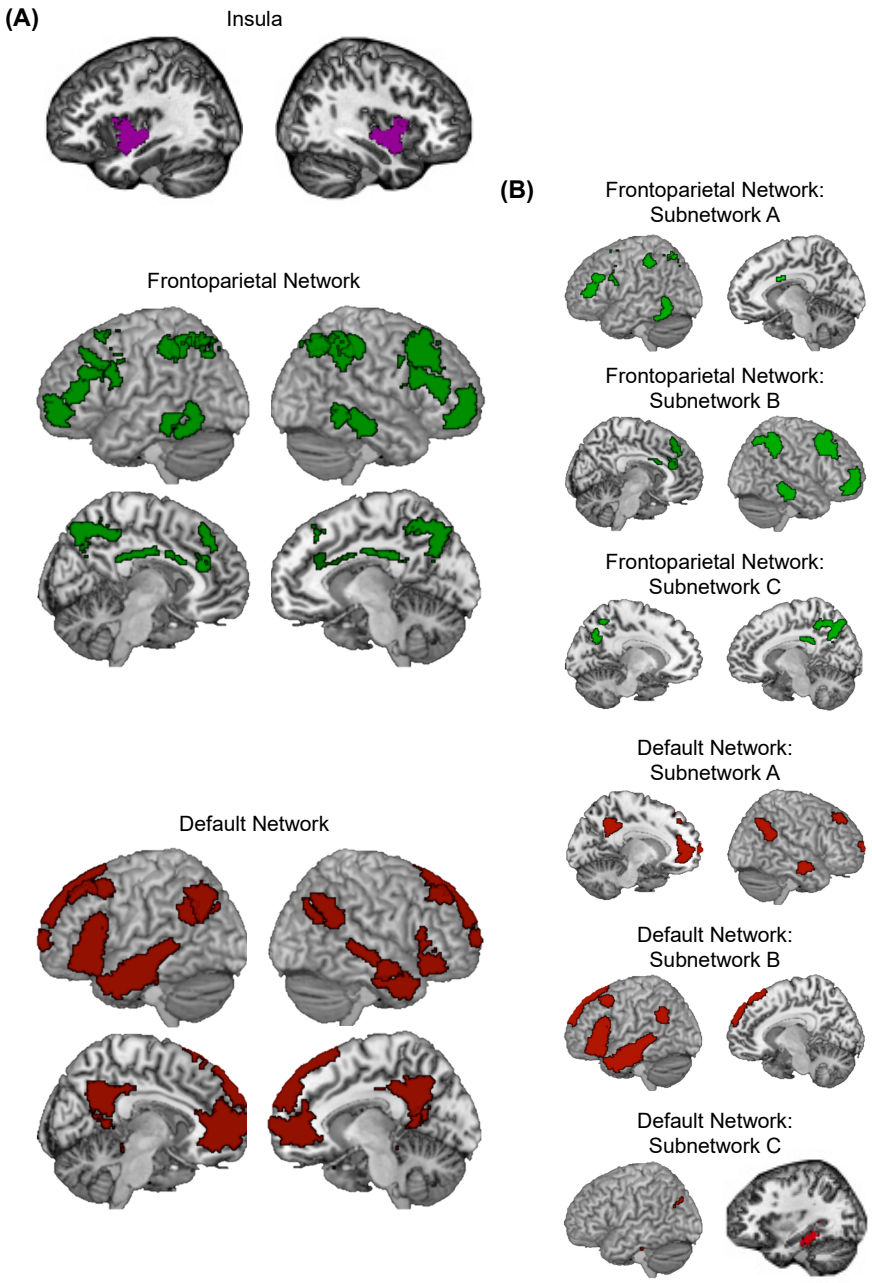

Figure S1. Functional networks. Regions of interest were defined on the basis of an *a priori* resting-state functional network parcellation (see (1,2) for source of parcellation and additional views) and included **(A)** 4 insula ROIs, 26 ROIs comprising the frontoparietal network (FN), and 24 ROIs comprising the default network (DN). **(B)** The *a priori* FN and DN can be decomposed into subnetworks (15) representing functional circuits that are especially highly coordinated in their activity, and have been differentially implicated in psychiatric illness (e.g.,(16)).

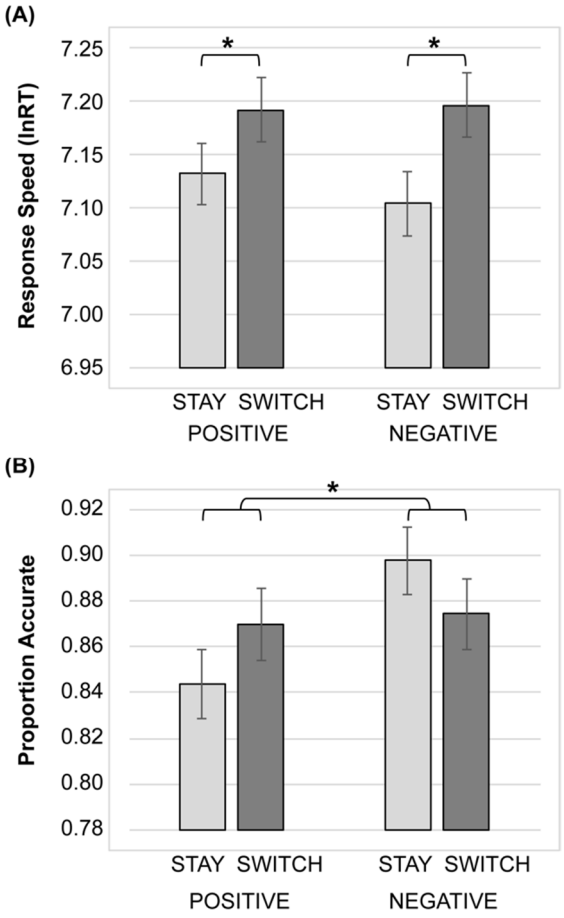

Figure S2. Emotional working memory task performance. Displayed are measures of average task performance, across the group (n=40), for the emotional working memory task. **(A)** Natural-log transformed reaction time (lnRT) to each trial type, in response to happy (Positive) or fearful (Negative) face stimuli that were either spatially manipulated in working memory (Switch) or were maintained in their original spatial position (Stay). **(B)** Proportion of accurate responses to each trial type, in response to happy (Positive) or fearful (Negative) face stimuli that were either spatially manipulated in working memory (Switch) or were maintained in their original spatial position (Stay). *Note:* Significant task effects, *p<0.05.

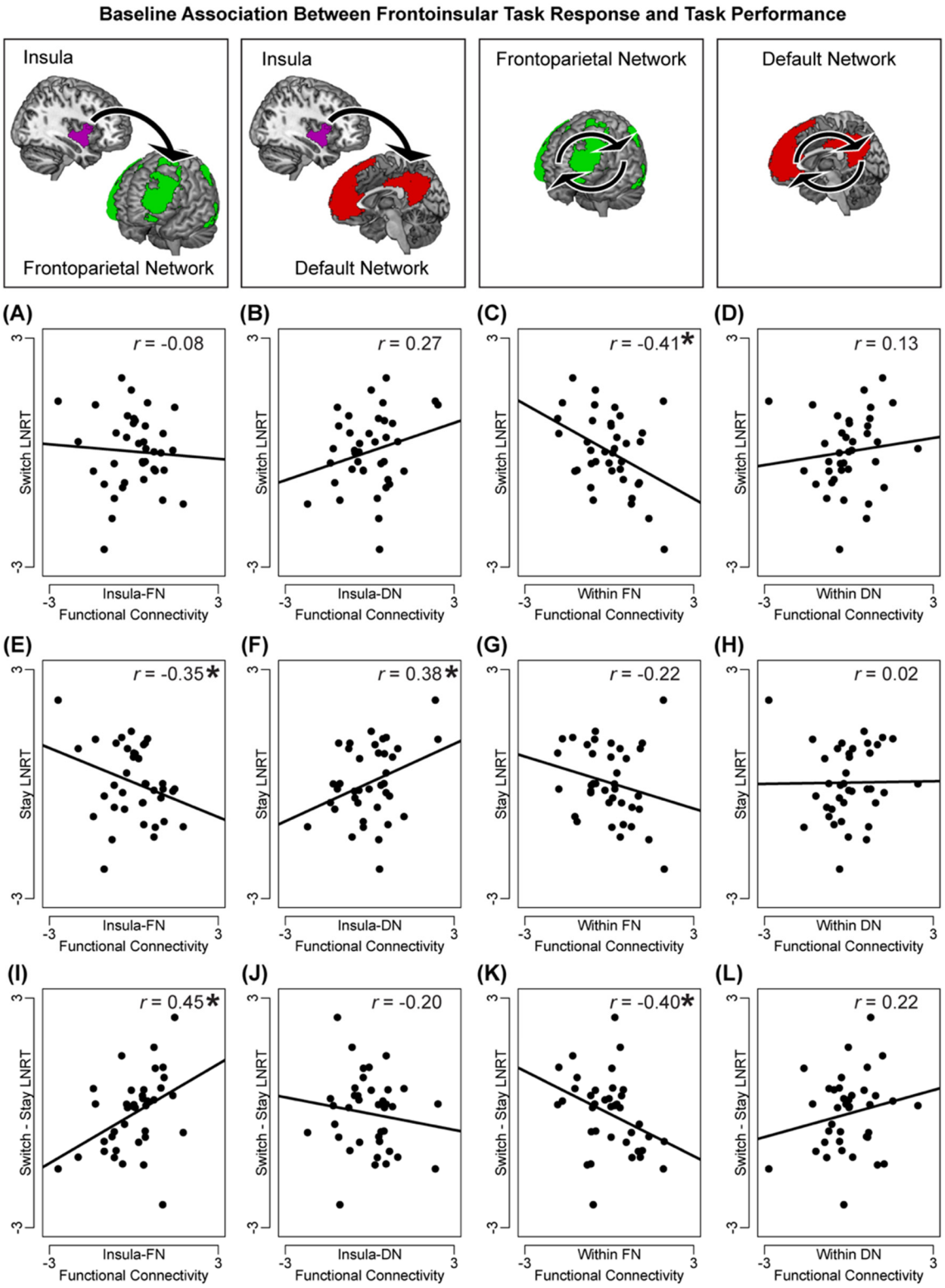

Figure S3. Functional network response to task is associated with task performance. To test associations between functional network responses to task demands and task performance, a single (M)ANOVA was performed with working memory load (switch vs. stay), image type

(negative vs. positive valence), and their interaction entered as within-subject variables; and task-related functional connectivity between insula-FN, between insula-DN, within FN, and within DN, as between-subjects variables (together with demographic covariates). Significant effects in this (M)ANOVA are reported in the main text. Displayed here are scatterplots showing associations between task-related functional connectivity in each network of interest, and response speed to **(A-D)** trials in which emotional face stimuli were spatially manipulated in working memory (Switch), **(E-H)** trials in which emotional face stimuli were maintained in their original spatial position in working memory (Stay), and **(I-L)** the contrast of Switch – Stay trials. *Note:* On y-axis, reaction time scores are normalized and residualized for demographic covariates (age and gender); on x-axis, task-related network functional connectivity is normalized and residualized for covariates. Reported are correlation coefficients from correlations performed to follow up (M)ANOVAs. Significant r values, $*p<0.05$.

Supplementary References

1. Kaiser RH, Whitfield-Gabrieli S, Dillon DG, Goer F, Beltzer M, Minkel J, et al. Dynamic Resting-State Functional Connectivity in Major Depression. (2016) *Neuropsychopharmacology*. 41(7):1822-30. doi: 10.1038/npp.2015.352. PubMed PMID: WOS:000375968600014.
2. Kaiser RH, Snyder HR, Goer F, Clegg R, Ironside M, Pizzagalli DA. (2018) Attention Bias in Rumination and Depression: CognitiveMechanisms and Brain Networks. *Clinical Psychological Science.* 6(6): 765-782
3. Liu RT. (2016) Taxometric evidence of a dimensional latent structure for depression in an epidemiological sample of children and adolescents. *Psychological Medicine.* 46(6):1265-75. doi: 10.1017/s0033291715002792. PubMed PMID: WOS:000372513000012.
4. Widiger TA, Samuel DB. (2005) Diagnostic categories or dimensions? A question for the diagnostic and statistical manual of mental disorders-fifth edition. *Journal of Abnormal Psychology.* 114(4):494-504. doi: 10.1037/0021-843x.114.4.494. PubMed PMID: WOS:000233760700002.
5. Sheehan DV, Sheehan KH, Shytle RD, Janavs J, Bannon Y, Rogers JE, et al. (2010) Reliability and Validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). *Journal of Clinical Psychiatry*. 71(3):313-26. doi: 10.4088/JCP.09m05305whi. PubMed PMID: WOS:000276273500014.
6. Kaiser R, Kang M, Lew Y, Van der Feen J, Aguirre B, Clegg R, et al. Abnormal Frontoinsular-Default Network Dynamics in Adolescent Depression and Rumination: A Resting-State Co-Activation Pattern Analysis
7. Watson D, Clark LA, Tellegen A. (1988) Development and validation of brief measures of positive and negative affect – the PANAS scales. *Journal of Personality and Social Psychology.* 54(6):1063-70. doi: 10.1037/0022-3514.54.6.1063. PubMed PMID: WOS:A1988N666500016.
8. Wallace ML, McMakin DL, Tan PZ, Rosen D, Forbes EE, Ladouceur CD, et al. (2017) The role of day-to-day emotions, sleep, and social interactions in pediatric anxiety treatment. *Behaviour Research and Therapy*. 90:87-95. doi: 10.1016/j.brat.2016.12.012. PubMed PMID: WOS:000393248800010.
9. Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, et al. (2009) The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research.* 168(3):242-9. doi: 10.1016/j.psychres.2008.05.006. PubMed PMID: WOS:000269118700014.
10. Kaiser RH, Andrews-Hanna JR, Metcalf CA, Dimidjian S. (2015) Dwell or Decenter? Rumination and Decentering Predict Working Memory Updating After Interpersonal Criticism. *Cognitive Therapy and Research.* 39(6):744-53. doi: 10.1007/s10608-015-9697-1. PubMed PMID: WOS:000363965400003.
11. Snyder HR, Kaiser RH, Whisman MA, Turner AEJ, Guild RM, Munakata Y. (2014) Opposite effects of anxiety and depressive symptoms on executive function: The case of selecting among competing options. *Cognition & Emotion.* 28(5):893-902. doi: 10.1080/02699931.2013.859568. PubMed PMID: WOS:000335846600011.
12. Henderson RK, Snyder HR, Gupta T, Banich MT. (2012) When does stress help or harm? The effects of stress controllability and subjective stress response on Stroop performance.

Frontiers in Psychology. 3. doi: 10.3389/fpsyg.2012.00179. PubMed PMID: WOS:000208863900190.

13. Whitfield-Gabrieli S, Nieto-Castanon A. (2012) Conn: A functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connectivity*. doi: 10.1089/brain.2012.0073.
14. Behzadi Y, Restom K, Liau J, Liu TT. (2007) A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage*. 37(1):90-101. doi: 10.1016/j.neuroimage.2007.04.042. PubMed PMID: WOS:000248526700010.
15. Yeo BTT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*. 106(3):1125-65. doi: 10.1152/jn.00338.2011. PubMed PMID: WOS:000294775500007.
16. Baker JT, Holmes AJ, Masters GA, Yeo BTT, Krienen F, Buckner RL, et al. (2014) Disruption of Cortical Association Networks in Schizophrenia and Psychotic Bipolar Disorder. *Jama Psychiatry*. 71(2):109-18. doi: 10.1001/jamapsychiatry.2013.3469. PubMed PMID: WOS:000331370600002.
17. Woo CW, Chang LJ, Lindquist MA, Wager TD. (2017) Building better biomarkers: brain models in translational neuroimaging. *Nature Neuroscience*. 20(3):365-77. doi: 10.1038/nn.4478. PubMed PMID: WOS:000394920400009.
18. Pfister R, Schwarz KA, Janczyk M, Daleand R, Freeman JB. (2013) Good things peak in pairs: a note on the bimodality coefficient. *Frontiers in Psychology*. 4:4. doi: 10.3389/fpsyg.2013.00700. PubMed PMID: WOS:000331234600001.
19. Judd CM, McClelland GH, Ryan CS. (2009) Data analysis: A model comparison approach. 2nd ed. New York: Routledge.