Anxiety and Stress Alter Decision-Making Dynamics and Causal Amygdala-Dorsolateral Prefrontal Cortex Circuits During Emotion Regulation in Children

Supplementary Information

Supplementary Methods

Drift diffusion modeling (DDM) of latent behavioral dynamics

Wiener distribution to characterize the drift diffusion process

The emotion evaluation process is characterized as a drift diffusion process, which posits that evidence accumulates over time resulting in a decision when a decision threshold is reached. (**Figure 3A**). The evaluations of an individual *i* on trial *t* with stimulus type *k* (neutral vs aversive) and instructions *p* (view vs reappraise) are denoted as x_{ipkt}, where $x_{ipkt} = 1$ if the evaluations are rated as 1 or 2 (positive) and $x_{ipkt} = -1$ if the evaluations are rated as 3 or 4 (negative). The response time of an individual is similarly denoted as r_{ipkt}. The choice coded response time is represented as $y_{ipkt} = x_{ipkt} r_{ipkt}$. The drift diffusion model implemented here assumes that x_{ipkt}, r_{ipkt}, y_{ipkt} are samples from random variables X_{ipk}, R_{ipk}, and Y_{ipk} respectively, and that Y_{ipk} is distributed according to a Wiener distribution (1, 2). The Wiener distribution thus describes the probability density function for Y_{ipk}, or in other words, the joint pdf for the random variables X_{ipk} and R_{ipk}, and is described by the following equations[1].

$$y_{ipkt} \sim \text{Wiener}(\alpha_{ip}, \tau_i, \beta_{ipk}, \delta_{ipk})$$

$$\text{Wiener}\,(\text{r}, \text{x} = 1|\alpha, \tau, \beta, \delta) = \frac{\pi}{\alpha^2} e^{(-\alpha(1-\beta)\delta)} \sum_{j=1}^{\infty} \text{j} \sin(\pi j(1-\beta))\, e^{\left(-\frac{1}{2}\left(\delta^2 + \frac{\pi^2 j^2}{\alpha^2}\right)(r-\tau)\right)}$$

$$\text{Wiener}\,(\text{r}, \text{x} = -1|\alpha, \tau, \beta, \delta) = \frac{\pi}{\alpha^2} e^{(-\alpha\beta\delta)} \sum_{j=1}^{\infty} \text{j} \sin(\pi j\beta)\, e^{\left(-\frac{1}{2}\left(\delta^2 + \frac{\pi^2 j^2}{\alpha^2}\right)(r-\tau)\right)}$$

[1] These equations assume a diffusion coefficient *s* set to 1 as implemented in the JAGS Wiener package. Further details can be obtained from Vandekerckhove, Tuerlinckx, & Lee (2011), and Wabersich & Vandekerckhove (2014).

The Wiener distribution is a four-parameter distribution that describes the first passage time of the diffusion process, that is, the distribution of times taken for the diffusion process to first hit the positive $(x = 1)$ and negative $(x = -1)$ decision boundaries. The key model parameters are described below.

Model parameters

α_{ip} is the *decision threshold* for individual *i* on trials with instruction *p* and represents the distance between the positive and negative decision boundaries. The decision threshold is commonly interpreted as a top down strategy, and we make the standard assumption of allowing the decision threshold for an individual to vary by instruction *p* (whether or not an individual is asked to reappraise) but not by stimulus type *k* (neutral versus negative images). A uniform prior ranging from (0.1,10) was placed on the decision threshold for each instruction, representing uninformative (flat) priors over a range that is based on prior theoretical and empirical results.

β_{ipk} is the *initial bias* for individual *i* represents the starting bias, or the starting point of the diffusion process between the two decision boundaries. For this task, the initial bias is assumed to reflect the initial reaction gained during the stimulus viewing period prior to the start of the decision period, and is thus allowed to vary by instruction *p* and type of stimulus *k*. A uniform prior over the entire plausible range (0,1) was placed on the initial bias for the aversive and neutral conditions. In addition, the initial bias for the reappraisal trials is constrained to lie between that for the neutral and aversive conditions.

δ_{ipk} is the *drift rate* for individual *i* represents the speed of evidence accumulation. High positive values indicate strong evidence accumulation towards positive evaluations and high negative values indicate strong evidence accumulation towards negative evaluations. Values closer to zero in any direction indicate relative ambiguity between positive and negative evaluations. The drift rate captures the key component of the decision process during the decision window and is allowed to vary by instruction *p* and type of stimulus *k*. An unconstrained prior N(0,1) was placed on the drift rates.

τ_i is the *non-decision time* for individual *i* represents the time taken by non-cognitive processes, such as motor response and perceptual processes. Since the non-decision time reflects non-cognitive processes, for each individual it is assumed to be fixed across both instruction and stimulus type. A uniform prior ranging from 0 to the shortest possible reaction time for each individual was placed on the non-decision time.

Model Fitting

The drift diffusion model DDM was implemented within a Bayesian inference framework (**Figure S3**) using JAGS for MCMC sampling. The sampling was checked for convergence (**Figure S4** demonstrates the convergences of 3 chains for a single parameter). The DDM produces posterior predictive reaction time and choice data based on the posterior distribution of parameters (**Figure S5** demonstrates the fit of the DDM model for a single participants' choice-coded reaction times).

Participants

Participants were recruited from a larger, longitudinal project testing the effectiveness of health and wellness methods in a suburban public-school district serving approximately 4,200 kindergarten through eighth grade students across eight schools in Northern California. Students in the district are predominantly Hispanic/Latinx (79%), African American (10%), Pacific Islander (9%) and other (2%). From this larger project, the present study recruited children in a narrow age range of 10 to 11 (fifth graders) to reduce age-related variability as brain regions involved in emotion regulation develop at variable rates in childhood (3). Parents and children provided written informed consent and assent, respectively, prior to participating in the study. All children were right-handed, medication-free, and met standard safety screening protocols for MRI. Subjects were scanned at the Richard M. Lucas Center for Imaging at Stanford University. Stanford University's Institutional Review Board approved all procedures.

fMRI data acquisition

Task-based functional data were acquired on a 3T GE scanner using a T2* weighted gradient echo-spiral in-out pulse sequence (TR = 2000ms, TE = 30ms, flip angle=80°, FOV=22cm, 3.4375 x

3.4375 x 4.5 mm^3 resolution; interleaved). A total of 31 axial slices were acquired, 4mm in thickness, and covering the whole brain. A T1-weighted, 166 slice high-resolution structural image was acquired (slice thickness 1mm; in-plane resolution: 256 x 256, voxel size x voxel size 0.9375 x 0.9375 mm^2) to facilitate registering each participant's data to standard space. Stimuli were presented using E-Prime and displayed using an LCD projector and a back-projection screen in the scanner suite. Participants responded to the rating scale using a 4-button box.

fMRI experimental design and emotion regulation task

Images were either neutral or aversive (meant to elicit a strong emotional reaction). Aversive and neutral images were selected from the International Affective Picture System (4) based on established norms for emotional valence and arousal dimensions. Selected aversive images were high-arousal whereas neutral images were low-arousal. Each image was shown only once for a given participant, and trials were counterbalanced across participants. Participants were coached in reappraisal strategies and practiced reappraising images that were not part of the fMRI task until the experimenter determined that they understood task instructions.

fMRI data preprocessing

Images were pre-processed and analyzed using SPM12. The preprocessing pipeline included realignment, slice-timing correction, volume repair, coregistration to subjects' T1 and normalization to a 2mm MNI152 template, and smoothing using a 6mm full-width half-maximum Gaussian kernel to decrease spatial noise. Volumes with greater than 0.2 mm scan to scan displacement along linear or rotational axes were deweighted, as well as volumes with greater than 5% change in global signal. The first three volumes of each time-series were discarded to allow for signal equilibrium.

Head motion

Scan-to-scan displacement was calculated based on parameters from realignment procedure. Subjects whose mean scan-to-scan displacement exceeded 1mm in either run were excluded from

analysis. After volume repair, subjects for whom more than 15% of volumes were repaired were also excluded from analysis. Subjects who did not comply with the task (did not make button presses) in more than 50% of the trials were also excluded from the analysis.

General linear model analysis

Task-related brain activation was assessed using a general linear model (GLM) implemented in SPM12. At the individual subject level, brain responses to each task condition (neutral, aversive, or reappraisal) was modeled using boxcar functions convolved with a canonical hemodynamic response function and a temporal derivative to account for voxelwise latency differences in hemodynamic response. Six head movement parameters generated from the realignment procedure were included to regress out effects of head movement on brain response. High-pass filtering using a cutoff of 1/128 Hz was applied. Serial correlations were accounted for by modeling the fMRI time series as a first-order autoregressive process (AR(1)) in the GLM framework. Stimulus-motion correlations across the two runs were significantly less than 0.2 mm on average (*ps* < 0.001), indicating that the inclusion of six head motion parameters in the GLM would not prevent us from detecting brain effects.

Contrast images for aversive vs neutral, reappraisal vs neutral, and reappraisal vs aversive conditions, generated at the individual level, were submitted to second-level, one-sample t-tests to examine changes in brain activity associated with emotion processing and regulation. Of note, as emotion regulation processes were best defined by the Reappraisal vs Aversive contrast, all brain and brain-behavioral analyses were conducted using the Reappraisal vs Aversive contrast.

In addition, although the Reappraisal vs Aversive contrast was our main focus, we chose not to use this contrast for ROI identification due to potential biases introduced by using the same contrast for selection (i.e., identification of brain regions of interests) and selective analysis (i.e., the subsequent MDS analysis using the selected brain regions). To alleviate such biases, we conducted an omnibus F-test to identify brain regions showing significant responses in either Reappraisal vs Neutral or Aversive vs Neutral contrast (**Figure 5A and S2**), and then identified activation peaks in bilateral amygdala and DLPFC regions (**Figure 5B**) to construct 6mm sphere

ROIs for subsequent causal analyses. Importantly, our DLPFC peaks identified from the omnibus F-test are consistent with those from previous meta-analysis of fMRI studies of emotion regulation (**Figure 5C**) (5-9). Additional prefrontal cortical regions were identified and used as control regions to probe specificity of our amygdala-DLPFC findings (**Table S5**).

To correct for multiple comparisons, and protect against detection of false activation clusters, we used Monte Carlo simulations to determine the cluster threshold at a false-positive cluster detection rate of $p < 0.01$. We generated random images, smoothed them with a 6mm FWHM smoothing kernel, used a whole brain mask excluding white matter, set a voxel-wise threshold of $p < 0.005$ and computed cluster sizes across 10,000 iterations (10). The minimum cluster size to control for false activation clusters and achieve the desired overall significance level (cluster-wise, $p < 0.01$) was determined to be 87 resampled voxels (or 696 mm^3). These thresholds are consistent with those previously reported in studies of emotion regulation in children and adults (11-14).

Computational modeling of dynamic causal interactions between amygdala and DLPFC

Multivariate dynamical systems (MDS) is a state-space model for estimating context-dependent causal interactions between multiple brain regions from fMRI data while accounting for variations in hemodynamic responses in these regions (16). MDS has been validated using extensive simulations (16-18) and has been successfully applied to our previous studies (19-21).

MDS consists of a state equation to model the latent "neuronal–like" states of the dynamic network and an observation equation to model BOLD-fMRI signals as a linear convolution of latent neural dynamics and HRF responses. The state equation in MDS is a multivariate linear difference equation or a first order multivariate auto regressive (MVAR) model that represents the time evolution of latent signals in M brain regions.

$$\boldsymbol{s}(t) = \sum_{j=1}^{J} v_j(t) C_j \, \boldsymbol{s}(t-1) + \boldsymbol{w}(t) \qquad (1)$$

The model for the observed BOLD responses is a linear convolution model

$$\boldsymbol{x}_m(t) = [\boldsymbol{s}_m(t)\ \boldsymbol{s}_m(t-1)\ldots.\boldsymbol{s}_m(t-L+1)]' \quad (2)$$

$$y_m(t) = b_m \Phi \boldsymbol{x}_m(t) + \boldsymbol{e}_m(t) \quad (3)$$

In Equation (1), $\boldsymbol{s}(t)$ is a $M \times 1$ vector of latent signals at time *t* of M regions, C_j is an $M \times M$ connection matrix ensued by modulatory input $v_j(t)$, and *J* is the number of modulatory inputs. The non-diagonal elements of C_j represent the coupling of brain regions in the presence of $v_j(t)$. $C_j(m,n)$ denotes the strength of causal connection from n-th region to m-th region for j-th type stimulus. Therefore, latent signals ***s(t)*** in *M* regions at time *t* is a bilinear function of modulatory inputs $v_j(t)$ and its previous state ***s(t-1)***. $\boldsymbol{w}(t)$ is an $M \times 1$ state noise vector whose distribution is assumed to be Gaussian distributed with covariance matrix *Q(* $\boldsymbol{w}(t) \sim N(0, Q)$). Additionally, state noise vectors at time instances *1,2,….,T* ($\boldsymbol{w}(1), \boldsymbol{w}(2) \ldots \boldsymbol{w}(T)$) are assumed to be identical and independently distributed (iid). The latent dynamics modeled in equations (1) and (2) give rise to observed fMRI time series represented by Equation (3).

The fMRI-BOLD time series in region *m* were modeled as a linear convolution of HRF and latent signal $\boldsymbol{s}_m(t)$ in that region. To represent this linear convolution model as an inner product of two vectors, the past *L* values of $\boldsymbol{s}_m(t)$ are stored as a $L \times 1$ vector $\boldsymbol{x}_m(t)$ in equation (2).

Equation (3) represents the linear convolution between the embedded latent signal $\boldsymbol{x}_m(t)$ and the basis vectors for HRF. Specifically, $y_m(t)$ is the observed BOLD signal at time *t* of *m*-th region. Φ is a $p \times L$ matrix whose rows contain bases for HRF. Here, we use the canonical HRF and its time derivative as bases, as is common in most fMRI studies. b_m is a $1 \times p$ coefficient vector representing the weights for each basis function in explaining the observed BOLD signal $y_m(t)$. Therefore, the HRF in *m-th* region is represented by the product $b_m\Phi$. The BOLD response in this region is obtained by convolving HRF ($b_m\Phi$) with the *L* past values of the region's latent signal ($\boldsymbol{x}_m(t)$) and is represented mathematically by the vector inner product $b_m\Phi\ \boldsymbol{x}_m(t)$. Uncorrelated observation noise $\boldsymbol{e}_m(t)$ with zero mean and variance σ_m^2 is then added to generate the observed signal $y_m(t)$. $\boldsymbol{e}_m(t)$ is also assumed to be uncorrelated with $\boldsymbol{w}(\tau)$, at all *t* and τ.

Equations (1-3) together represent a state-space model for estimating the causal interactions in latent signals based on observed multivariate fMRI time series. Crucially, MDS also takes into account variations in HRF as well as the influences of modulatory stimuli in estimating causal interactions between the brain regions.

Estimating causal interactions between *M* regions specified in the model is equivalent to estimating the parameters $C_j, j = 1,2..J$. In order to estimate C_j's, we need to first estimate the other unknown parameters Q, $\{b_m\}_{m=1}^{M}$ and $\{\sigma_m^2\}_{m=1}^{M}$ and the latent signal $\{\boldsymbol{s}^s(t)\}_{t=1}^{T}$ based on the observations $\{y_m^s(t)\}_{m=1,s=1}^{M,S}, t = 1,2..T$, where *T* is the total number of time samples and *S* is the number of subjects. A variational Bayes approach (VB) was used for estimating the posterior probabilities of the unknown parameters of the MDS model given fMRI time series observations for *S* number of subjects (16).

Prior to applying MDS, the fMRI time-series for each region *m* and subject *s*, $y_m^s(t)$, was linearly de-trended, its temporal mean removed and normalized by its standard deviation. For all regions, time-series (1st eigenvector) were extracted from preprocessed fMRI data using MarsBar toolbox in SPM12. Spherical ROIs were defined as the sets of voxels contained in 6 mm spheres centered on the MNI coordinates of each ROI. MDS was applied to estimate causal interactions between the amygdala and DLPFC for neutral, aversive, and reappraisal conditions. Only two brain regions – amygdala and DLPFC within the same hemisphere – were entered into MDS each time to estimate their causal interactions. For each directed causal connection (e.g., right amygdala to right DLPFC), the difference in causal connection magnitude between reappraisal and aversive conditions were calculated for each subject, which were further fed into a one sample t-test to assess statistical significance. Moreover, we examined whether the difference in causal connection between the reappraisal and aversive condition was associated with anxiety and stress reactivity severity or behavioural reappraisal ability using Pearson correlations. False Discovery Rate (FDR) correction was employed to correct for multiple comparisons. Results with FDR corrected $p < 0.05$ were considered significant.

Supplementary Results and Discussion

Latent behavioral dynamics are associated with self-reports of ability to reappraise emotion

Hierarchical linear regressions with reappraisal success as the dependent variable and changes in initial bias, drift rate, and decision threshold under Reappraisal vs. Aversive conditions as the independent variables revealed an excellent model fit (adjusted R2 = 0.78, F(3, 41) = 53.64, p < 0.001). The R^2 increased by 0.639 ($F(1, 41) = 129.05$, $p < 0.001$) when adding drift rate to the model containing initial bias and decision threshold. In contrast, R^2 increased by only 0.005 ($F(1, 41) = 1.07$, $p = 0.31$) when adding initial bias to the model containing drift rate and decision threshold, and by 0.018 ($F(1, 41) = 3.69$, $p = 0.06$) when adding decision threshold to the model containing drift rate and initial bias. Change in drift rate from the Aversive to Reappraisal conditions was the only independent variable that contributed unique variance and thus emerged as the dominant predictor ($t(41) = 11.36$, $\beta = 0.99$, $p < 0.001$).

DDM revealed that emotion regulation in children is characterized by an increase in initial bias and higher drift rate. However, hierarchical linear regression revealed that children's reappraisal behavioral ratings were primarily driven by higher drift rates associated with their ability to regulate evaluation of their emotional reactivity to negative stimuli rather than initial bias (**Figure S1**). This result suggests that success in emotion regulation is primarily related to an evaluative decision-making process, rather than reaction suppression during initial viewing. In other words, the regulation of immediate emotional reactions captured by the bias parameter, which characterizes the positive or negative bias induced during the initial viewing of the aversive images, is not described by the overt reappraisal score, which measures mainly the ability to regulate an evaluation during the post-viewing response period.

Correlation between anxiety and stress

In the current sample (n = 45), anxiety and stress measures were correlated, with $t(43) = 5.93$, $r = 0.67$, and $p < 0.001$. We performed additional analyses to differentiate the roles of anxiety and stress in their relation to amygdala → DLPFC causal interaction during emotion regulation

(Reappraisal vs Aversive condition). Specifically, (i) residualized anxiety was derived by regressing stress out from anxiety and (ii) residualized stress was derived by regressing anxiety out from stress. The strength of causal influence from the right amygdala to right DLPFC was not correlated with residualized anxiety ($t(42) = 1.54$, $r = 0.23$, $p = 0.13$) or residualized stress ($t(42) = 1.06$, $r = 0.16$, $p = 0.30$). Formal structural equation modelling revealed a significant relationship between the strength of right amygdala → DLPFC causal interactions and a latent factor underlying anxiety and stress (**Figure 6D;** CFI = 1, TLI = 1, RMSEA = 0.00). These results suggest that shared variance between anxiety and stress reactivity drives bottom-up amygdala → DLPFC signaling during emotion regulation.

Brain activity during emotion regulation

Compared with neutral stimuli, aversive stimuli during passive viewing elicited greater activations in bilateral amygdala, right DLPFC, as well as bilateral anterior insula, bilateral DMPFC, right VLPFC, right caudate, and bilateral occipital cortices (**Figure S2A and Table S1**). Compared with neutral stimuli, aversive stimuli during Reappraisal elicited greater activations in bilateral amygdala, bilateral DLPFC, bilateral DMPFC, bilateral VLPFC, right SPL, and bilateral occipital cortices (**Figure S2B and Table S2**). Compared with aversive stimuli during passive viewing, aversive stimuli during Reappraisal elicited greater activations in left DLPFC, bilateral DMPFC, bilateral VLPFC, left angular gyrus, right supramarginal gyrus, left precuneus, bilateral posterior cingulate cortex, bilateral middle temporal gyrus, and bilateral occipital cortices (**Figure S2C and Table S3**). These findings were consistent with the notion that reappraisal modulates affective responding via recruitment of a frontoparietal network (5-8, 22). Moreover, we did not identify decreases in amygdala activation under the Reappraisal vs Aversive contrast. While decreases in amygdala activation (a proxy of successful down-regulation of negative affect) are often identified in adult studies using reappraisal to down regulate negative affect, developmental studies find that the ability to regulate emotion improves with age (22). Our failure to identify decreases in amygdala activation in children (10-11 years-old) during emotion regulation is consistent with findings in previous developmental studies which included children of similar age (13, 14, 23). Additionally, participants were asked to reappraise aversive images by telling themselves a story

to make the pictures seem less negative, or more positive, which may have confounded down- and up-regulation strategies affecting amygdala activity.

Amygdala and PFC responses during emotion regulation are not correlated with anxiety and stress

Anxiety and stress were not correlated with activation levels in amygdala or any of the PFC ROIs, including anterior insula (MNI: 34, 24, 0; -36, 20, -6), DMPFC (MNI: 8, 28, 46; -6, 24, 42), VLPFC (MNI: 52, 24, 4; -50, 22, 12), and DLPFC (MNI: 40, 8, 38; -42, 12, 44), during emotion regulation (Reappraisal vs Aversive).

Strong causal effects from amygdala to DLPFC and the reverse

We found strong causal effects from amygdala to DLPFC and the reverse in both hemispheres under Neutral, Aversive, and Reappraisal conditions (**Figure S6**). However, the strength of DLPFC → amygdala signalling was not correlated with anxiety or stress. It is possible that the DLPFC action on the amygdala to moderate the effects of anxiety and stress may require multiple brain regions.

Right lateralization of causal amygdala→DLPFC circuits associated with anxiety and stress provides an actionable target

Stress and anxiety measures were associated with right hemisphere connectivity, consistent with neuroimaging findings of right lateralized anxiety related signals (25-27). Control analyses revealed that the effects were specific to right hemisphere amygdala-DLPFC as the left hemisphere did not show the same effects. Present results are consistent with neuropsychological and frontal-cortical models of emotion which suggest that unpleasant, withdrawal-related emotions depend on the right hemisphere and pleasant, approach-related emotions are dependent on the left hemisphere (28-30). As anxiety is characterized by unpleasant valence and high arousal, greater right hemisphere engagement is predicted but a novel aspect of our findings is the right lateralization of causal amygdala→DLPFC circuits associated with anxiety and stress. These models and

neuroimaging evidence are relevant to advancing our understanding of the neurobiological mechanisms by which both anxiety and stress impact cortical function.

Our identification of a specific right lateralized amygdala-PFC circuit associated with children's emotion regulation provides a novel target for probing early stress etiology and for facilitating early interventions to prevent the development of long-term psychopathology. Indeed, rTMS is now being widely harnessed for the treatment of anxiety in adults by targeting the right DLPFC as one key mechanism of intervention (31). Convergent on this approach, our findings suggest that amygdala-DLPFC interactions may represent a vulnerable functional circuit for early interventions of treating children at risk for psychopathology. Furthermore, our results demonstrated right hemisphere amygdala-DLPFC causal dysfunction across multiple sources of stress and anxiety which may have transdiagnostic implications. Taken together with the wealth of literature on TMS/tDCS effects on affective circuits via the DLPFC, these findings suggest an actionable circuit for transdiagnostic targeting of developmental psychopathology.

Supplementary Figures

Figure S1. Relationship between drift rate and reappraisal ability. Individual reappraisal scores were significantly correlated with changes in drift rate between the Reappraisal and Aversive conditions.

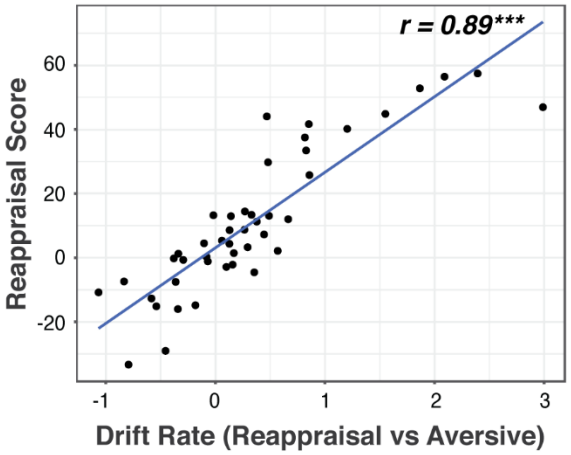


Figure S2. The effects of Reappraisal and Aversive emotion processing on brain activity. (**A**) Compared with neutral stimuli, aversive stimuli during passive viewing elicited greater activations in bilateral amygdala, right DLPFC, as well as bilateral anterior insula, bilateral DMPFC, right VLPFC, right caudate, and bilateral occipital cortices. (**B**) Compared with neutral stimuli, aversive stimuli during Reappraisal elicited greater activations in bilateral amygdala, bilateral DLPFC, bilateral DMPFC, bilateral VLPFC, right SPL, and bilateral occipital cortices. (**C**) Compared with aversive stimuli during passive viewing, aversive stimuli during Reappraisal elicited greater activations in left DLPFC, bilateral DMPFC, bilateral VLPFC, left angular gyrus, right supramarginal gyrus, left precuneus, bilateral posterior cingulate cortex, bilateral middle temporal gyrus, and bilateral occipital cortices.

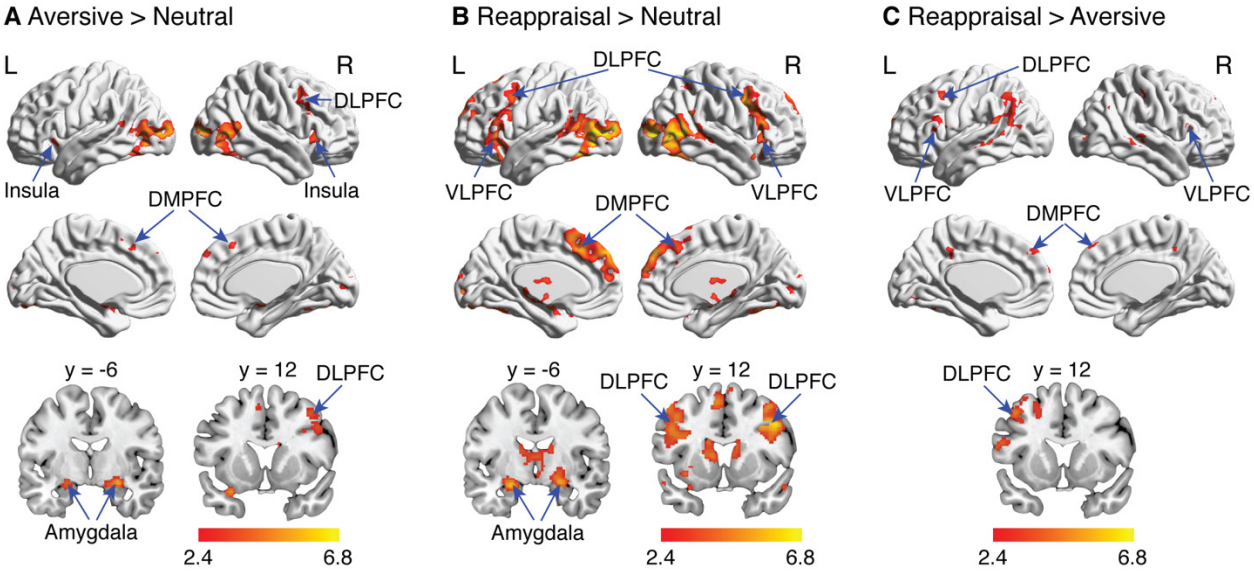


Figure S3. Graphical representation of the drift diffusion model (DDM). Graphical representation of the DDM with nodes capturing the relationship between model parameters (unshaded) and observed data (shaded). The model for emotion evaluation was implemented within a Bayesian inference framework using JAGS (32) to implement Markov chain Monte Carlo sampling. This allowed us to obtain the full posterior distribution for all parameters of the model, conditional on the observed data. The mean values of these posterior distributions were used in further analysis. The absolute model fit was validated by comparing the posterior predictive values generated by the model, given the inferred posterior parameter values, for both the emotion evaluations and response times under the three different conditions to the observed data. The posterior samples were checked for convergence using a convergence statistic (33). Inference about the posterior distribution of parameters requires specification of prior distributions.

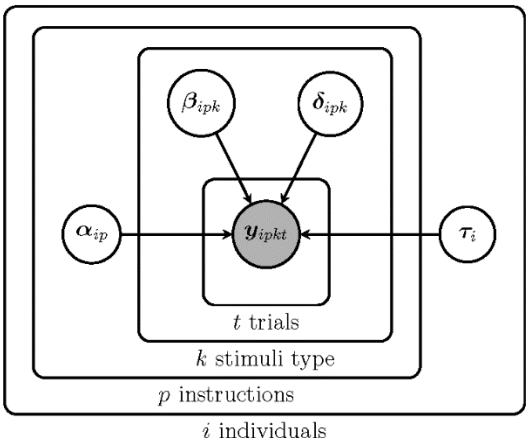


Figure S4. MCMC sampling showing convergence of chains. An example of the MCMC sampling for 1000 retained samples for one parameter, obtained after discarding 1000 burn-in samples. This shows the convergence of the chains, showing that the parameter is being samples from stationary distributions.

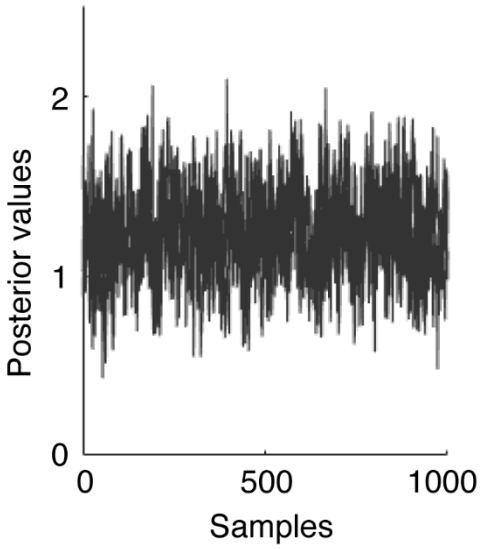


Figure S5. DDM fits to choice coded reaction time. An example of the posterior predictive distribution of choice coded reaction times produced by the generative model for a single participant. The black line shows the model generated posterior predictive samples that approximate the observed behavior for this participant as shown by the gray bars.

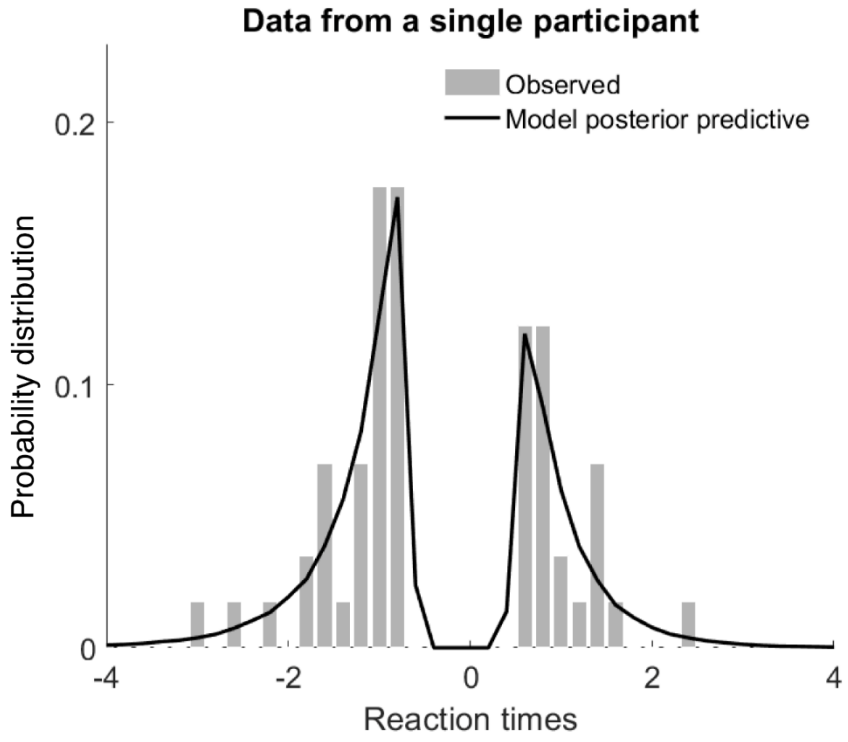
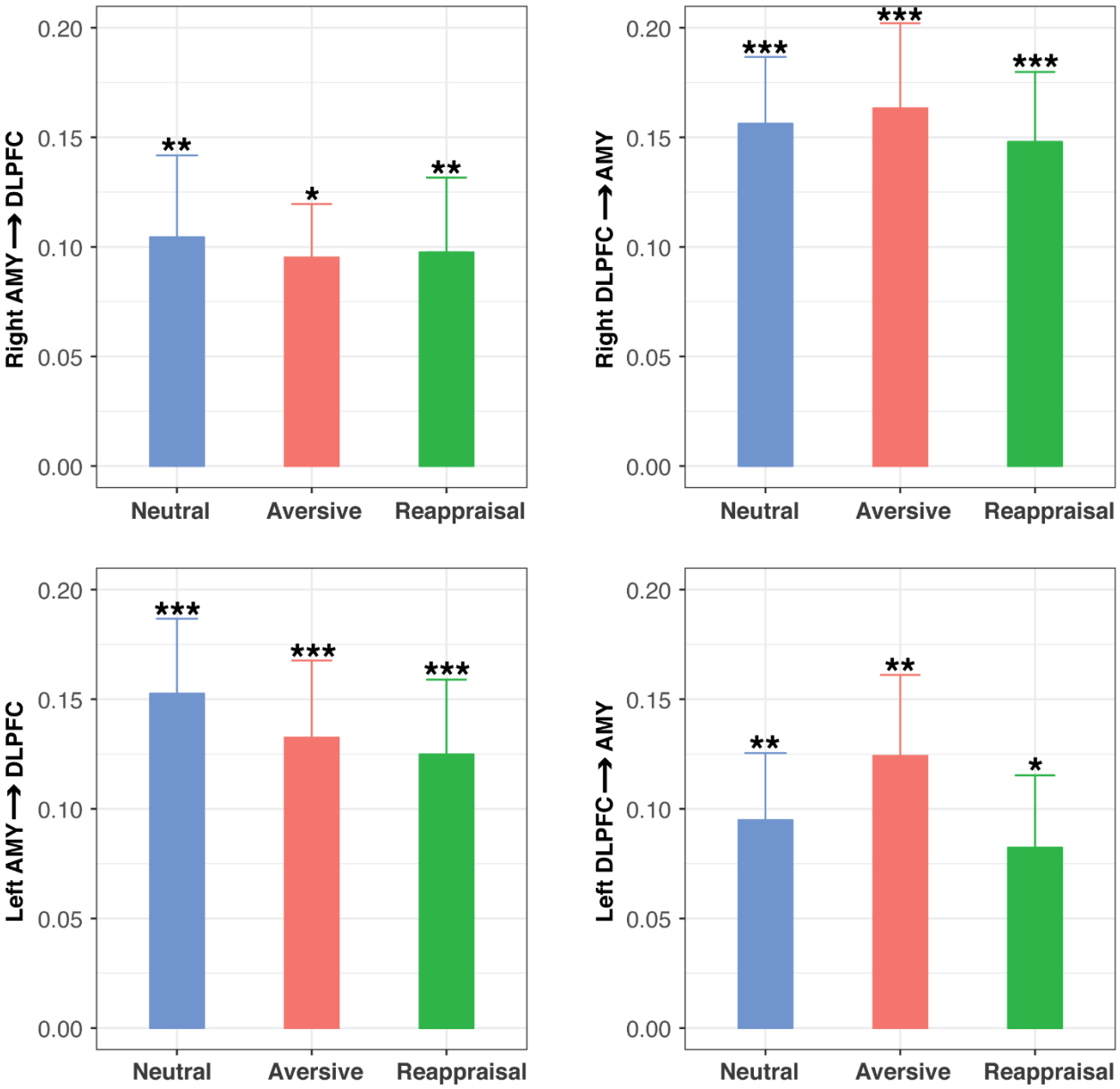


Figure S6. Causal signaling from amygdala to DLPFC and reverse. Causal signaling from amygdala to DLPFC/IFJ and the reverse within in each hemisphere under Neutral, Aversive, and Reappraisal conditions. DLPFC: dorsolateral prefrontal cortex; IFJ: inferior frontal junction. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.



Supplementary Tables

Table S1. Participant characteristics

Characteristic	Mean ± SD
Gender (male/female)	56% / 44%
Hispanic/Latinx	80%
Non-Hispanic/Latinx	4%
Ethnicity not reported	16%
Age	10.76 ± 0.52
RSQ Involuntary Stress Response	0.55 ± 0.37
BASC-SR Anxiety (T Score)	52.2 ± 8.95
MPI Global IQ Score	99.63 ± 11.56

RSQ: Response to Stress Questionnaire; BASC-SR: Behavioral Assessment System for Children, Self-Report; MPI: Mental Processing Index from the Kaufman Assessment Battery for Children, Second Edition.

Table S2. Brain activations of Aversive vs. Neutral contrast (height threshold $p < 0.005$, extent threshold $k = 87$ voxels).

Region	Hemisphere	Cluster size (#Voxels)	Peak T-value	MNI coordinates		
				x	y	z
Occipital pole/LOC	R	2,985	6.81	28	-94	6
	L	2,913	6.57	-26	-90	4
Temporal pole	L	199	5.17	-32	10	-22
Amygdala	L		3.59	-22	-6	-14
	R	161	4.87	26	-6	-14
DLPFC (MFG)	R	544	4.57	52	18	32
Insula	R	343	4.45	34	24	-2
VLPFC (IFG)	R		3.96	46	28	10
DMPFC	R	167	4.00	6	56	42
	R	175	3.97	6	26	48
	L		3.42	-6	24	44
Caudate	R	131	3.87	16	6	18
OFC/Insula	L	178	3.86	-36	24	-8

DLPFC: dorsolateral prefrontal cortex; DMPFC: dorsomedial prefrontal cortex; IFG: inferior frontal gyrus; LOC: Lateral occipital cortex; MFG: middle frontal gyrus; OFC: orbitofrontal cortex; VLPFC: ventrolateral prefrontal cortex.

Table S3. Brain activations of Reappraisal vs. Neutral contrast (height threshold $p < 0.005$, extent threshold $k = 87$ voxels).

Region	Hemisphere	Cluster size (#Voxels)	Peak T-value	MNI coordinates		
				x	y	z
LOC/Occipital pole	L	5,404	9.31	-44	-80	4
	R	5,735	8.78	50	-70	0
DLPFC (MFG)	R	2,386	6.75	40	8	38
OFC/Insula	R		6.00	42	28	-12
VLPFC (IFG)	R		5.59	58	26	18
Amygdala	R		4.69	22	-6	-14
DMPFC	L	3,314	6.59	-4	52	38
	R		6.26	6	50	40
VLPFC (IFG)/ DLPFC (MFG)	L	5,923	5.80	-48	26	20
Amygdala	L		5.44	-24	-6	-16
SPL	R	103	4.57	32	-54	54

DLPFC: dorsolateral prefrontal cortex; DMPFC: dorsomedial prefrontal cortex; IFG: inferior frontal gyrus; LOC: lateral occipital cortex; MFG: middle frontal gyrus; OFC: orbitofrontal cortex; SPL: superior parietal lobule; VLPFC: ventrolateral prefrontal cortex.

Table S4. Brain activations of Reappraisal vs. Aversive contrast (height threshold $p < 0.005$, extent threshold $k = 87$ voxels).

Region	Hemisphere	Cluster size (#Voxels)	Peak T-value	MNI coordinates		
				x	y	z
Angular gyrus	L	1,970	5.02	-58	-56	26
MTG	L		4.67	-64	-44	6
VLPFC (IFG)	L	378	4.34	-48	22	22
MTG/STG	R	403	4.31	68	-32	0
LOC	L	133	4.30	56	-62	36
DLPFC (MFG)	L	395	4.13	-42	12	46
OFG	L	169	4.08	-22	-82	-10
Precuneus	L	332	3.93	-4	-40	46
PCC	L		3.90	-6	-44	28
	R		3.29	8	-44	32
VLPFC (IFG)	R	103	3.92	58	22	2
	R		3.43	58	26	14
LOC	R	123	3.83	36	-86	2
DMPFC	L	193	3.82	-6	66	22
	L		3.52	-16	64	24
	R		3.01	2	60	8
postCG/SMG	R	136	3.63	50	-28	50
DMPFC	L	130	3.60	-6	54	44
	R		3.34	6	48	44
	R		3.04	4	40	50

DLPFC: dorsolateral prefrontal cortex; DMPFC: dorsomedial prefrontal cortex; IFG: inferior frontal gyrus; LOC: Lateral occipital cortex; MFG: middle frontal gyrus; MTG: middle temporal gyrus; OFC: orbitofrontal cortex; PCC: posterior cingulate cortex; postCG: postcentral gyrus; SMG: supramarginal gyrus; STG: superior temporal gyrus; VLPFC: ventrolateral prefrontal cortex.

Table S5. Brain activations of Reappraisal vs Neutral or Aversive vs Neutral contrast (height threshold $p < 0.005$, extent threshold $k = 87$ voxels).

Region	Hemisphere	Cluster size (#Voxels)	Peak F-value	MNI coordinates		
				x	y	z
LOC	L	5922	40.11	-44	-80	4
	R	4552	38.87	50	-70	2
Precuneus	L	2435	15.45	-6	-66	24
	R		12.96	12	-68	24
DMPFC	L	1863	19.12	-4	50	40
	R		13.96	6	50	42
	L		12.23	-6	24	42
	R		11.13	8	28	46
VLPFC (IFG)	L	1785	14.73	-50	22	12
	L		14.44	-48	24	20
DLPFC (MFG)	L		11.58	-42	12	44
Insula	L		10.12	-36	20	-6
Amygdala	L		9.84	-24	-6	-14
DLPFC (MFG)	R	1112	17.94	40	8	38
	R		14.11	48	14	34
DLPFC (MFG)	R		10.30	46	20	24
preCG	L	792	15.82	-2	-36	50
Lingual gyrus	L	336	14.73	-12	-70	-8
Caudate	L	351	12.66	-14	8	20
	R	123	11.00	14	6	18
Insula	R	280	9.52	34	24	0
SMG	R	181	12.09	56	-42	24
Amygdala	R	106	11.18	22	-6	-14

DLPFC: dorsolateral prefrontal cortex; DMPFC: dorsomedial prefrontal cortex; IFG: inferior frontal gyrus; LOC: lateral occipital cortex; MFG: middle frontal gyrus; preCG: precentral gyrus; SMG: supramarginal gyrus; VLPFC: ventrolateral prefrontal cortex.

Table S6. MNI coordinates of amygdala and prefrontal ROIs identified using the omnibus F-test (height threshold $p < 0.005$, extent threshold $k = 87$ voxels).

Region	Hemisphere	MNI coordinates		
		x	y	z
Amygdala	R	22	-6	-14
	L	-24	-6	-14
DLPFC/IFJ	R	40	8	38
	L	-42	12	44
DMPFC	R	8	28	46
	L	-6	24	42
VLPFC	R	52	24	4
	L	-50	22	12
Insula	R	34	24	0
	L	-36	20	-6

DLPFC: dorsolateral prefrontal cortex; DMPFC: dorsomedial prefrontal cortex; IFJ: inferior frontal junction; VLPFC: ventrolateral prefrontal cortex; L: left; R: right.

Supplementary References

1. Vandekerckhove J, Tuerlinckx F, Lee MD (2011): Hierarchical diffusion models for two-choice response times. *Psychological methods*. 16:44.
2. Wabersich D, Vandekerckhove J (2014): Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behavior research methods*. 46:15-28.
3. Mills KL, Goddings AL, Clasen LS, Giedd JN, Blakemore SJ (2014): The developmental mismatch in structural brain maturation during adolescence. *Dev Neurosci*. 36:147-160.
4. Lang PJ, Bradley MM, Cuthbert BN (1999): International affective picture system (IAPS): Instruction manual and affective ratings. *The center for research in psychophysiology, University of Florida*.
5. Brandl F, Corbi ZLH, Bratec SM, Sorg C (2019): Cognitive reward control recruits medial and lateral frontal cortices, which are also involved in cognitive emotion regulation: A coordinate-based meta-analysis of fMRI studies. *NeuroImage*. 200:659-673.
6. Buhle JT, Silvers JA, Wager TD, Lopez R, Onyemekwu C, Kober H, et al. (2014): Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cereb Cortex*. 24:2981-2990.
7. Frank DW, Dewitt M, Hudgens-Haney M, Schaeffer DJ, Ball BH, Schwarz NF, et al. (2014): Emotion regulation: quantitative meta-analysis of functional activation and deactivation. *Neurosci Biobehav Rev*. 45:202-211.
8. Kohn N, Eickhoff SB, Scheller M, Laird AR, Fox PT, Habel U (2014): Neural network of cognitive emotion regulation--an ALE meta-analysis and MACM analysis. *Neuroimage*. 87:345-355.
9. Morawetz C, Bode S, Derntl B, Heekeren HR (2017): The effect of strategies, goals and stimulus material on the neural mechanisms of emotion regulation: A meta-analysis of fMRI studies. *Neurosci Biobehav Rev*. 72:111-128.
10. Ward BD (2000): Simultaneous inference for fMRI data.
11. Gee DG, Humphreys KL, Flannery J, Goff B, Telzer EH, Shapiro M, et al. (2013): A developmental shift from positive to negative connectivity in human amygdala-prefrontal circuitry. *J Neurosci*. 33:4584-4593.
12. Goldin PR, McRae K, Ramel W, Gross JJ (2008): The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol Psychiatry*. 63:577-586.
13. McRae K, Gross JJ, Weber J, Robertson ER, Sokol-Hessner P, Ray RD, et al. (2012): The development of emotion regulation: an fMRI study of cognitive reappraisal in children, adolescents and young adults. *Soc Cogn Affect Neurosci*. 7:11-22.
14. Silvers JA, Shu J, Hubbard AD, Weber J, Ochsner KN (2015): Concurrent and lasting effects of emotion regulation on amygdala response in adolescence and young adulthood. *Dev Sci*. 18:771-784.

15. Zilverstand A, Parvaz MA, Goldstein RZ (2017): Neuroimaging cognitive reappraisal in clinical populations to define neural targets for enhancing emotion regulation. A systematic review. *Neuroimage*. 151:105-116.

16. Ryali S, Supekar K, Chen T, Menon V (2011): Multivariate dynamical systems models for estimating causal interactions in fMRI. *Neuroimage*. 54:807-823.

17. Ryali S, Chen TW, Supekar K, Tu T, Kochalka J, Cai WD, et al. (2016): Multivariate dynamical systems-based estimation of causal brain interactions in fMRI: Group-level validation using benchmark data, neurophysiological models and human connectome project data. *J Neurosci Meth*. 268:142-153.

18. Ryali S, Shih YY, Chen T, Kochalka J, Albaugh D, Fang Z, et al. (2016): Combining optogenetic stimulation and fMRI to validate a multivariate dynamical systems model for estimating causal brain interactions. *Neuroimage*. 132:398-405.

19. Cai W, Chen T, Ryali S, Kochalka J, Li C-SR, Menon V (2015): Causal interactions within a frontal-cingulate-parietal network during cognitive control: Convergent evidence from a multisite–multitask investigation. *Cerebral Cortex*. 26:2140-2153.

20. Cho S, Metcalfe AW, Young CB, Ryali S, Geary DC, Menon V (2012): Hippocampal–prefrontal engagement and dynamic causal interactions in the maturation of children's fact retrieval. *Journal of cognitive neuroscience*. 24:1849-1866.

21. Supekar K, Menon V (2012): Developmental maturation of dynamic causal control signals in higher-order cognition: a neurocognitive network model. *PLoS Comput Biol*. 8:e1002374.

22. Martin RE, Ochsner KN (2016): The neuroscience of emotion regulation development: implications for education. *Current opinion in behavioral sciences*. 10:142-148.

23. Stephanou K, Davey CG, Kerestes R, Whittle S, Pujol J, Yücel M, et al. (2016): Brain functional correlates of emotion regulation across adolescence and young adulthood. *Human Brain Mapping*. 37:7-19.

24. Powers JP, LaBar KS (2018): Regulating emotion through distancing: a taxonomy, neurocognitive model, and supporting meta-analysis. *Neuroscience & Biobehavioral Reviews*.

25. Avram J, Baltes FR, Miclea M, Miu AC (2010): Frontal EEG activation asymmetry reflects cognitive biases in anxiety: evidence from an emotional face Stroop task. *Appl Psychophysiol Biofeedback*. 35:285-292.

26. Engels AS, Heller W, Mohanty A, Herrington JD, Banich MT, Webb AG, et al. (2007): Specificity of regional brain activity in anxiety types during emotion processing. *Psychophysiology*. 44:352-363.

27. Miskovic V, Schmidt LA (2010): Frontal brain electrical asymmetry and cardiac vagal tone predict biased attention to social threat. *Int J Psychophysiol*. 75:332-338.

28. Heller W (1993): Neuropsychological mechanisms of individual differences in emotion, personality, and arousal. *Neuropsychology*. 7:476.

29. Heller W, Nitschke JB, Miller GA (1998): Lateralization in emotion and emotional disorders. *Curr Dir Psychol Sci*. 7:26-32.

30. Miller GA, Crocker LD, Spielberg JM, Infantolino ZP, Heller W (2013): Issues in localization of brain function: The case of lateralized frontal cortex in cognition, emotion, and psychopathology. *Front Integr Neurosci*. 7:2.

31. Paes F, Machado S, Arias-Carrion O, Velasques B, Teixeira S, Budde H, et al. (2011): The value of repetitive transcranial magnetic stimulation (rTMS) for the treatment of anxiety disorders: an integrative review. *CNS Neurol Disord Drug Targets*. 10:610-620.

32. Plummer M (2003): JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd international workshop on distributed statistical computing*: Vienna, Austria., pp 10.

33. Gelman A, Carlin JB, Stern HS, Rubin D (2004): Bayesian data analysis 2nd edn Chapman & Hall. *CRC, Boca Raton FL*.