Supplementary Material

Recruitment and Assessment of Subjects

Depressed adolescents were recruited from adolescent psychiatric and primary care clinics in San Diego, while healthy controls were recruited from the same geographic area via e-mail, internet, or flyers. Male and female adolescents from all ethnicities were allowed to participate. All participating adolescents provided written informed assent and their parents/legal guardian(s) provided written informed consent in accordance with the Declaration of Helsinki. All subjects received financial compensation for their participation.

The Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (KSADS-PL) was administered to all potentially depressed adolescents. All KSADS-PL diagnoses were verified by a board-certified child and adolescent psychiatrist. All MDD subjects in the study met full criteria for a current primary diagnosis of MDD and were medication-naïve at the time of scanning (see Table 1 in the main text for more details on the clinical characteristics of our MDD sample).

The computerized Diagnostic Interview Schedule for Children 4.0 and the Diagnostic Predictive Scale was used to screen for the presence of any Axis I diagnoses in the HCL adolescents.
In addition to completing forms on basic demographics and general medical and developmental history, all subjects completed the following within three days of their scan session: Edinburgh Handedness Inventory (Oldfield, 1971), Customary Drinking and Drug Use Record (Brown et al., 1998), Family Interview for Genetics Studies (Maxwell, 1992), Ishihara Color Plates Test (8

plates, 2005 edition), Standard Snellen Eye Chart, and Beck Depression Inventory-II (BDI-II; Beck et al., 1996). One subject (HCL) failed to complete the BDI-II and was therefore excluded from any analysis regarding this measure.

The exclusionary criteria for adolescents with MDD included any psychiatric comorbidities, left-handedness, being color blind or having less than 20/40 correctable vision, contraindication to MR imaging (e.g., pregnancy, claustrophobia, metallic implants), a serious medical or neurological illness, a learning disability, prior or present use of antidepressants, the use of medication with CNS effects within the past 2 weeks, evidence of illicit drug use or misuse of prescription drugs, and more than 2 alcoholic drinks per week or within the previous month at the time of scanning. Please see Table 1 for a summary of the clinical characteristics of our depressed subjects.

HCL adolescents were excluded from the study for any of the exclusionary criteria for the MDD group, as well as any current or lifetime Axis I psychiatric disorder, any family history of mood or psychotic disorders in first- or second-degree relatives.

Image Acquisition

All scanning was carried out on a General Electric Signa Excite 3T scanner (General Electric, Milwaukee, WI) with Twin Speed gradients and a GE 8-channel head coil. A fast spoiled gradient recalled sequence was used to collect T1-weighted images: TR=8 ms, TE=3 ms, TI=450 ms, flip angle=12°, 256x256 matrix, FOV=25x25 mm, 150 sagittal slices 1 mm thick with an in-plane resolution of 0.98x0.98 mm. T2*-weighted echo planar images were acquired using the following pulse sequence: TR=2000 ms, TE=32 ms, flip angle=90°, 64×64 matrix, FOV=23x23 mm, 192 repetitions, 30 oblique slices 2.6 mm thick with an in-plane resolution of 3.6x3.6 mm.

During scanning, subjects lay supine in the bore of the magnet and were instructed to relax but remain awake and as still as possible. Visual stimuli were projected onto a screen and viewed through a small, angled mirror mounted above the subject's head.

Image Preprocessing

The first 2 volumes of each EPI scan were excluded from analysis to allow for magnetic saturation and thus, signal stabilization. The derivative of the total outlier voxel count for each EPI acquisition was computed in order to find the volume with the least amount of head movement (i.e., the baseline volume). All EPI scans were slice time-corrected and realigned to their respective baseline volume. Each individual's anatomical scan was also aligned to their respective baseline EPI volume. Extreme outliers (2.5 standard deviations from mean) in the time series data were replaced on a voxel-by-voxel basis with values from the neighboring voxels. The total outlier voxel count was then computed again in order generate a list of time points to be censored from further analysis. Subjects were excluded from further analysis if more than 15% of the scan duration exhibited excessive motion; 4 subjects were therefore excluded on the basis of excessive motion. To account for individual variations in anatomical landmarks, a Gaussian filter with a full-width half-maximum (FWHM) of 4 mm was applied to the functional data. A multiple regression model was then used to fit the time series data. Regressors-of-interest included the three trial types: fear-strong (FS), fear-moderate (FM), and fear-neutral (FN). Six motion parameters and the time points flagged as outliers were also considered nuisance regressors to account for motion artifacts. Linear trend was also modeled in the time series of each voxel in order to account for correlated drift. Finally, the data were converted to percent signal change by dividing the time series of each voxel by the mean global signal and

transformed to stereotaxic coordinates (Talairach and Tournoux, 1988). Thus, all subsequent analyses were conducted at 4x4x4 mm resolution in Talairach space. Finally, since the primary focus of this study was on negative emotion processing, we limited our voxel-based analyses to the linear contrast of FearStrong-FearNeutral to maximize fear-related activation.

LBA Parameter Estimation

A hierarchical Bayesian method was used to estimate parameters in the LBA (Turner et al., 2012) at two levels: participant-level and group-level. Due to the limited number of trials in this experiment, all analyses were made collapsing across fear type. At the participant-level, the analysis estimates values for the parameters per participant (n=29). Figure S3 displays LBA fits for the individual participants. Using Markov Chain Monte Carlo simulations, we obtained full posterior distributions for each parameter. The participant-level parameters reported in the text and used for correlational analyses were therefore the median of this posterior distribution.

One advantage of this hierarchical Bayesian estimation method is that we are also able to obtain group-level estimates. In other words, we can get information about the MDD and HCL participants as two separate populations and statistically assess whether there is enough evidence to favor there being cognitive processing differences between these two groups. This analysis estimates the distribution of the participant-level parameters within the population of interest (MDD and HCL). All group-level distributions were assumed to be normal, truncated to positive values and defined by a mean (μ) and standard deviation (σ).

Figure S4 displays the posterior distributions of the mean (μ, top row) and variability (σ, bottom row) for each group-level parameter (columns). Figure 2 in the main text shows the difference in group distribution for each LBA parameter between MDD and HCL adolescents. This was done by firstly using the μ and σ hyper distributions for each parameter to calculate the mean of the associated truncated normal distribution. We then computed the differences in these means (MDD-mean minus HCL-mean). If the resulting distribution centered on a positive number on the x-axis, this indicates a larger mean parameter value for the MDD group, while a distribution centered on a negative value indicates a smaller parameter value for the MDD group. Consequently, a distribution centered on 0 (indicated in Figure 2 by the solid red line) would indicate no difference between the hyper distributions and thus, no differences between groups for the latent cognitive process corresponding to that LBA parameter.

Finally, we calculated odds ratios (ORs) for each difference distribution to provide a measure of statistical evidence for a difference between groups. For each group and each parameter we compared samples drawn from the true distribution (n=2000). A count was produced reflecting when the value from the MDD draw was larger than the value from the HCL draw. The mean count was then simply divided by 1 minus this count. All ORs were calculated to be greater than 1, for ease of interpretation.

Table S1. Location and size of significant clusters from the group effect. Results are based on a two-sample *t*-test of percent signal change values for the condition of interest (FearStrong-FearNeutral) between groups. Locations are reported according to center of mass of cluster in Talairach coordinates (radiological convention). See Figure S6 for more details. HCL=healthy controls; MDD=major depressive disorder; L=left; R=right.

Direction	**Cluster**	**x**	**y**	**z**	**# of voxels (volume)**
HCL > MDD	L Precuneus	-7	64	40	22 (1408 µL)
HCL > MDD	L ACC	-19	-33	21	11 (704 µL)
HCL > MDD	R Precentral Gyrus	49	11	30	11 (704 µL)

Table S2. Location and size of significant clusters resulting from an analysis of task effect. Results are based on a one-sample *t*-test of percent signal change values for the condition of interest (FearStrong-FearNeutral). Locations are reported according to center of mass of cluster in Talairach coordinates (radiological convention). See Figure S7 for more details. L=left; R=right; FFG=fusiform gyrus.

Cluster	**x**	**y**	**z**	**# of voxels (volume)**
R FFG	33	65	-9	165 (10560 μL)
L FFG	-32	67	-5	155 (9920 μL)

Figure S1. Schematic of the LBA model. Imagine a subject has to decide if a face is male or female. The LBA conceives of this two-choice perceptual decision as a race between two "accumulators" that accrue sensory evidence in favor of each choice over time. In this model, where the decision process starts is termed *starting point* (*A*) and the rate of evidence accumulation is termed the *drift rate* (v_c for correct responses and v_e for error responses) and the first accumulator to gather the criterion amount of evidence determines the subject's choice and response time (equivalent to the time taken for the accumulator to hit the *response threshold* (*b*) plus *non-decision time* (t_0) to account for sensory and motor processing time. Here, the accumulator for "male" hits the response threshold first, thereby determining the subject's final perceptual decision to be a fearful face. For more details, see Brown and Heathcote, 2008.

Figure S2. Summary of behavioral data. Accuracy **(a)** and RT data **(b)** for each group for each fear level. Overall accuracy (mean ± SEM) for MDD and HCL was 81.76% ± 1.9% and 86.96% ± 0.63%, respectively. Overall RT (mean ± SEM) for MDD and HCL was 1358.2ms ± 67.3ms and 1252.8ms ± 44.1ms, respectively. See *Results* in the main text for more details.

Figure S3. LBA fits for individual participants. Response time data plotted as histograms separately for correct (green) and error (red) responses. Model fits are overlaid with a solid line.

Figure S4. Histograms of full posterior distributions for LBA group-level parameters for each group. Posterior predictive distributions of μ (top row) and σ (bottom row) for each group-level parameter (columns).

Figure S5. Group differences in LBA parameters. Each panel shows the posterior predictive distribution of σ for the difference between the HCL and MDD groups for each parameter of the LBA model. Positive differences indicate larger parameter estimates for the MDD group, while negative differences indicate smaller parameter estimates for the MDD group. A distribution peaking at zero (denoted by the red line at x=0) indicates no difference between groups for that parameter. Odds ratios (ORs) indicating amount of evidence in favor of a difference are reported beneath each panel. See Figure 2 in the main text for posterior predictive distributions of μ for each LBA group-parameter.

Figure S6. Group differences in mean BOLD activation. Results are based on a two-sample *t*-test of percent signal change values during the condition of interest (FearStrong-FearNeutral) between groups and are significant at a cluster-wise $p<0.05$ (see *Controlling for multiple comparisons* under Methods in main text for more details). Results are overlaid over a standardized Talairach template. All locations are reported in Talairach coordinates (radiological convention). See Table S1 for more details.

Figure S7. Mean BOLD activation differences in task. Results are based on a one-sample *t*-test of percent signal change values for the condition of interest (FearStrong-FearNeutral) and are significant at a cluster-wise $p<0.01$ (see *Controlling for multiple comparisons* under Methods in the main text for more details). Locations are reported in Talairach coordinates (radiological convention). See Table S2 for more details.

Figure S1

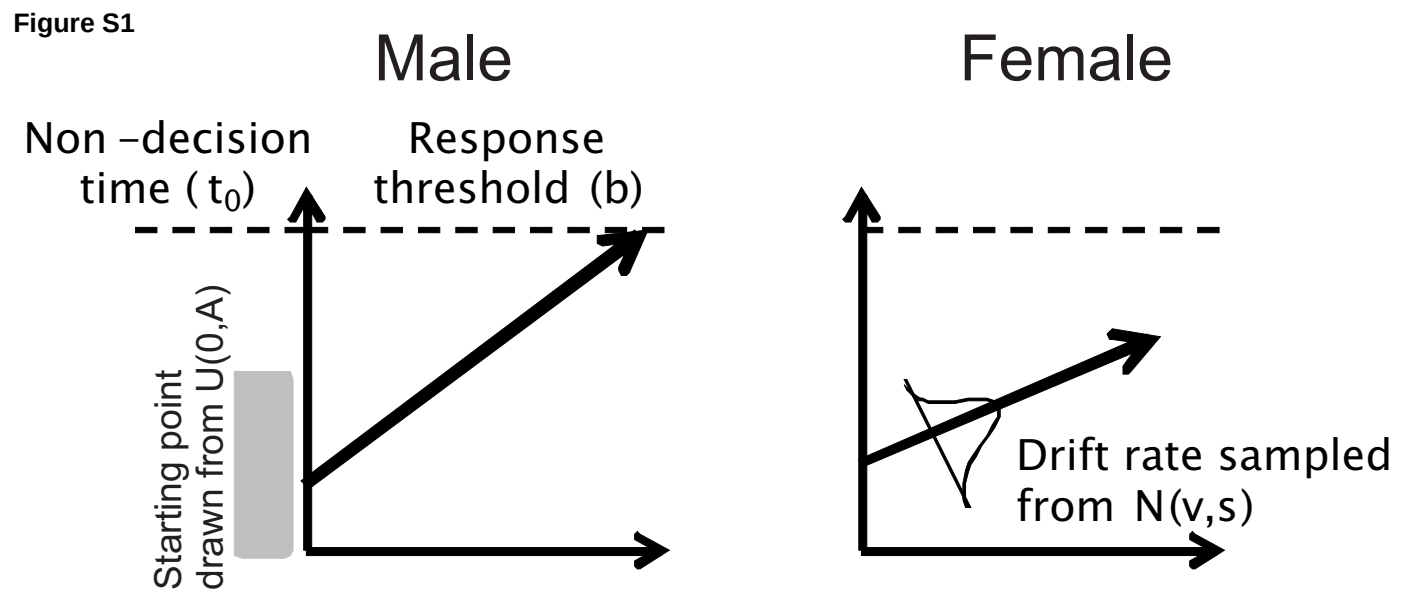

Male
Female
Non –decision time (t_0)
Response threshold (b)
Starting point drawn from U(0,A)
Drift rate sampled from N(v,s)

Figure S2

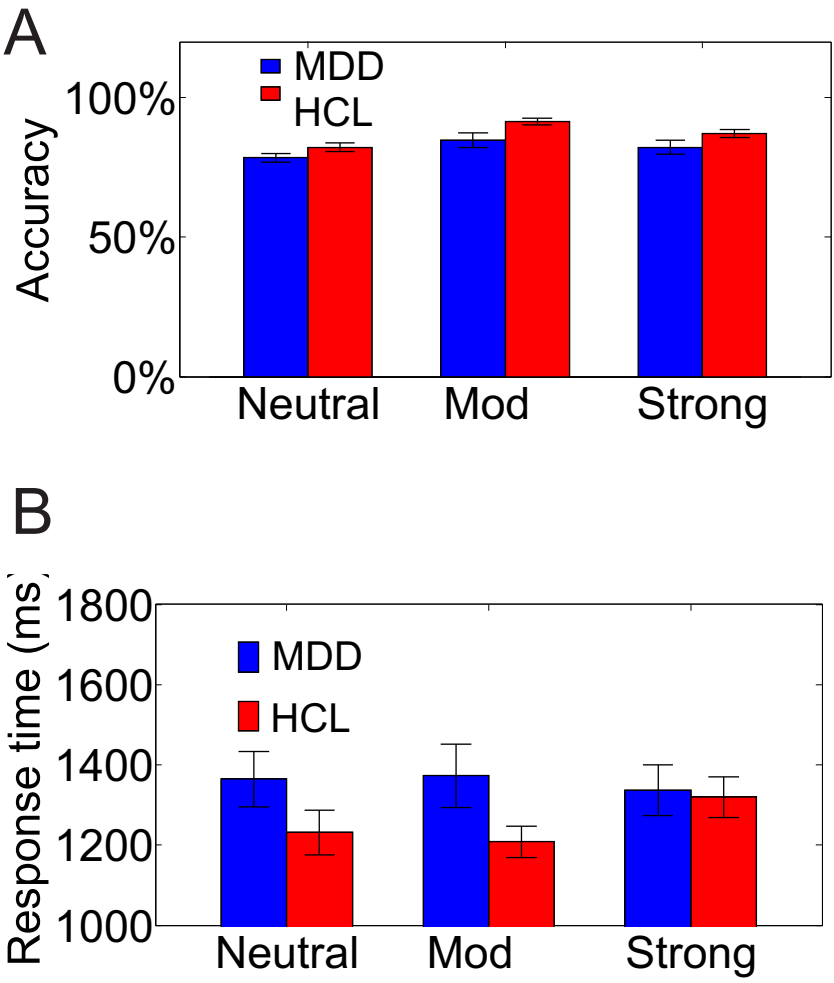

A
MDD
HCL
100%
50%
0%
Accuracy
Neutral
Mod
Strong
B
1800
1600
1400
1200
1000
Response time (ms)
MDD
HCL
Neutral
Mod
Strong

Figure S3

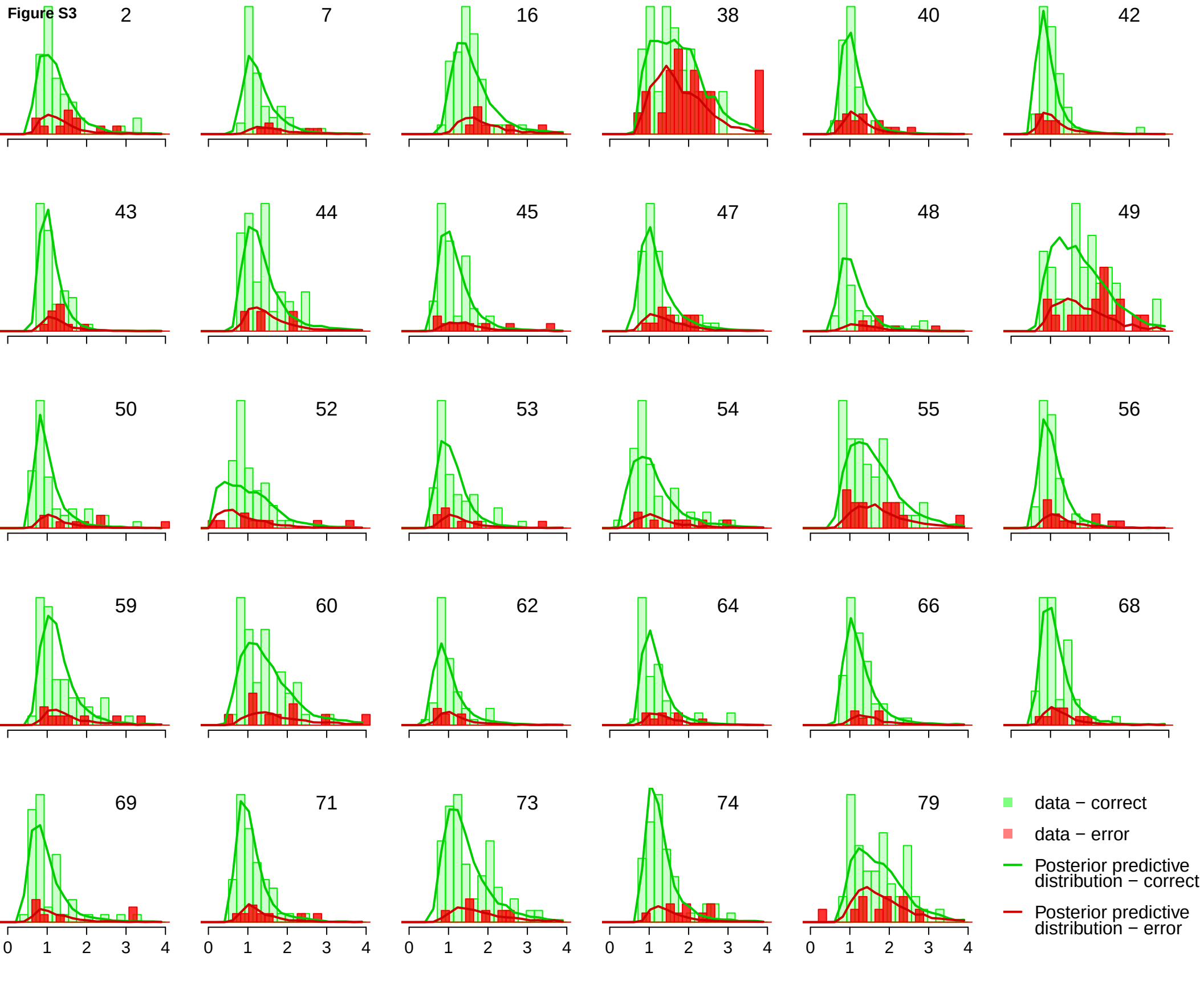

2
7
16
38
40
42
43
44
45
47
48
49
50
52
53
54
55
56
59
60
62
64
66
68
69
71
73
74
79
0
1
2
3
4
data − correct
data − error
Posterior predictive distribution − correct
Posterior predictive distribution − error

Figure S4

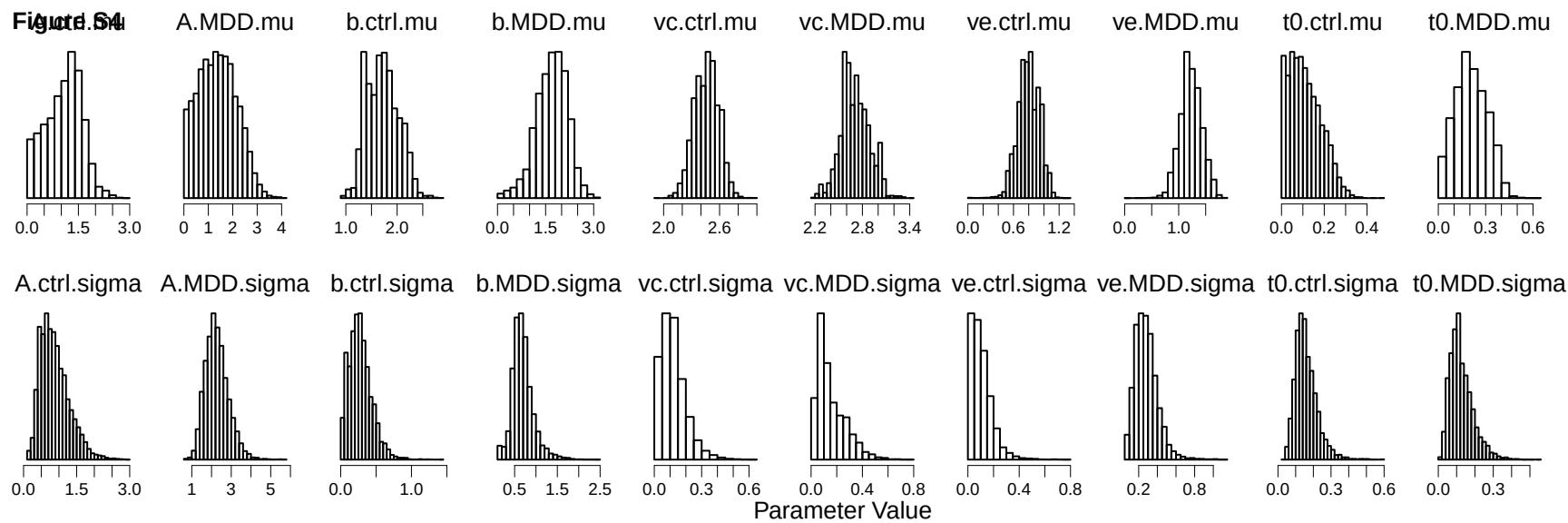

A.ctrl.mu
A.MDD.mu
b.ctrl.mu
b.MDD.mu
vc.ctrl.mu
vc.MDD.mu
ve.ctrl.mu
ve.MDD.mu
t0.ctrl.mu
t0.MDD.mu
0.0 1.5 3.0
0 1 2 3 4
1.0 2.0
0.0 1.5 3.0
2.0 2.6
2.2 2.8 3.4
0.0 0.6 1.2
0.0 1.0
0.0 0.2 0.4
0.0 0.3 0.6
A.ctrl.sigma
A.MDD.sigma
b.ctrl.sigma
b.MDD.sigma
vc.ctrl.sigma
vc.MDD.sigma
ve.ctrl.sigma
ve.MDD.sigma
t0.ctrl.sigma
t0.MDD.sigma
0.0 1.5 3.0
1 3 5
0.0 1.0
0.5 1.5 2.5
0.0 0.3 0.6
0.0 0.4 0.8
0.0 0.4 0.8
0.2 0.8
0.0 0.3 0.6
0.0 0.3
Parameter Value

Figure S5

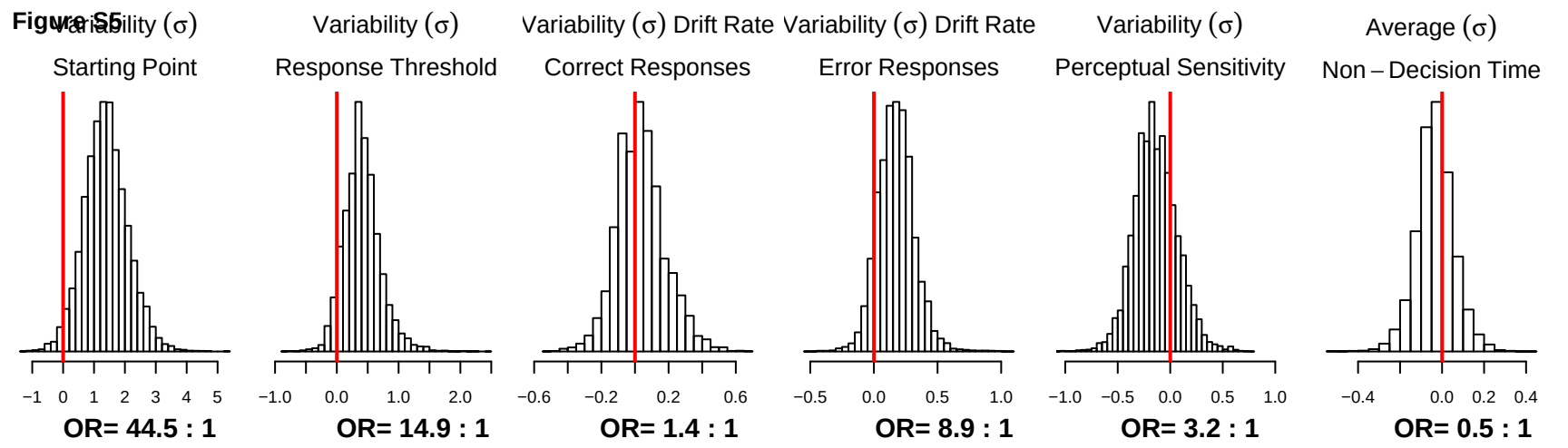
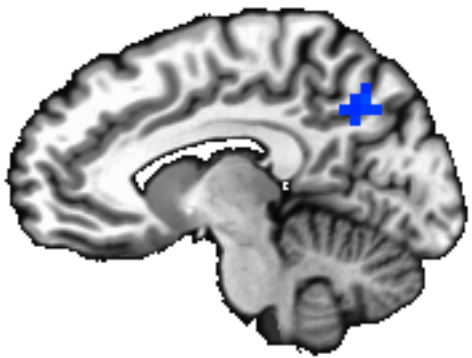
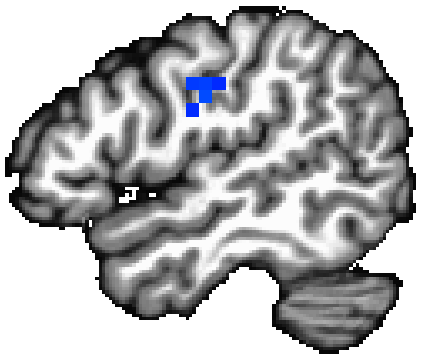

Variability (σ)
Starting Point
−1 0 1 2 3 4 5
OR= 44.5 : 1
Variability (σ)
Response Threshold
−1.0 0.0 1.0 2.0
OR= 14.9 : 1
Variability (σ) Drift Rate
Correct Responses
−0.6 −0.2 0.2 0.6
OR= 1.4 : 1
Variability (σ) Drift Rate
Error Responses
−0.5 0.0 0.5 1.0
OR= 8.9 : 1
Variability (σ)
Perceptual Sensitivity
−1.0 −0.5 0.0 0.5 1.0
OR= 3.2 : 1
Average (σ)
Non – Decision Time
−0.4 0.0 0.2 0.4
OR= 0.5 : 1
Difference (MDD minus HCL)

Figure S6



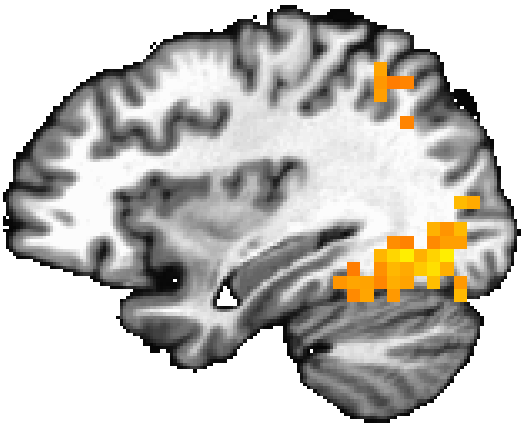
x=-8

x=-35

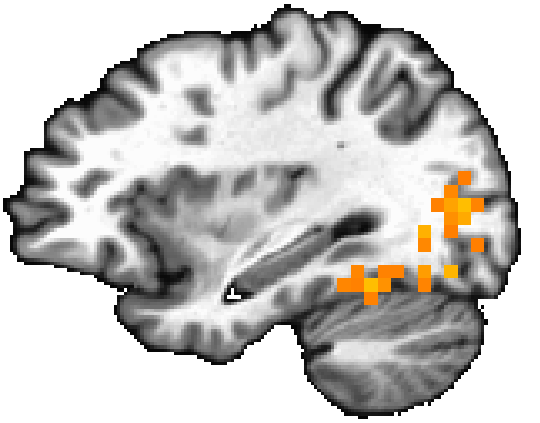


x=49

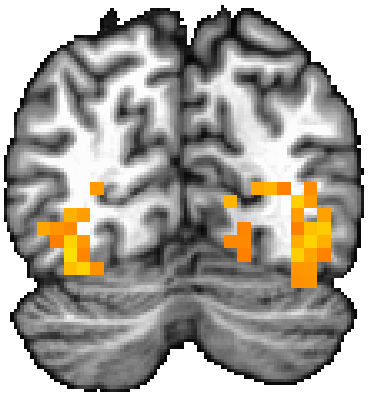
Figure S7



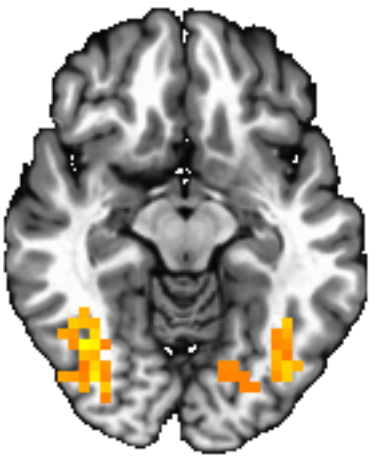
x=33



x=-31



y=10



z=-9