Neuron
Article

Prefrontal-Subcortical Pathways Mediating Successful Emotion Regulation

Tor D. Wager,[1,*] Matthew L. Davidson,[1] Brent L. Hughes,[1] Martin A. Lindquist,[2] and Kevin N. Ochsner[1]
[1]Department of Psychology
[2]Department of Statistics
Columbia University, New York, NY 10027, USA
*Correspondence: tor@psych.columbia.edu
DOI 10.1016/j.neuron.2008.09.006

SUMMARY

Although prefrontal cortex has been implicated in the cognitive regulation of emotion, the cortical-subcortical interactions that mediate this ability remain poorly understood. To address this issue, we identified a right ventrolateral prefrontal region (vlPFC) whose activity correlated with reduced negative emotional experience during cognitive reappraisal of aversive images. We then applied a pathway-mapping analysis on subcortical regions to locate mediators of the association between vlPFC activity and reappraisal success (i.e., reductions in reported emotion). Results identified two separable pathways that together explained ~50% of the reported variance in self-reported emotion: (1) a path through nucleus accumbens that predicted *greater* reappraisal success, and (2) a path through ventral amygdala that predicted *reduced* reappraisal success (i.e., more negative emotion). These results provide direct evidence that vlPFC is involved in both the generation and regulation of emotion through different subcortical pathways, suggesting a general role for this region in appraisal processes.

INTRODUCTION

If our emotions are woven into the fabric of human life, then our ability to regulate them keeps us from coming unraveled. In the best of circumstances, successful regulation leaves us feeling frayed around the edges. In the worst of circumstances, regulatory failures take a severe toll and contribute to the genesis and symptomatology of many psychiatric disorders (Davidson et al., 2000; Phillips et al., 2003).

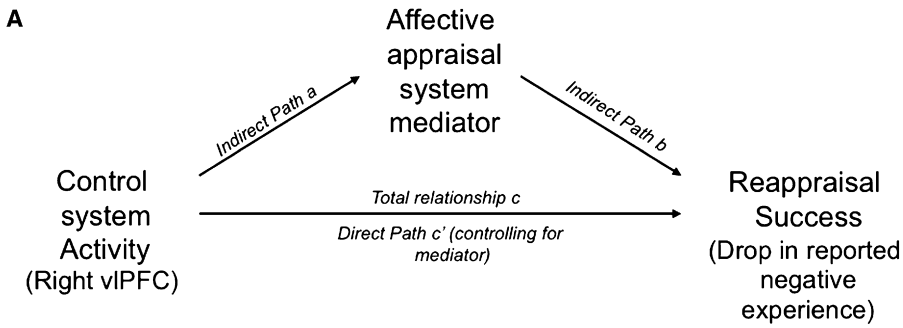
In the past few years, brain-based models of emotion regulation have been developed that can be extended to clinical contexts. These models have established the prefrontal cortex (PFC) as a key player in the cognitive regulation of emotion (Davidson, 2002; Ochsner and Gross, 2005). Numerous fMRI studies have observed increases in activity in the ventrolateral, dorsolateral, and dorsomedial prefrontal cortices (vlPFC, dlPFC, and dmPFC, respectively) when participants are instructed to deploy cognitive strategies that reduce negative emotional experience (Ochsner and Gross, 2005). Perhaps the most well studied such strategy is reappraisal, which involves reinterpreting the meaning of affective stimuli in ways that alter their emotional impact.

It is typically assumed that the beneficial effects of reappraisal are accomplished via interactions between PFC regions and subcortical networks related to emotional responding (Beauregard et al., 2001; Eippert et al., 2007; Goldin et al., 2007; Kalisch et al., 2006; Kim and Hamann, 2007; Ochsner et al., 2004; Phan et al., 2005; Urry et al., 2006; van Reekum et al., 2007). According to this *mediation hypothesis* (Figure 1A), PFC activity reduces negative emotion by influencing subcortical systems implicated in affective appraisal and learning processes, which in turn impact reported emotional experience.

Alternatively, the *direct pathway hypothesis* (Figure 1A) suggests that successful reappraisal is directly related to cortical activity and minimally impacts evolutionarily older subcortical systems. This hypothesis suggests that reappraisal—as a form of cognitive reinterpretation—primarily impacts cortical systems involved either in cognitive appraisals (Barrett et al., 2007; Ochsner and Barrett, 2001; Scherer, 2001) or in assessing changes in one's subjective emotional state (McRae et al., 2008b). Discriminating between the mediation and direct hypotheses is critical because it concerns how deeply reappraisal penetrates the emotional appraisal process.

The results of some prior studies are consistent with the mediation hypothesis, but none have tested the complete mediation pathway and instead have examined only separate portions of it. For example, prefrontal correlates of reappraisal success (Ochsner et al., 2002, 2004; Ohira et al., 2006; Phan et al., 2005; Urry et al., 2006), prefrontal-subcortical correlations (reviewed in Ochsner and Gross, 2005, 2008) and subcortical activity-experience correlations (Ochsner et al., 2004; Phan et al., 2005) have all been reported. In part because these results come from separate analyses, the PFC and subcortical regions involved differ from study to study, and no coherent mediation model has emerged. What is needed is a direct, formal test of whether subcortical regions of interest (ROIs) mediate the relationship between key PFC regions and reappraisal success.

In this paper, we tested alternative models of how PFC activity is related to reappraisal success. This relationship could be mediated by activity in subcortical ROIs, or it could be direct, without mediation through subcortical pathways. Testing these alternative models entailed identifying candidate PFC regions using a conventional statistical parametric mapping analysis of correlates of reappraisal success, and then testing for mediation of the PFC-reappraisal success relationship in subcortical regions.



B **Question 1:** Does activity in affective appraisal systems mediate the vlPFC-experience relationship?

Mediation Hypothesis:	Yes, paths *a*, *b*, and *a*b* are significant
Direct pathway Hypothesis:	No, only direct path *c'* is significant

Question 2: If there is mediation, what type of appraisal system does it involve?

Negative Mediation Hypothesis:	Systems preferential to negative affect, like the amygdala
Positive Mediation Hypothesis:	Systems preferential to positive affect, like the ventral striatum

Figure 1. Mediation Model and Hypotheses

We selected right ventrolateral prefrontal cortex (vlPFC) as a predictor region, and performed analyses to search in subcortical ROIs and throughout the brain for mediators of the relationship between reappraisal-induced vlPFC activity and reappraisal success.

(A) Path diagram with standard notation for path coefficients.

(B) Main hypotheses for mediation search analyses.

Assuming that the relationship between PFC activity and reported emotional experience is subcortically mediated, we hypothesized that systems involved in negative or positive affect—or both—could be mediators (Figure 1B). On one hand, we predicted that the PFC could influence activity in regions associated with negative affect, such as the amygdala and insula (Etkin and Wager, 2007; Wager et al., 2008). The amygdala is the affect-related region most commonly modulated during reappraisal (Ochsner and Gross, 2005, 2008); it has been implicated in detecting, attending to, and encoding into memory affectively arousing and especially threatening stimuli (Holland and Gallagher, 1999; Phelps, 2006; Whalen, 1998), and it is hyperactive across a range of anxiety disorders (Etkin and Wager, 2007).

On the other hand, we predicted that PFC could also influence activity in regions associated with positive affect, such as the nucleus accumbens (NAC) and nearby ventral striatum (VS). These regions have been consistently implicated in reward learning (McClure et al., 2003; Schultz, 2004), reward expectancies (Breiter et al., 2001; Cromwell and Schultz, 2003), and pain-diminishing placebo effects (Scott et al., 2007; Wager et al., 2007), as well as approach motivation more generally (Tindell et al., 2006; Wager et al., 2007). Striatal activity has been reported in reappraisal studies, although its functional role has not been clarified (McRae et al., 2008a; Ochsner et al., 2002, 2004; van Reekum et al., 2007).

Importantly, the PFC could be associated with activity in both kinds of cortical-subcortical pathways simultaneously. Because most studies have investigated brain-behavior relationships one brain region at a time, their results cannot be used to develop models in which two or more brain pathways make independent contributions to emotion regulation or any other behavior. Thus, in addition to providing direct tests of whether subcortical regions mediate the PFC-reappraisal success relationship, we developed an analysis strategy to parse the contributions of multiple brain pathways to a psychological measure of interest—in this case, reappraisal success.

To this end, we developed a procedure for locating multiple brain mediators, which we term Mediation Effect Parametric Mapping (MEPM). The approach is built on standard approaches from structural equation modeling and can be used to test specific hypotheses about brain pathway-behavior relationships. The use of mediation analysis to test relationships among variables has a long history (Baron and Kenny, 1986; Hyman, 1955; MacCorquodale and Meehl, 1948). What is new, however, is (1) the construction of statistical parametric maps of mediation tests, which permits mapping of multiple brain regions that satisfy the formal criteria for mediators, and (2) application of clustering methods to examine how multiple mediating regions are organized into networks. These developments allowed us to test whether multiple PFC-subcortical pathways make independent contributions to reappraisal success.

RESULTS

We performed a series of analyses designed to identify prefrontal-subcortical pathways linked to reappraisal success. We defined reappraisal success as the decrease in reported emotion when applying a cognitive reappraisal strategy to aversive images (ReappNeg trials) versus experiencing the natural emotional response to aversive images (LookNeg trials). Thus, reappraisal success scores were the [LookNeg − ReappNeg] values for each participant. We chose emotion self-reports as an outcome because they are stable and reliable, they predict

numerous mental and physical health outcomes (Gross and Munoz, 1995; Moskowitz, 2003; Tugade et al., 2004), and no other measure provides a direct correlate of emotional experience. We applied MEPM and multivariate clustering methods to relate reappraisal success with reappraisal-related increases in brain activity, which we defined as values for the [ReappNeg − LookNeg] contrast for each participant.

Analyses of emotion reports showed that aversive images evoked substantially stronger negative emotion reports (mean = 3.60 points, SE = 0.10) than neutral images [mean = 1.33, SE = 0.03; paired t test $t(35) = 22.3$, $p < 0.0001$]. Reappraisal of negative images resulted in robust reductions in emotion reports [ReappNeg mean = 2.65, SE = 0.10, an average reduction of 0.95 points, paired $t(35) = 8.83$, $p < 0.0001$; see Figure S2 available online and Experimental Procedures for additional details]. We refer to this reduction in reported emotion as reappraisal success in subsequent analyses.

fMRI analysis proceeded in a sequential series of steps. (1) First, we identified regions showing significant reappraisal-related activation. (2) Then, we used standard statistical parametric mapping to locate voxels from Step 1 in which reappraisal-related activation predicted reappraisal success. From among the resulting regions, we focused specifically on the right vlPFC (rvlPFC) as a predictor in subsequent analyses. (3) The next step was to conduct mediation analysis in subcortical ROIs. We defined ROIs in the amygdala and NAC/VS and tested whether voxels in each region mediated the vlPFC-reappraisal success relationship. (4) To locate additional regions, we used MEPM to generate a whole-brain map of mediators (4a), and then used cluster analysis to group these mediators into coherent networks and tested their independence from one another (4b). (5) Finally, we considered whether other candidate PFC regions showed similar relationships with reappraisal success mediated by NAC and amygdala.

1. Reappraisal-Related Activations

Results from the [ReappNeg − LookNeg] contrast are shown in Figure 2A and Table S1 (available online), controlling the expected false discovery rate (FDR) in whole-brain search at $q < 0.05$ (Genovese et al., 2002). Consistent with previous work, we observed increases (orange/yellow in Figure 2A) in a number of PFC regions. Dorsal and ventral PFC were active bilaterally, including the inferior frontal junction (IFJ), inferior frontal gyrus (IFG), middle frontal gyrus (Brodmann's Area [BA] 8/9), and anterior PFC (BA 10). Medial PFC regions included anterior cingulate (BA 24) and dmPFC in and anterior to pre-SMA (BA 9). Posterior cortical regions included the inferior parietal lobule (IPL), angular gyrus, and middle and inferior temporal gyri extending into anterior temporal cortex (aTC). Activations appeared to be more reliable on the left, though we did not perform detailed analysis of laterality effects. Decreases during reappraisal (blue in Figure 2A) were found in the parietal operculum, around SII, and in the parahippocampal cortex.

2. Identification of Potential Predictor Regions: Correlates of Reappraisal Success

Voxel-wise regression analyses identified eight regions in which reappraisal-related activation was significantly predicted by reappraisal analysis, shown in Figure 2B and reported in detail in Table 1 ($p < 0.005$ and $k = 3$ contiguous voxels; see Experimental Procedures for details). Only voxels identified in Step 1 above were subjected to analysis. Reappraisal-success-related regions included bilateral vlPFC (see also Figure 2C) and temporal regions, pre-SMA and dmPFC, left IPL, and right caudate. The cluster in rvlPFC, identified by its position on the IFG on the mean normalized anatomical image for our participants, was used as the predictor ROI in subsequent mediation analyses. Analysis steps 3 and 4a use this region as a frontal predictor region in MEPM analyses. Analysis steps 4b and 5 explore the relationships between rvlPFC and other cortical regions correlated with reappraisal success.

We chose vlPFC as the focus region for three reasons. First, vlPFC—and especially rvlPFC—is well known as a region critical for cognitive control in general, and response inhibition or the selection of information in particular (Aron and Poldrack, 2005; Badre and Wagner, 2005; Nee et al., 2007; Phillips et al., 2003). Second, studies of reappraisal and other cognitive means of regulating affective responses (Cunningham and Zelazo, 2007; Lieberman et al., 2007) have consistently activated rvlPFC. Third, activation measures in vlPFC have been shown to be abnormal in emotional disorders (Drevets et al., 1992; Lawrence et al., 2004). This choice was necessary to search for mediators, but we did not expect it to be the only important PFC region. As we describe below, subsequent analyses searched for additional PFC regions whose relationship with emotion was mediated by subcortical ROIs.

We also conducted a whole-brain exploratory search for correlations with reappraisal success to locate regions that may have been missed in our more constrained masked analysis ($q < 0.05$, FDR-corrected, $p < 0.0005$). The results largely confirmed those in our constrained search, including bilateral vlPFC and pre-SMA, but also identified additional regions shown in Figure S3, including the dorsal thalamus, caudate body, and caudate head extending into the NAC/VS; and parahippocampal cortex, cingulate isthmus (which bridges the posterior cingulate and medial temporal cortices), and area around periaqueductal gray (PAG). Whole-brain mediation and network analyses (Section 4a) identified some of these regions as significant mediators, and we explored their relationship to the key ROIs in our study in Section 4b below.

3. Testing the Amygdala and NAC/VS as Mediators Using A Priori ROIs

With the rvlPFC region identified as the predictor, and reappraisal success as the outcome (Figure 1), the next step in the MEPM analysis was to search, voxel by voxel, for mediators of the rvlPFC-reappraisal success relationship within two subcortical ROIs that were of a priori interest: the amygdala and NAC/VS.

For a variable to be considered a significant mediator, we required that it reach statistical significance in each of three tests (at $p < 0.005$ and three contiguous voxels in each), which were conducted as part of the same path model. First, the predictor variable (rvlPFC) must be related to the mediator (voxel in amygdala or NAC/VS), a link which we refer to in the text and all figures and tables as path *a*. Second, the mediator must be directly

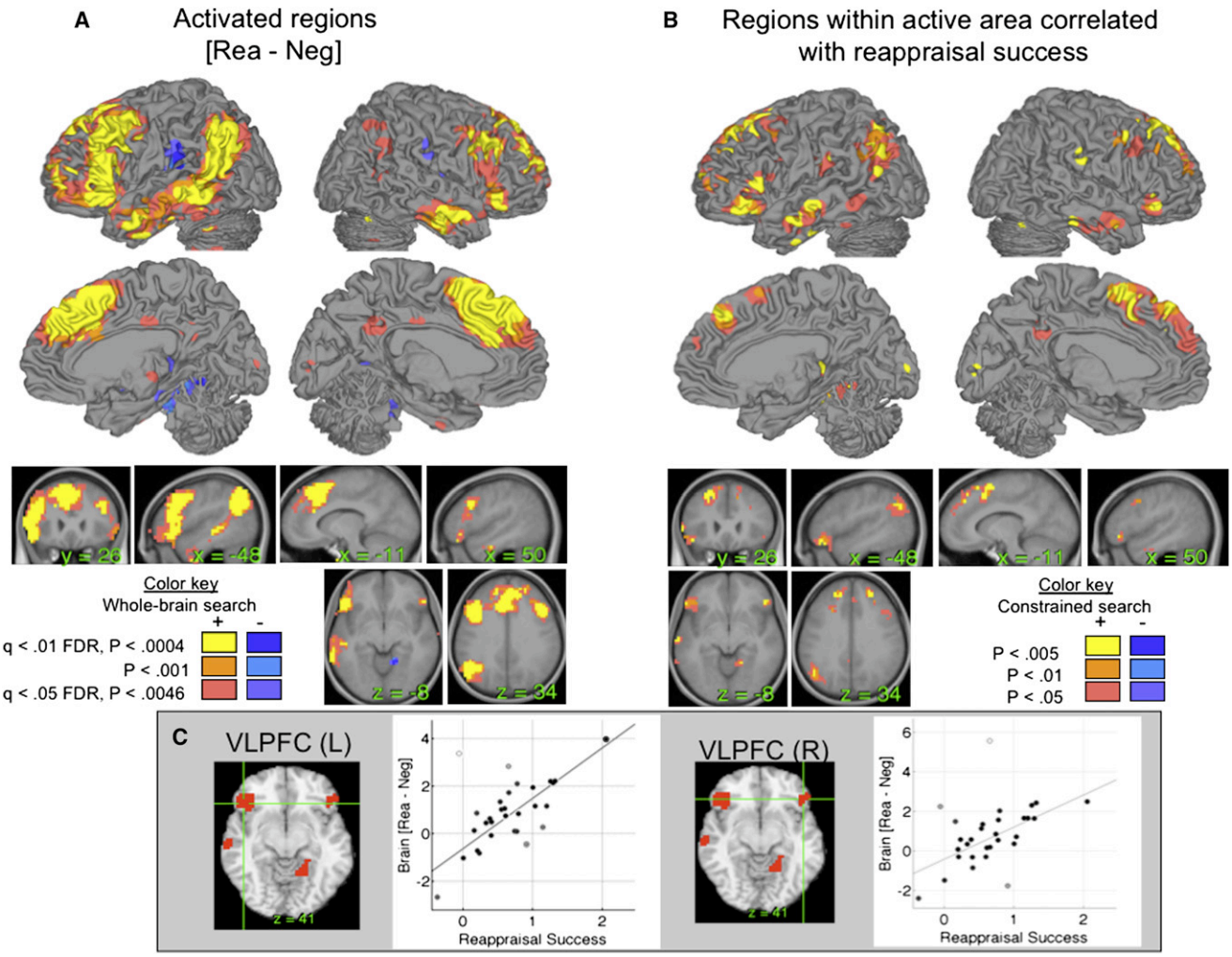


Figure 2. Results for the [ReappNeg – LookNeg] Contrast

This contrast was the main comparison of interest for the mediation analyses.

(A) Significant reappraisal-induced activation. Positive effects are in warm colors (yellow-red), and negative effects are in cool colors (blue-purple).

(B) Significant correlations between reappraisal-induced activation and reappraisal success, limited to a search area composed of activated regions in (A). Positive correlations are shown in red/yellow, and negative correlations are shown in blue. Positive correlations indicate a greater relative increase in activity for participants who report more successful reappraisal, and negative correlations indicate a greater relative decrease in activity for participants who report more successful reappraisal. Thresholds are shown in the color key on the figure.

(C) Correlation scatterplots for the average activity in left and rvlPFC. L, left; R, right.

related to the outcome (reappraisal success), controlling for vlPFC (path *b*). Finally, the mediation effect must be significant (effect *a*b*), which amounts to a statistical test on the product of the *a* and *b* path coefficients, or equivalently, a test that the predictor-outcome relationship (vlPFC-reappraisal success) is significantly reduced by including the mediator in the path model. The threshold of $p < 0.005$ and three contiguous voxels controlled the family-wise error rate at $p < 0.05$ corrected in each region (see Experimental Procedures).

According to standard conventions for mediation analysis, we refer to the overall predictor-outcome relationship as effect *c*, and the direct effect controlling for the mediator as *c′*. Thus, the *a*b* effect tests the significance of $c - c'$. In statistical reports and figures, *a* refers to the vlPFC-mediator relation, *b* refers to the direct mediator-reappraisal success relation (controlling for vlPFC), and *a*b* refers to the mediation effect.

For the amygdala analysis, we manually identified amygdala ROIs in gray matter on the average T1 image from our sample after warping to Montreal Neurologic Institute (MNI) space (left: xyz = [−24 0 −23], 5958 mm^3; right: xyz = [21 0 −22], 4354 mm^3 [Figure 3A; see Experimental Procedures]). MEPM analysis identified a portion of ventral left amygdala that negatively mediated the vlPFC-reappraisal success relation (xyz = [−24 0 −32], 319 mm^3). This portion of the left amygdala was positively associated with vlPFC, but predicted *reduced* reappraisal success: $a_{amy} = 0.57$, $Z = 4.02$, $b_{amy} = -0.24$, $Z = -5.01$, and $ab_{amy} = -0.14$, $Z = -3.30$, all $p < 0.001$. This finding suggests that the vlPFC plays a role in *generating* negative affective responses

Table 1. Correlations between Reappraisal Activation [Rea – Neu] and Reappraisal Success [Neu – Rea] in Reported Experience

	Coordinates			Cluster Size		Max Statistics				Significant Voxels		
Name	x	y	z	Vox.	Vol.	r	t	Z	p	p < 0.005	p < 0.01	p < 0.05
MTG (L)	−62	−3	−22	22	1170	0.60	3.56	3.22	0.0014	4	6	22
PHCP (R)	21	−48	−14	30	1595	0.69	4.47	3.85	0.0001	11	15	30
vlPFC (L)	−52	31	−9	87	4626	0.62	3.78	3.39	0.0008	19	35	87
MTG (L)	−65	−21	−14	16	851	0.62	3.54	3.21	0.0014	3	5	16
vlPFC (R)	52	31	−9	15	798	0.68	4.22	3.69	0.0002	4	5	15
CAU (R)	24	−34	9	9	479	0.63	3.85	3.44	0.0006	3	4	9
PreSMA/dmPFC	−14	31	45	456	24247	0.82	6.02	4.81	< 0.0001	113	197	456
IPL (L)	−48	−69	36	143	7604	0.59	3.42	3.13	0.0019	10	32	143

Results are reported for search only within regions that showed significant reappraisal in the group at q < 0.05 corrected (FDR). Significant regions within this reduced search space were thresholded at p < 0.005, three contiguous voxels. The extent of each region is listed at p < 0.01 and p < 0.05, uncorrected. CAU, caudate; dmPFC, dorsomedial prefrontal cortex; IPL, intraparietal sulcus; MTG, middle temporal gyrus; PHCP, parahippocampal cortex; Pre-SMA, pre-supplementary motor area; vlPFC, ventrolateral prefrontal cortex; Vox, voxels; Vol, volume in cubic mm.

through the amygdala, perhaps due to PFC-related appraisal processes that support the generation of negative responses to emotional pictures. Figure S4 shows path diagrams and scatterplots for path analyses in both amygdala and NAC/VS. In this and other figures showing ROI analyses, the region contiguous with the significant cluster is shown at p < 0.005 (dark blue), p < 0.01 (light blue), and p < 0.05 (purple; all two-tailed), including the extent beyond the ROI boundary.

For the NAC/VS analysis, we manually identified ROIs in NAC/VS gray matter adjacent to the posterior ventral caudate head (left: xyz = [−10 14 −9], 2765 mm^3; right: xyz = [7 14 −9], 2818 mm^3 [Figure 3A]). MEPM analysis identified a left NAC/VS region that *positively* mediated the vlPFC-reappraisal success relation (xyz = [−10 14 −14], 106 mm^3). The left NAC/VS was positively associated with both vlPFC and reappraisal success: $a_{NAC/VS}$ = 0.42, Z = 2.86, $b_{NAC/VS}$ = 0.29, Z = 3.20, and $ab_{NAC/VS}$ = 0.12,

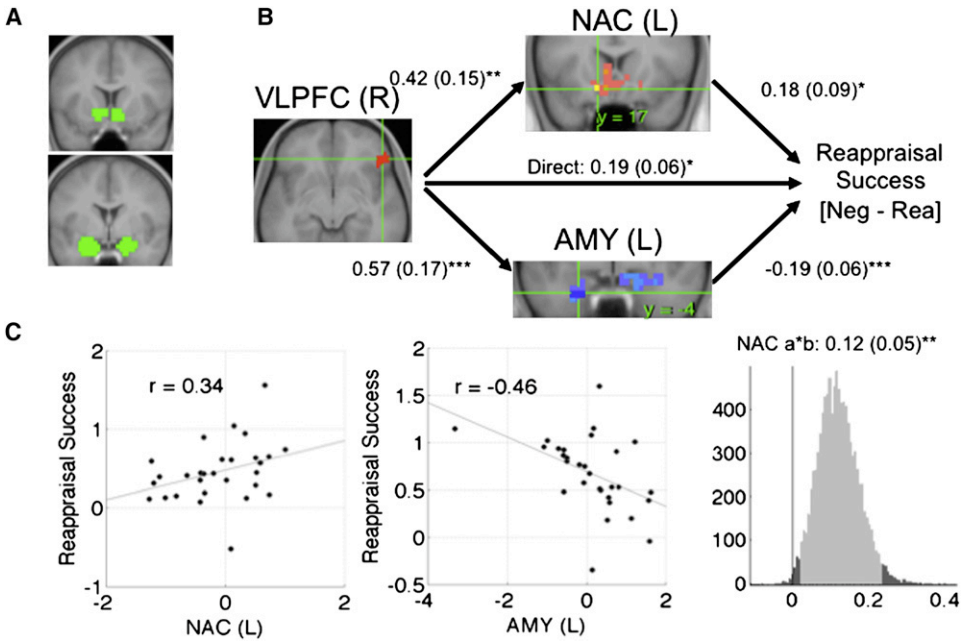


Figure 3. Mediation Analysis Results within ROIs

(A) ROIs in the nucleus accumbens (NAC, top) and amygdala (AMY, bottom).

(B) Path diagram showing the relationships between regions in the path model. The predictor region in vlPFC is shown at left, which predicts activity in each of the NAC (top) and amygdala (bottom) regions shown. These are the *a* paths for each mediating region. The lines are labeled with path coefficients, and standard errors are shown in parentheses. The mediator regions' (NAC and amygdala) connections to reappraisal success (the outcome) are the *b* paths for each mediator. They are calculated controlling for vlPFC activity and for other mediators, as is standard in mediation models. ***p < 0.001, **p < 0.01, *p < 0.05, two-tailed. The direct path is the *c′* path, and this is calculated controlling for both mediators.

(C) Partial regression scatterplots for the relationships between left (L) NAC and reappraisal success (left panel) and L amygdala and reappraisal success (center panel), controlling for vlPFC and the other mediator. The right panel shows an example of a bootstrapped mediation effect (path *ab*) for the left NAC. The range on the x axis spanned by the lighter gray portion of the histogram is the 95% confidence interval for the effect.

$Z = 3.04$, all $p < 0.002$. Figure S4 shows the significant region and contiguous voxels at $p < 0.005$ (yellow), $p < 0.01$ (orange), and $p < 0.05$ (pink; all two-tailed).

We then tested whether each region was a significant mediator of rvlPFC-reappraisal success correlations while controlling for effects attributable to the other region. We averaged over voxels in each region and included both amygdala and NAC/VS in the same path model. This analysis tests whether the results, shown in Figure 3, indicated that each subcortical pathway makes an independent contribution to reappraisal success. rvlPFC was positively associated with both amygdala (a_{amy} = 0.42, SE = 0.15, $p < 0.01$) and NAC/VS ($a_{NAC/VS}$ = 0.57, SE = 0.17, $p < 0.001$). Consistent with the individual mediation analyses, the two subcortical regions had opposite associations with reappraisal success: increasing amygdala activity was associated with reduced success ($b_{amy} = -0.19$, SE = 0.06, $p < 0.001$), whereas increasing NAC/VS activity was associated with increased success ($b_{NAC/VS}$ = 0.18, SE = 0.09, $p < 0.05$). Finally, both regions were significant mediators ($ab_{amy} = -0.11$, $Z = -2.83$, $p = 0.01$; $ab_{NAC/VS} = 0.07$, $Z = 2.31$, $p = 0.002$), indicating that each explains a part of the rvlPFC-reappraisal success covariance while controlling for effects attributable to the other mediator.

Because each pathway had an opposite effect on reappraisal success, this suggests that the rvlPFC-reappraisal success correlation is partially masked by their opposing effects. The correlation between vlPFC and reappraisal success was only moderately significant ($r = 0.43$, $p = 0.018$) without controlling for any additional regions, and was much stronger ($r = 0.66$, $p < 0.0001$) when controlling for amygdala activity.

These results demonstrate the advantage of the mediation approach over a simple bivariate (i.e., two-variable, no mediators or confounding variables) correlation approach. To more fully illustrate this advantage, we tested bivariate correlations with reappraisal success in the amygdala and NAC/VS. The left panel of Figure 4A shows the results in the amygdala ROIs (again including contiguous significant voxels extending outside the ROI), demonstrating significant negative correlations with reappraisal success (peak $r = -0.54$, $p < 0.005$, one-tailed; Figure 4D). However, a small subset of voxels show significant effects in the amygdala ROIs (one voxel at $p < 0.005$, three at $p < 0.01$, and five at $p < 0.05$, one-tailed). Figure 4B shows direct amygdala-reappraisal success associations, controlling for rvlPFC. Much stronger effects are apparent (32 voxels in the amygdala at $p < 0.005$, 44 at $p < 0.01$, and 78 at $p < 0.05$, two-tailed; Figure 4F). The NAC/VS showed strong positive correlations with reappraisal success (Figure 4E, $p < 0.0004$, whole-brain FDR corrected $q < 0.05$), which were somewhat reduced but still strongly significant after controlling for vlPFC (Figure 4F). Overall, the results confirm our prior expectations that amygdala and NAC/VS activation would show negative and positive brain-behavior correlations, respectively, and they demonstrate the value of the path modeling approach.

4. Identifying Mediating Networks Localized with MEPM Analysis

One important issue with the preceding results is that activity in other emotion-related regions might also play a mediating role in the vlPFC-reappraisal success relationship. In order to examine how the amygdala and NAC/VS fit into distributed networks that might underlie reappraisal success, we performed a whole-brain search for mediators of the rvlPFC-reappraisal success relationship. Then, significant regions from this mediation effect map were subjected to component analysis and clustering into interconnected networks. Finally, the average contrast values within each network were entered into a path model as mediating variables, with one mediator per network. This analysis tested whether multiple distributed networks make independent contributions to reappraisal success.

Results of the whole-brain search for mediators are shown in Figure 5A on medial and limbic surfaces (left and center panels, respectively) and on sagittal and coronal slices (right panel). These included several regions with positive indirect effects, including NAC/VS and neighboring subgenual cingulate and VS, pre-SMA, precuneus, superior frontal gyrus, dorsal caudate, and the cingulate isthmus. Regions showing negative indirect effects included bilateral amygdala, ventral anterior insula, rostral dorsal anterior cingulate (rdACC), and two regions in the subthalamic zone (near the ventral thalamus and hypothalamus). Stereotaxic coordinates and statistics for all mediators are shown in Table 2.

Cluster analysis combined with nonparametric inference (Etkin and Wager, 2007; see Supplemental Data for details) revealed evidence of two clusters, one consisting of the positive mediators (yellow in Figure 5B) including NAC/VS, and the other of the negative ones (blue in Figure 5B) including amygdala. These networks are shown in the context of the path model in Figure 5B. Regions in each network are shown on brain surfaces, with solid black lines connecting pairs of regions that are significantly correlated ($q < 0.05$, FDR-corrected) and whose correlation is not mediated by any single other region. Thus, grouping the set of positive mediators and the set of negative ones seemed to provide a sensible first-order grouping of the results, and further subdivision was not warranted by the data.

To assess the overall organization of these regions and their relationship with vlPFC activity and reappraisal success, we next averaged over voxels from all the regions within each network to obtain average [ReappNeg − LookNeg] contrast values for each network. We included the average scores from each network in the mediation model. This step was employed to avoid multicolinearity problems that make many multivariable SEM models inestimable or unstable and difficult to interpret. Path coefficients and standard errors are shown in Figure 5B. As before, both the positive and negative mediating networks were positively correlated with rvlPFC, but their direct effects on reappraisal success had opposite signs: the amygdala/insula/rdACC network (blue) was associated with reduced success, and the NAC/VS/cingulate isthmus/pre-SMA/precuneus network was associated with increased success. Scatterplots and estimates of the probability distribution of the indirect effects ($a*b$, see Experimental Procedures) are shown for each network in Figure 5C.

5. Search for Additional Prefrontal Predictor Regions with MEPM Analysis

In our previous analyses, we chose rvlPFC as a frontal predictor or "seed" region based on prior anatomical and functional

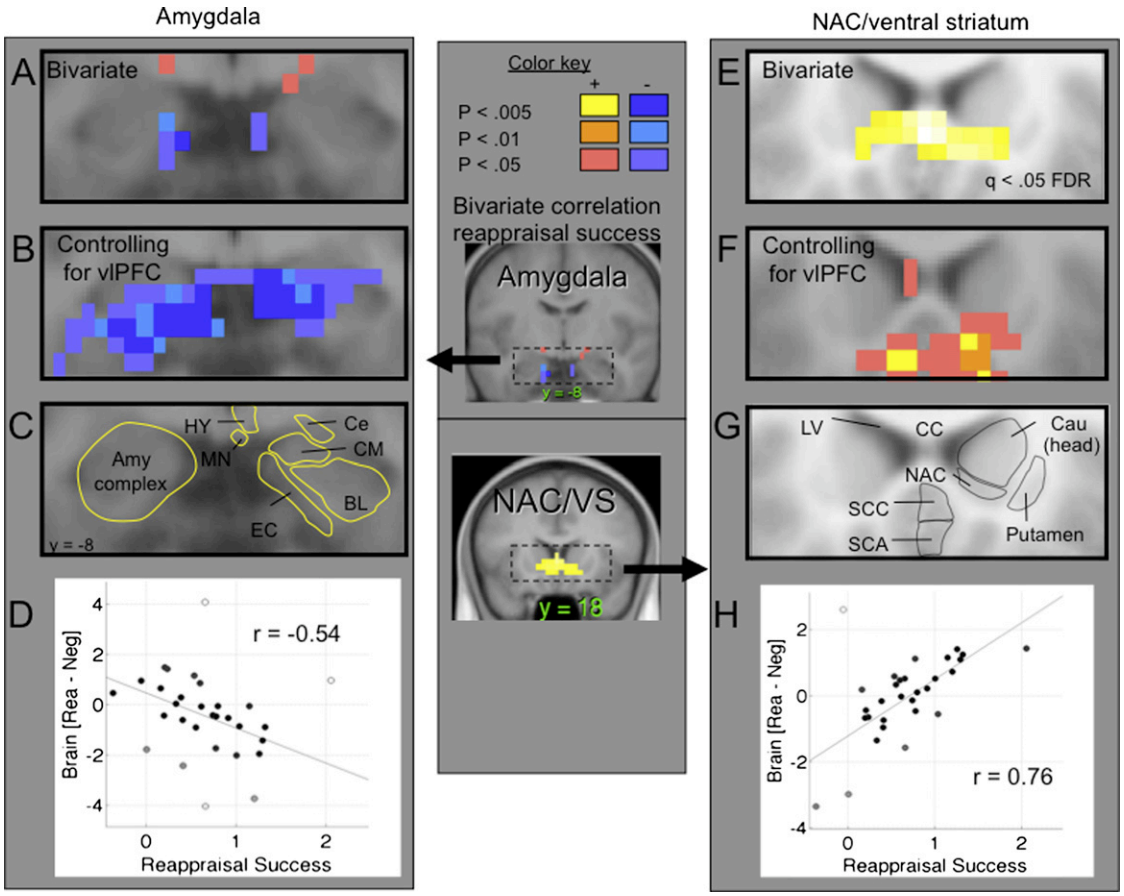


Figure 4. Detail of Effects in the Amygdala (Left) and NAC/Ventral Striatum (Right)

(A) Coronal slice (y = −5) showing reappraisal-induced activation in and contiguous with the amygdala ROI correlated with reappraisal success. Blue: negative correlations, showing a greater relative decrease in amygdala activity (LookNeg − ReappNeg) with more successful reappraisal. Red/yellow: positive correlations.

(B) Significant voxels in and contiguous with the amygdala showing a *b* effect in the mediation (the same relationship as in [A], but controlling for vlPFC activity). Results are two-tailed.

(C) The anatomical boundaries of amygdalar subregions. BL, basolateral complex; Ce, central nucleus; CM, cortomedian group; EC, entorhinal cortex; HY, hypothalamus; MN, mamillary nucleus. The strongest results appear to be in the entorhinal cortex bordering the basolateral and centromedian nuclear groups.

(D) Robust regression scatterplot showing the relationship between reported reappraisal success (x axis) and reappraisal-induced activation (relative to the mean) for the peak voxel in (A). More successful reappraisal was associated with relative reduction in amygdala activity (min r = −0.54, t = −2.82, Z = 2.61, p = 0.0043, one-tailed).

(E) Voxels in and contiguous with the NAC/ventral striatum (VS) ROI correlated with reappraisal success, as in (A). The display threshold is set at a false discovery rate of q < 0.05 for the whole brain.

(F) Significant voxels in and contiguous with the NAC/VS showing a *b* effect in the mediation.

(G) The anatomical boundaries of NAC and surrounding regions. Cau, caudate; CC, corpus callosum; LV, lateral ventrical; SCA, subcallosal area; SCC, subgenual cingulate cortex.

(H) Robust regression scatterplot showing relationship between reported reappraisal success (x axis) and reappraisal-induced activation (relative to the mean). More successful reappraisal was associated with relative increases in NAC/VS activity (mean r = 0.76, t = 5.50, p = 7×10^{-6}, two-tailed).

evidence. A natural question, however, is whether other frontal regions are similarly related to emotional experience through the cortical-subcortical pathways we describe. The MEPM analysis framework also provides a facility for searching for these mediated regions (i.e., regions that satisfy all requirements for mediated regions), given a chosen outcome and mediating pathway. We searched for frontal regions mediated by either left NAC/VS or left amygdala (averaging over voxels in each region) and controlling for the other (see Experimental Procedures for details). The results of each analysis are shown in Figure 6, and coordinates and statistics for each significant region are listed in Table S1. The regions include the frontal regions that are part of the network shown in Figure 5, including pre-SMA and superior frontal sulcus, as well as bilateral vlPFC and several other regions on the medial wall: medial frontal pole (MFP) and ventromedial PFC, including subgenual cingulate cortex (SGACC) and medial orbitofrontal cortex (OFC). Notably, the rostral medial PFC was also strongly, positively correlated with reappraisal

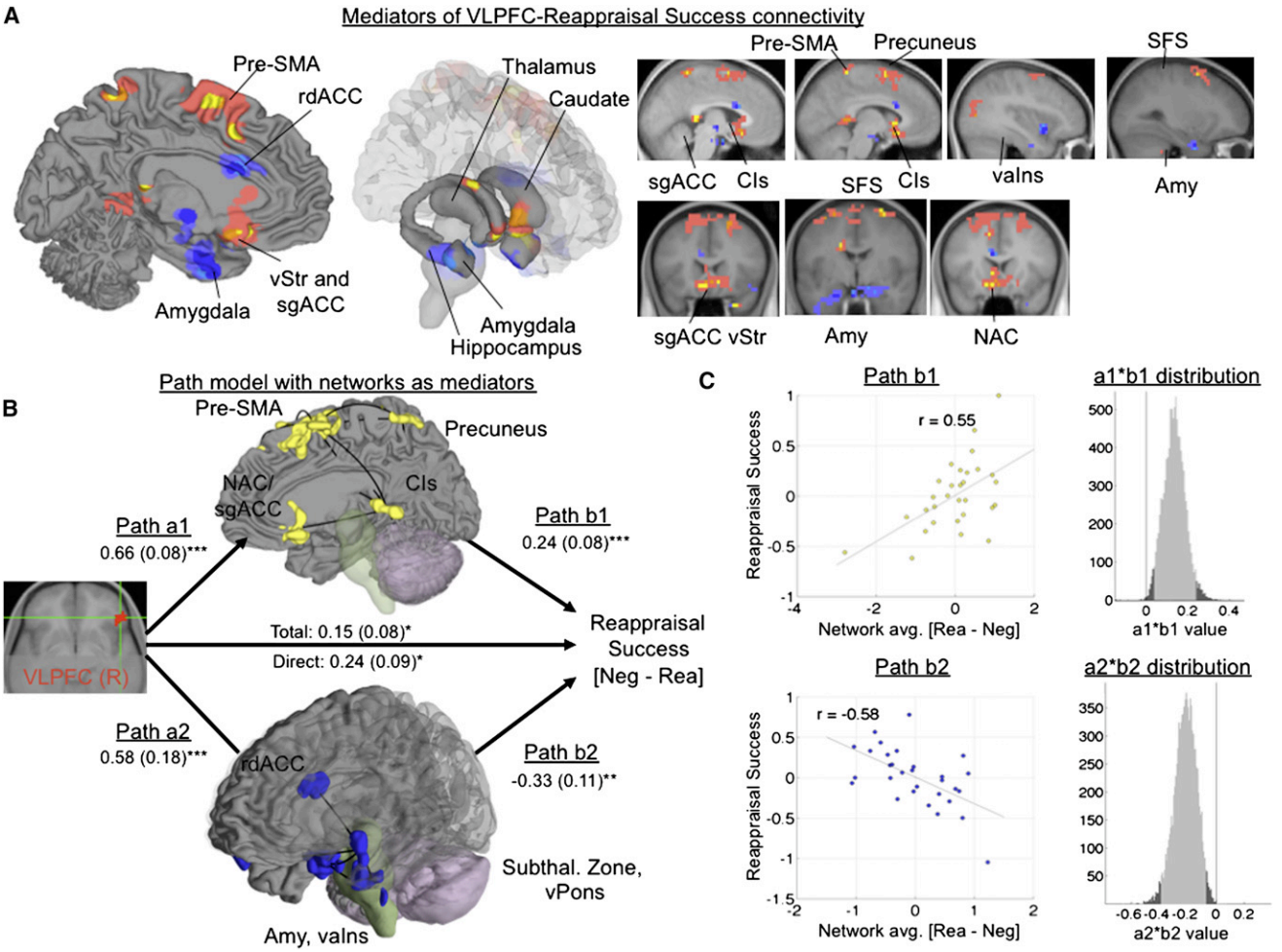


Figure 5. Mediation Results in Whole-Brain Search

(A) Significant regions ($p < 0.005$, two-tailed, and three contiguous voxels in each of paths *a*, *b*, and *a*b*) mediating the rvlPFC-reappraisal success relationship throughout the brain. Positive mediators are shown in yellow, and negative mediators are shown in dark blue. The extent around significant effects is shown at $p < 0.01$ (orange/light blue) and $p < 0.05$ (pink/purple).

(B) Network clustering of the mediating regions showed the strongest evidence for two functional networks, a group of positive mediators (yellow) and a group of negative mediators (blue). Lines show significant connectivity between regions that could not be explained by any other single mediating region. The path diagram shows the mediation model including activity in each of the two networks (averaged across voxels for each participant) as independent predictors. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, two-tailed.

(C) Partial regression scatterplots for the *b* effect for each network, and histograms showing the bootstrap distribution for each mediation (*a*b*) effect.

success in the bivariate correlation analysis (Figure S2), providing additional support for the involvement of the medial frontal cortex in successful downregulation of negative emotion. Thus, as expected, the rvlPFC is part of a set of frontal regions whose activation during reappraisal predicts drops in reported experience, to the degree that they coactivate with NAC/VS.

DISCUSSION

Previous studies of reappraisal of negative emotional stimuli have found activations in a number of prefrontal cortical areas, reductions of amygdala activity, and changes in reported emotion (Ochsner and Gross, 2005, 2008). However, the full-path model (Figure 1) linking these effects has not been tested. Furthermore, reappraisal success and related outcomes have typically been regressed against brain activity one region (or voxel) at a time, precluding the development of models in which multiple pathways make independent contributions. We addressed these issues by developing and applying MEPM, the voxel-by-voxel analysis of formal mediation effects in a path modeling framework. Specifically, we addressed (1) whether the prefrontal-emotional response relationship was direct or instead mediated via subcortical brain systems involved in emotional appraisal, and (2) whether multiple subcortical pathways might make separable contributions to successful reappraisal.

Our findings both replicated prior work and provided information that brain regions implicated in reappraisal are organized into at least two independent cortical-subcortical networks.

Table 2. Mediators of the rvlPFC Relationship with Reappraisal Success in Whole-Brain Search

		Coordinates			Correlation		*a* Path		*b* Path		*ab* Path		Conjunction (Num. Vox.)		
	Name	x	y	z	r	p	Z	p	Z	p	Z	p	p < 0.005	p < 0.01	p < 0.05
Positive mediators	ITG (R)	58	−38	−22	0.48	0.009	4.22	<0.001	3.22	0.001	3.15	0.002	5	7	41
	NAC/SGACC	3	21	−9	0.49	0.007	3.60	<0.001	3.25	0.001	3.34	0.001	4	9	85
	Isth Cing	−3	−41	4	0.48	0.009	4.02	<0.001	3.05	0.002	3.50	0.001	2	4	34
	Cau (R)	21	7	22	0.44	0.020	4.23	<0.001	2.89	0.004	3.13	0.002	1	2	6
	DMPFC (R)	24	41	40	0.52	0.004	4.53	<0.001	2.92	0.004	3.41	0.001	1	8	32
	Pre-SMA/SFG	−14	7	63	0.52	0.004	4.92	<0.001	3.79	<0.001	3.79	<0.001	9	16	159
	SFG (R)	24	21	58	0.44	0.020	3.81	<0.001	3.21	0.001	2.98	0.003	1	7	51
	Precuneus	−3	−52	68	0.50	0.005	4.50	<0.001	3.09	0.002	3.14	0.002	1	4	27
Negative mediators	AMY (L)	−28	−3	−32	−0.58	0.001	4.77	<0.001	−5.69	<0.001	−3.92	<0.001	6	7	42
	AMY (R)	14	−7	−22	−0.56	0.002	5.72	<0.001	−4.31	<0.001	−2.95	0.003	1	20	104
	LOFC (R)	48	24	−18	−0.47	0.010	5.20	<0.001	−3.13	0.002	−3.02	0.003	1	1	17
	VAINS (L)	−38	10	−14	−0.44	0.020	3.33	0.001	−3.30	0.001	−3.47	0.001	1	5	29
	Subthalamus	−3	−10	−4	−0.29	0.130	2.97	0.003	−3.00	0.003	−2.93	0.003	1	1	3
	RACC	−7	17	22	−0.37	0.050	3.28	0.001	−3.46	0.001	−3.29	0.001	2	4	18

These results complement focused searches within a priori subcortical ROIs. For this exploratory analysis, we considered voxels to be significant mediators if statistical significance reached $p < 0.005$ in each of the three following effects. *a* path: correlation between vlPFC and voxel activity; *b* path: correlation between voxel and reported emotional experience, controlling for vlPFC; and *a*b* path: significant indirect effect, whose magnitude is the product of *a* and *b*. All statistics were assessed with bootstrapping (1000 samples) at each voxel. These regions were subjected to subsequent multivariate clustering analysis for grouping into networks. Correlation refers to correlation with reported experience. AMY, amygdala; ITG, inferior temporal gyrus; Isth Cing, cingulate isthmus; LOFC, lateral orbitofrontal cortex; NAC, nucleus accumbens; RACC, rostral anterior cingulate; SGACC, subgenual anterior cingulate; SFG, superior frontal gyrus; VAINS, ventral anterior insula. Other abbreviations are as in Table 1.

Replicated findings include: (1) reappraisal-related increases in multiple frontal, parietal, and temporal regions (reviewed in Ochsner and Gross, 2005), (2) decreases in the amygdala correlated with self-reported negative affect, and (3) correlations of prefrontal activation with reappraisal success.

Of the eight brain regions both activated and correlated with reappraisal success, the rvlPFC was of particular interest, because of its association in previous studies with reappraisal and other cognitive forms of emotion regulation (Cunningham and Zelazo, 2007; Kim and Hamann, 2007; Lieberman et al., 2007; Ochsner et al., 2004) and selection and inhibition of information in general (Aron and Poldrack, 2005; Nee et al., 2007), and its dysfunction in clinical populations (Phillips et al., 2003). Mediation analyses showed that reappraisal-induced activation in this region was positively associated with activity in both amygdala and NAC/VS, and that both of these regions independently mediated the relationship between vlPFC and reported reappraisal success. Successful regulation was related to both increases in the NAC/VS pathway and decreases in the amygdala pathway.

Implications of Dual Routes for Reappraisal Effects on Emotion

The finding that two independent networks mediated the impact of PFC on emotional experience suggests that reappraisal success can be understood in terms of how PFC controls the nature and relative balance of negative and positive appraisals of a given stimulus. This interpretation is supported by our findings that PFC activity is associated with at least two subcortical pathways, one including NAC/VS and the other including the amygdala. Though NAC/VS and amygdala activity are not expected to be pure markers of positive and negative emotion (Cromwell and Schultz, 2003; Jensen et al., 2003; Levita et al., 2002; Paton et al., 2006), recent meta-analyses suggest that activity in NAC/VS and amygdala is biased toward positive and negative affective experience, respectively (Etkin and Wager, 2007; Wager et al., 2008). Thus, a plausible interpretation of our mediation results is that two processes are engaged by vlPFC: a negative appraisal process that leads to the generation of negative emotional responses to the stimuli and involves the amygdala, and a positive appraisal process that leads to effective positive reappraisal and involves the NAC/VS. Successful reappraisal therefore involves both dampening activity in the PFC-amygdala pathway and increasing activity in the PFC-NAC/VS pathway.

The idea that the NAC/VS is involved in cognitive emotion change is, to our knowledge, a new development in understanding emotion regulation. Although striatal activity has been implicated in cognitive reappraisal of negative emotion, a functional role for this region has not previously been identified (McRae et al., 2008a; Ochsner et al., 2002, 2004; van Reekum et al., 2007).

Also of interest is that PFC activity can be related to negative appraisal processes and amygdala increases. This diverges from some findings in other studies of emotion regulation, which have reported negative ventromedial PFC-amygdala connectivity (Johnstone et al., 2007; Urry et al., 2006). These differences may relate to differences in the regulatory strategies employed. For example, one kind of reappraisal accentuates positive

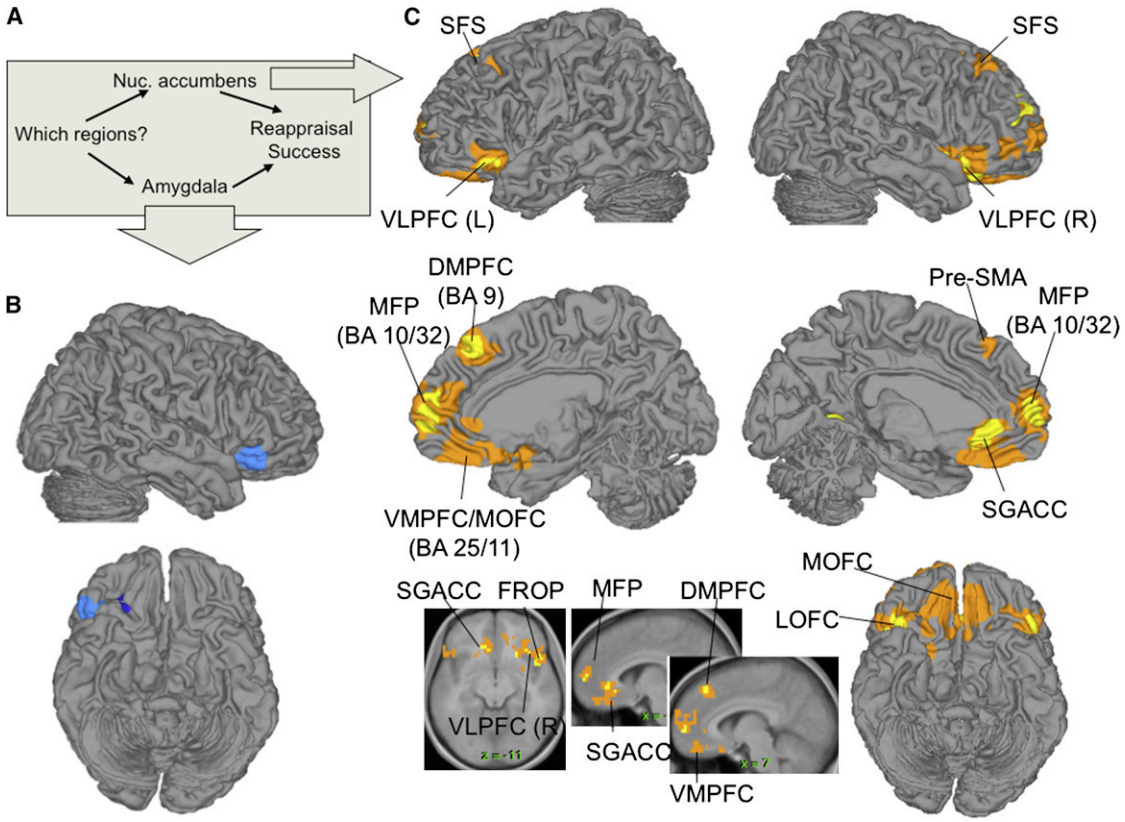


Figure 6. Frontal Regions Mediated by NAC/VS and Amygdala

(A) Diagram of the mediation model, in which both left NAC/VS and left amygdala regions were included as mediators of reappraisal success. The analysis searched over brain voxels for areas showing a significant mediation (*a*b*) effect.

(B) Regions mediated by amygdala.

(C) Regions mediated by NAC/VS. Yellow/dark blue: positive/negative effects at $p < 0.001$, three contiguous voxels; orange/light blue: positive/negative effects at $p < 0.005$, ten contiguous voxels. BA, Brodmann's Area; DMPFC, dorsomedial prefrontal cortex; FROP, frontal operculum; MFP, medial frontal pole; MOFC, medial orbitofrontal cortex; SGACC, subgenual anterior cingulate; SFS, superior frontal sulcus; Pre-SMA, presupplementary motor area; VLPFC, ventrolateral prefrontal cortex; VMPFC, ventromedial prefrontal cortex.

potential interpretations, as when one imagines that a sick person might soon be well. Another type of reappraisal might blunt the negativity of a stimulus, as when one imagines that an image of a mutilated body was taken on a movie set rather than at the scene of an accident. Finally, a third type might involve dissociation or distancing from the contents of a stimulus. Here, it is noteworthy that our reappraisal training focused largely on generating *positive* reappraisals. By contrast, prior reappraisal instructions have either given equal emphasis to positive-generation and negative-blunting appraisals (Johnstone et al., 2007; Ochsner et al., 2002, 2004; Phan et al., 2005; Urry et al., 2006) or have emphasized distancing and detachment (Eippert et al., 2007; Kalisch et al., 2005). Linking specific cognitive strategies to separable mediators of regulatory effects and their effects on measures of emotional responding (see Supplemental Data for extended discussion) will be important for future research.

Implications for the Role of PFC in Emotional Appraisal

The finding of both negative and positive appraisal-related pathways also has implications for the role that PFC plays in the generation and regulation of emotion. It is commonly assumed that frontal activation during reappraisal reflects "cool" emotion-dampening processes (Metcalfe and Mischel, 1999). An alternative view is that PFC is involved in the appraisal process itself, with the type of appraisal (positive or negative) determined by the specific cognitions involved (Lazarus, 1991). Consider that bilateral vlPFC is activated by both positive and negative emotion-induction tasks (often in tandem with subcortical activation; Kober et al., 2008), that it shows heightened activity in major depression (Drevets et al., 1992) that decreases with successful treatment (Brody et al., 1999, 2001; Goldapple et al., 2004), and that its activity increases when both decreasing and increasing negative emotion via reappraisal (Eippert et al., 2007; Kim and Hamann, 2007; Ochsner et al., 2004). This all supports an "appraisal-process" view of the vlPFC. The present findings of positive associations between vlPFC and both amygdala and NAC/VS activity, with opposite effects on reappraisal success, suggest that the information processing performed by the vlPFC (perhaps related to the selection among elements of the scene and/or different possible affective interpretations) both

generated negative emotion and mitigated it during reappraisal (for additional discussion, see Supplemental Data).

This prefrontal appraisal-process hypothesis helps explain why correlations between PFC and reappraisal success may not have been more commonly reported. If the PFC is associated with both positive and negative appraisals, the two opposing effects can effectively cancel each other out, resulting in weak overall PFC-reappraisal connections. This is known in the path analysis literature as a suppression effect (MacKinnon et al., 2000). Our findings are consistent with the view that a negative appraisal-related PFC-amygdala pathway suppresses an otherwise strong positive PFC-reappraisal success relationship. The appraisal-process hypothesis also helps understand why few studies have reported correlations between the magnitudes of self-reported emotion and amygdala activity (Abercrombie et al., 1998; Ochsner et al., 2004; Phan et al., 2005; Schaefer et al., 2002). Here we showed that a strong negative relationship between amygdala and reported emotion was masked by a positive influence on both amygdala and reported emotion from vlPFC, through another pathway (the NAC/VS pathway). This positive influence from PFC effectively obscured the negative amygdala-emotion experience relationship, which was significant but could not be detected until the influence of PFC was removed.

Implications of Network Analyses for Understanding Cognition-Emotion Interactions

It has been increasingly recognized that higher-order cognitive processes like appraisal arise from information processing within large-scale distributed networks (Cole and Schneider, 2007; Kober et al., 2008; Schmitz and Johnson, 2007). The present paper supported this view in two ways. Our whole-brain MEPM revealed that the amygdala and NAC/VS were parts of two larger networks (Figure 5B) that appear to operate in a partially independent manner to support cognitive reappraisal. The first network involved the amygdala and other regions associated with negative emotion, including the lateral OFC, anterior insula, and rdACC (see the Supplemental Data for an extended discussion). The second network involved the NAC/VS and regions implicated in action selection and memory, including the pre-SMA, precuneus, and subgenual and retrosplenial cingulate cortices. The positive association of vlPFC with activity in these two networks is consistent with the idea that that reappraisal success involves both limiting negative appraisals and generating positive appraisal by retrieving appropriate information from memory.

The notion that psychological processes typically map onto distributed systems also suggests that the vlPFC may be part of a larger distributed cortical network involved in cognitive construal. Here, we found evidence supporting that hypothesis. We conducted additional MEPM analyses that specified the amygdala and NAC/VS as mediators and searched for other frontal predictor regions whose impact on reappraisal success was mediated by these subcortical regions. This analysis identified other frontal regions that show the same relationship with subcortical pathways as rvlPFC (Figure 6), including left vlPFC, superior frontal sulcus, and regions of medial PFC associated with amygdala reductions in other studies (Johnstone et al., 2007; Urry et al., 2006). Though it was necessary for the MEPM analysis to specify a single seed region—the rvlPFC—these additional analyses provide evidence for distributed prefrontal control networks.

Alternative Models for PFC-Subcortical-Reappraisal Interactions

Although this study advances our knowledge of the regions mediating PFC-emotion relationships, it is limited in ways that may fruitfully be addressed in future studies. A chief limitation is that we estimated the mediation effect on the basis of naturally occurring variance over subjects. This is suboptimal since we do not have experimental control over this variability, but under the standard assumptions that all our subjects have the same functional anatomy and that intersubject differences do not affect the coupling between dependent variables, we can interpret our mediation effects in terms of the functional pathways tested here.

However, because we did not experimentally manipulate the brain, alternative path models may also be consistent with the data—particularly those that involve different patterns of causal relationships. For example, amygdala activity related to negative emotional appraisals could perhaps signal the need for regulatory control and trigger activation of the PFC, which in turn activates the negative emotion-reducing NAC/VS pathway. We think this unlikely, however, because the largest projections are from PFC to amygdala rather than vice versa (Pitkanen, 2000), and there is direct evidence that PFC plays an important role in modulating amygdala responses (Davis, 1992; Milad and Quirk, 2002), whereas we know of no direct evidence that supports the idea that the amygdala regulates the PFC.

Another alternative involves common influences of unmodeled variables on the variables in the mediation model, which can cause variables to be correlated even without a direct functional anatomical link between the two regions. For example, endogenous variation in diffuse neuromodulatory systems such as dopamine, opioids, etc. could affect both PFC and NAC/VS activity in a similar way, causing activity measures in these regions to be correlated. Future work may serve to address these issues directly by combining the path modeling methods we present here with direct experimental manipulation of the PFC and/or amygdala through brain stimulation, which would allow a stronger basis for drawing causal inferences.

Future studies could also expand on the work we present here by investigating how pathway strength is moderated by other variables, such as clinical diagnosis, personality, genetic differences (Munafo et al., 2008), and sex (McRae et al., 2008a). We found that the pathways we report were not moderated by sex (data not shown), but a complete test requires studies designed specifically to target this issue. Likewise, future work could focus on manipulating other aspects of reappraisal, such as the use of different strategies (e.g., cognitive reappraisal versus distraction or self-distancing) and the regulation of different kinds of responses (e.g., to aversive versus appetitive stimuli), which have begun to be addressed in other papers (Eippert et al., 2007; Goldin et al., 2007; Johnstone et al., 2007; Kim and Hamann, 2007; Ochsner et al., 2004; van Reekum et al., 2007).

Conclusions

In sum, this study shows evidence for a distributed set of lateral frontal, medial frontal, and orbitofrontal regions that together orchestrate reappraisal of the meaning of emotional events. Our results suggest that there are at least two separable pathways linking prefrontal cortical activity with reductions in negative emotion during reappraisal: one through the NAC/VS, which may generate positive appraisals, and one through amygdala, which may generate or enhance negative appraisals. These results suggest that it may be important to consider each of these functional pathways when examining the role of frontal-subcortical systems in emotion.

EXPERIMENTAL PROCEDURES

Participants

Thirty healthy right-handed participants (median age = 22.3 years; eighteen female) were recruited in compliance with the human subjects regulations of Columbia University and paid US$20/hour. Handedness was assessed with the Edinburgh Handedness Inventory, and eligibility was assessed with a general health questionnaire and fMRI safety screening form.

Materials and Procedures

Stimuli and Task Conditions

Forty-eight aversive images were selected from the International Affective Picture Set (IAPS; Lang et al., 1993; mean normative valence = 2.24, mean normative arousal = 6.28). The images were a subset of those used in Ochsner et al. (2004). Twenty-four neutral images were also selected from the set (valence = 5.27, arousal = 3.51). An additional set of 18 aversive images and 7 neutral images was used during a training session. These images were presented during fMRI scanning (~12° visual angle), and participants viewed each image only once.

Images were grouped into three conditions. In the LookNeu condition, participants viewed neutral images. In the LookNeg condition, participants viewed aversive images. In both conditions, participants were asked to view the image, understand its content, and allow themselves to experience/feel any emotional response it might elicit. In the ReappNeg condition, participants viewed aversive images, and were asked to reappraise the emotional value of those images so that the emotional impact was less negative. More specifically, they were instructed to generate a positive interpretation of the scene depicted in each picture that reduced the emotional impact, as in previous published work from our laboratory (Ochsner et al., 2002, 2004). A comprehensive prescanning training procedure was used to assure that participants understood the cue-task associations and the reappraisal strategy (see Supplemental Data). The assignment of negative images to conditions was randomized and counterbalanced across subjects.

Before presentation of each image, participants viewed a cue that signaled both the image type (aversive or neutral) and the instruction type (Look or Reappraise). Cues were white shapes—a circle, a square, and a triangle (~0.5° visual angle; see Figure S1)—presented on a black background. The assignment of shape to condition was counterbalanced across participants.

fMRI Task Design

Previous studies of reappraisal have not separated brain activity related to anticipation and instruction processing, stimulus viewing, and picture rating, and a goal of our task design was to provide the ability to estimate separately brain activation magnitude related to each of these three phases of the image viewing and rating process. To accomplish this, a partial trial design was employed (Ollinger et al., 2001; Stern et al., 2007). Within each task condition, LookNeu, LookNeg, and ReappNeg, three different trial types were used: full trials, anticipation-only trials, and stimulus-only trials (see Figure S1).

On full trials, a 2 s condition cue was followed by a 4 s anticipatory interval during which a fixation cross was presented on the screen. The image was subsequently presented for 8 s. Following image presentation, a fixation cross was presented during a 4 or 7 s jittered interstimulus interval (ISI; uniform distribution of 4 and 7 s intervals). Following the ISI period, the words "how negative do you feel?" appeared onscreen for 2.1 s, and participants were asked to rate negative affect on a five-point scale by pressing a button with one of five fingers on a button-response unit (1 = "not at all negative," indicated by a thumb button press, up to 5 = "extremely negative," indicated by a fifth-finger button press). Following the rating, a 4 or 7 s jittered ISI concluded the trial.

The anticipation-only trials consisted of cue, anticipatory, and rating intervals. The stimulus-only trials were identical to the full trials, except that the 4 s anticipation interval was omitted. This design allowed us to construct predictors for Cue-, Anticipation-, and Image-related brain activity related to each task condition in the General Linear Model (GLM) that were uncorrelated enough to provide efficient estimates of activation in each phase of the trial for each emotion condition (see Supplemental Data). Negative emotion ratings did not differ between full and stimulus-only trials (see Supplemental Data).

Data Acquisition and Analysis

For space reasons, fMRI acquisition and analysis procedures are described only briefly here. Details on all procedures are provided in the Supplemental Data.

Whole-brain fMRI data were acquired on a 1.5T GE Signa Twin Speed Excite HD scanner (GE Medical Systems). Functional and anatomical images were acquired with a T2*-sensitive EPI BOLD sequence with a TR of 2000 ms, TE of 40 ms, flip angle of 60°, field of view of 22 cm, 24 slices, and 3.44 × 3.44 × 4.5 mm voxels. Stimulus presentation and data acquisition were controlled using E-Prime software (PST Inc.). An LCD projector displayed stimuli on a back-projection screen mounted in the scanner suite. Responses were made with the right hand via a five-finger button-response unit with a molded hand brace (Avotec Inc. and Resonance Technologies).

Functional images were subjected to standard preprocessing using FSL (FMRIB Centre, University of Oxford) and SPM2 (Wellcome Department of Cognitive Neurology, UCL) software, and first-level (within-participant) statistical analysis using SPM2. Briefly, separate regressors in the GLM were specified for fMRI responses to the cue, anticipation, stimulus viewing, and rating response periods. Values for the [ReappNeg image viewing − LookNeg image viewing] contrast were subjected to second-level robust random effects analysis (Wager et al., 2005) to localize regions activated during reappraisal and regions in which activity correlated with reappraisal success.

The MEPM analysis is based on a standard three-variable path model (Baron and Kenny, 1986) with a bootstrap test for the statistical significance of the product *a*b* (Efron and Tibshirani, 1993; Shrout and Bolger, 2002), as diagrammed in Figure 1. MEPM analysis was conducted on [ReappNeg − LookNeg] contrast values. For ROI analyses in amygdala and NAC, the primary chosen threshold of $p < 0.005$ controlled the false positive rate below $p < 0.05$ corrected in each ROI.

The same threshold was used for exploratory whole-brain mediation analyses, which located multiple candidate mediators of the vlPFC-reappraisal success relationship. After locating multiple regions that were positive and negative mediators of the vlPFC-reappraisal success relationship, we sought to identify functional networks of interconnected regions. We adopted a procedure of principle components analysis-based dimension reduction followed by clustering of regions into functional networks used in several previous papers (Kober et al., 2008; Wager et al., 2007). This procedure used a nonparametric permutation test to assess whether there is significant grouping of regions into clusters (networks), and to identify brain regions that belong to each network. The null hypothesis of no interregion clustering was rejected, and a two-cluster solution provided the best fit, suggesting that there may be at least two independent functional networks related to reappraisal success. To test whether these networks were independent mediators of reappraisal success, we calculated average fMRI contrast values across voxels belonging to each network for each participant, and entered average values for each of the two networks as mediators in our path models.

In a final analysis, we used the MEPM approach to localize frontal regions whose relationship with reappraisal success was mediated by subcortical activity. We performed voxel-wise searches for frontal regions whose [ReappNeg − LookNeg] contrast values were mediated (significant *a*b* effect) by average contrast values over voxels in the amygdala or NAC/VS.

SUPPLEMENTAL DATA

The Supplemental Data for this article include one table, five figures, a Supplemental Discussion, and Supplemental Experimental Procedures and can be found online at http://www.neuron.org/cgi/content/full/59/6/1037/DC1/.

ACKNOWLEDGMENTS

We would like to thank Niall Bolger for helpful discussion on path analysis and the authors of SPM software for making it freely available. This paper was made possible with the support of NIH Grant MH076137 (K.O.), NIH Grant MH076136 (T.W.), and NSF 0631637 (T.W.) Author contributions are as follows: design, T.W. and K.O.; data collection, B.H.; data analysis, T.W., B.H., M.D., and M.L.; and writing, T.W., K.O., and B.H. Matlab code implementing mediation analyses is freely available at http://www.columbia.edu/cu/psychology/tor/.

Accepted: September 8, 2008
Published: September 24, 2008

REFERENCES

Abercrombie, H.C., Schaefer, S.M., Larson, C.L., Oakes, T.R., Lindgren, K.A., Holden, J.E., Perlman, S.B., Turski, P.A., Krahn, D.D., Benca, R.M., and Davidson, R.J. (1998). Metabolic rate in the right amygdala predicts negative affect in depressed patients. Neuroreport *9*, 3301–3307.

Aron, A.R., and Poldrack, R.A. (2005). The cognitive neuroscience of response inhibition: relevance for genetic research in attention-deficit/hyperactivity disorder. Biol. Psychiatry *57*, 1285–1292.

Badre, D., and Wagner, A.D. (2005). Frontal lobe mechanisms that resolve proactive interference. Cereb. Cortex *15*, 2003–2012.

Baron, R.M., and Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J. Pers. Soc. Psychol. *51*, 1173–1182.

Barrett, L.F., Mesquita, B., Ochsner, K.N., and Gross, J.J. (2007). The Experience of Emotion. Annu. Rev. Psychol. *58*, 373–403.

Beauregard, M., Levesque, J., and Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. J. Neurosci. *21*, RC165.

Breiter, H.C., Aharon, I., Kahneman, D., Dale, A., and Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. Neuron *30*, 619–639.

Brody, A.L., Saxena, S., Silverman, D.H.S., Alborzian, S., Fairbanks, L.A., Phelps, M.E., Huang, S.C., Wu, H.M., Maidment, K., and Baxter, L.R. (1999). Brain metabolic changes in major depressive disorder from pre-to post-treatment with paroxetine. Psychiatry Res. *91*, 127–139.

Brody, A.L., Saxena, S., Stoessel, P., Gillies, L.A., Fairbanks, L.A., Alborzian, S., Phelps, M.E., Huang, S.C., Wu, H.M., and Ho, M.L. (2001). Regional brain metabolic changes in patients with major depression treated with either paroxetine or interpersonal therapy: preliminary findings. Arch. Gen. Psychiatry *58*, 631–640.

Cole, M.W., and Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. Neuroimage *37*, 343–360.

Cromwell, H.C., and Schultz, W. (2003). Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. J. Neurophysiol. *89*, 2823–2838.

Cunningham, W.A., and Zelazo, P.D. (2007). Attitudes and evaluations: a social cognitive neuroscience perspective. Trends Cogn. Sci. *11*, 97–104.

Davidson, R.J. (2002). Anxiety and affective style: role of prefrontal cortex and amygdala. Biol. Psychiatry *51*, 68–80.

Davidson, R.J., Putnam, K.M., and Larson, C.L. (2000). Dysfunction in the neural circuitry of emotion regulation–a possible prelude to violence. Science *289*, 591–594.

Davis, M. (1992). The role of the amygdala in fear and anxiety. Annu. Rev. Neurosci. *15*, 353–375.

Drevets, W.C., Videen, T.O., Price, J.L., Preskorn, S.H., Carmichael, S.T., and Raichle, M.E. (1992). A functional anatomical study of unipolar depression. J. Neurosci. *12*, 3628–3641.

Efron, B., and Tibshirani, R. (1993). An Introduction to the Bootstrap (New York: Chapman & Hall/CRC).

Eippert, F., Veit, R., Weiskopf, N., Erb, M., Birbaumer, N., and Anders, S. (2007). Regulation of emotional responses elicited by threat-related stimuli. Hum. Brain Mapp. *28*, 409–423.

Etkin, A., and Wager, T.D. (2007). Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am. J. Psychiatry *164*, 1476–1488.

Genovese, C.R., Lazar, N.A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage *15*, 870–878.

Goldapple, K., Segal, Z., Garson, C., Lau, M., Bieling, P., Kennedy, S., and Mayberg, H. (2004). Modulation of cortical-limbic pathways in major depression: treatment-specific effects of cognitive behavior therapy. Arch. Gen. Psychiatry *61*, 34–41.

Goldin, P.R., McRae, K., Ramel, W., and Gross, J.J. (2007). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. Biol. Psychiatry *63*, 577–586.

Gross, J.J., and Munoz, R.F. (1995). Emotion regulation and mental health. Clin. Psychol. Sci. Pract. *2*, 151–164.

Holland, P.C., and Gallagher, M. (1999). Amygdala circuitry in attentional and representational processes. Trends Cogn. Sci. *3*, 65–73.

Hyman, H.H. (1955). Survey Design and Analysis: Principles, Cases, and Procedures (Glencoe, IL: Free Press).

Jensen, J., McIntosh, A.R., Crawley, A.P., Mikulis, D.J., Remington, G., and Kapur, S. (2003). Direct activation of the ventral striatum in anticipation of aversive stimuli. Neuron *40*, 1251–1257.

Johnstone, T., van Reekum, C.M., Urry, H.L., Kalin, N.H., and Davidson, R.J. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. J. Neurosci. *27*, 8877–8884.

Kalisch, R., Wiech, K., Critchley, H.D., Seymour, B., O'Doherty, J.P., Oakley, D.A., Allen, P., and Dolan, R.J. (2005). Anxiety reduction through detachment: subjective, physiological, and neural effects. J. Cogn. Neurosci. *17*, 874–883.

Kalisch, R., Wiech, K., Herrmann, K., and Dolan, R.J. (2006). Neural correlates of self-distraction from anxiety and a process model of cognitive emotion regulation. J. Cogn. Neurosci. *18*, 1266–1276.

Kim, S.H., and Hamann, S. (2007). Neural correlates of positive and negative emotion regulation. J. Cogn. Neurosci. *19*, 776–798.

Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., and Wager, T.D. (2008). Functional grouping and cortical-subcortical interactions in emotion: A meta-analysis of neuroimaging studies. Neuroimage *42*, 998–1031.

Lang, P.J., Greenwald, M.K., Bradley, M.M., and Hamm, A.O. (1993). Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology *30*, 261–273.

Lawrence, N.S., Williams, A.M., Surguladze, S., Giampietro, V., Brammer, M.J., Andrew, C., Frangou, S., Ecker, C., and Phillips, M.L. (2004). Subcortical and ventral prefrontal cortical neural responses to facial expressions distinguish patients with bipolar disorder and major depression. Biol. Psychiatry *55*, 578–587.

Lazarus, R.S. (1991). Cognition and motivation in emotion. Am. Psychol. *46*, 352–367.

Levita, L., Dalley, J.W., and Robbins, T.W. (2002). Nucleus accumbens dopamine and learned fear revisited: a review and some new findings. Behav. Brain Res. *137*, 115–127.

Lieberman, M.D., Eisenberger, N.I., Crockett, M.J., Tom, S.M., Pfeifer, J.H., and Way, B.M. (2007). Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. Psychol. Sci. *18*, 421–428.

MacCorquodale, K., and Meehl, P.E. (1948). On a distinction between hypothetical constructs and intervening variables. Psychol. Rev. *55*, 307–321.

MacKinnon, D.P., Krull, J.L., and Lockwood, C.M. (2000). Equivalence of the mediation, confounding and suppression effect. Prev. Sci. *1*, 173–181.

McClure, S.M., Berns, G.S., and Montague, P.R. (2003). Temporal prediction errors in a passive learning task activate human striatum. Neuron *38*, 339–346.

McRae, K., Ochsner, K.N., Mauss, I.B., Gabrieli, J.J.D., and Gross, J.J. (2008a). Gender differences in emotion regulation: An fMRI study of cognitive reappraisal. Group Processes and Intergroup Relations *11*, 143.

McRae, K., Reiman, E.M., Fort, C.L., Chen, K., and Lane, R.D. (2008b). Association between trait emotional awareness and dorsal anterior cingulate activity during emotion is arousal-dependent. Neuroimage *41*, 648–655.

Metcalfe, J., and Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: dynamics of willpower. Psychol. Rev. *106*, 3–19.

Milad, M.R., and Quirk, G.J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. Nature *420*, 70–74.

Moskowitz, J.T. (2003). Positive affect predicts lower risk of AIDS mortality. Psychosom. Med. *65*, 620–626.

Munafo, M.R., Brown, S.M., and Hariri, A.R. (2008). Serotonin transporter (5-HTTLPR) genotype and amygdala activation: a meta-analysis. Biol. Psychiatry *63*, 852–857.

Nee, D.E., Wager, T.D., and Jonides, J. (2007). A meta-analysis of neuroimaging studies of interference resolution. Cogn. Affect. Behav. Neurosci. *7*, 1–17.

Ochsner, K.N., and Barrett, L.F. (2001). A multiprocess perspective on the neuroscience of emotion. In Emotions: Currrent issues and future directions, T.J. Mayne and G.A. Bonanno, eds. (New York, NY: The Guilford Press), pp. 38–81.

Ochsner, K.N., and Gross, J.J. (2005). The cognitive control of emotion. Trends Cogn. Sci. *9*, 242–249.

Ochsner, K.N., and Gross, J.J. (2008). Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. Curr. Dir. Psychol. Sci. *17*, 153–158.

Ochsner, K.N., Bunge, S.A., Gross, J.J., and Gabrieli, J.D. (2002). Rethinking feelings: an FMRI study of the cognitive regulation of emotion. J. Cogn. Neurosci. *14*, 1215–1229.

Ochsner, K.N., Ray, R.D., Cooper, J.C., Robertson, E.R., Chopra, S., Gabrieli, J.D., and Gross, J.J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. Neuroimage *23*, 483–499.

Ohira, H., Nomura, M., Ichikawa, N., Isowa, T., Iidaka, T., Sato, A., Fukuyama, S., Nakajima, T., and Yamada, J. (2006). Association of neural and physiological responses during voluntary emotion suppression. Neuroimage *29*, 721–733.

Ollinger, J.M., Corbetta, M., and Shulman, G.L. (2001). Separating processes within a trial in event-related functional MRI. Neuroimage *13*, 218–229.

Paton, J.J., Belova, M.A., Morrison, S.E., and Salzman, C.D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. Nature *439*, 865–870.

Phan, K.L., Fitzgerald, D.A., Nathan, P.J., Moore, G.J., Uhde, T.W., and Tancer, M.E. (2005). Neural substrates for voluntary suppression of negative affect: a functional magnetic resonance imaging study. Biol. Psychiatry *57*, 210–219.

Phelps, E.A. (2006). Emotion and cognition: insights from studies of the human amygdala. Annu. Rev. Psychol. *57*, 27–53.

Phillips, M.L., Drevets, W.C., Rauch, S.L., and Lane, R. (2003). Neurobiology of emotion perception II: Implications for major psychiatric disorders. Biol. Psychiatry *54*, 515–528.

Pitkanen, A. (2000). Connectivity of the rat amygdaloid complex. In The amygdala: a functional analysis, J.P. Aggleton, ed. (New York: Oxford University Press), pp. 31–116.

Schaefer, S.M., Jackson, D.C., Davidson, R.J., Aguirre, G.K., Kimberg, D.Y., and Thompson-Schill, S.L. (2002). Modulation of amygdalar activity by the conscious regulation of negative emotion. J. Cogn. Neurosci. *14*, 913–921.

Scherer, K.R. (2001). Appraisal considered as a process of multilevel sequential checking. In Appraisal Processes in Emotion: Theory, Methods, Research, K.R. Scherer, A. Schorr, and T. Johnstone, eds. (New York: Oxford University Press), pp. 92–120.

Schmitz, T.W., and Johnson, S.C. (2007). Relevance to self: A brief review and framework of neural systems underlying appraisal. Neurosci. Biobehav. Rev. *31*, 585–596.

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. Curr. Opin. Neurobiol. *14*, 139–147.

Scott, D.J., Stohler, C.S., Egnatuk, C.M., Wang, H., Koeppe, R.A., and Zubieta, J.K. (2007). Individual differences in reward responding explain placebo-induced expectations and effects. Neuron *55*, 325–336.

Shrout, P.E., and Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. Psychol. Methods *7*, 422–445.

Stern, E.R., Wager, T.D., Egner, T., Hirsch, J., and Mangels, J.A. (2007). Preparatory neural activity predicts performance on a conflict task. Brain Res. *1176*, 92–102.

Tindell, A.J., Smith, K.S., Pecina, S., Berridge, K.C., and Aldridge, J.W. (2006). Ventral pallidum firing codes hedonic reward: when a bad taste turns good. J. Neurophysiol. *96*, 2399–2409.

Tugade, M.M., Fredrickson, B.L., and Feldman Barrett, L. (2004). Psychological resilience and positive emotional granularity: examining the benefits of positive emotions on coping and health. J. Pers. *72*, 1161–1190.

Urry, H.L., van Reekum, C.M., Johnstone, T., Kalin, N.H., Thurow, M.E., Schaefer, H.S., Jackson, C.A., Frye, C.J., Greischar, L.L., Alexander, A.L., and Davidson, R.J. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. J. Neurosci. *26*, 4415–4425.

van Reekum, C.M., Johnstone, T., Urry, H.L., Thurow, M.E., Schaefer, H.S., Alexander, A.L., and Davidson, R.J. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. Neuroimage *36*, 1041–1055.

Wager, T.D., Keller, M.C., Lacey, S.C., and Jonides, J. (2005). Increased sensitivity in neuroimaging analyses using robust regression. Neuroimage *26*, 99–113.

Wager, T.D., Scott, D.J., and Zubieta, J.K. (2007). Placebo effects on human mu-opioid activity during pain. Proc. Natl. Acad. Sci. USA *104*, 11056–11061.

Wager, T.D., Barrett, L.F., Bliss-Moreau, E., Lindquist, K., Duncan, S., Kober, H., Joseph, J., Davidson, M., and Mize, J. (2008). The neuroimaging of emotion. In Handbook of Emotion, M. Lewis, J.M. Haviland-Jones, and L.F. Barrett, eds. (New York: Guilford Press), pp. 249–271.

Whalen, P.J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. Curr. Dir. Psychol. Sci. *7*, 177–188.

Neuron, volume *59*

Supplemental Data

Prefrontal-Subcortical Pathways Mediating Successful Emotion Regulation

Tor D. Wager, Matthew L. Davidson, Brent L. Hughes, Martin A. Lindquist, and Kevin N. Ochsner

List of Supplemental Tables and Figures

Supplemental Text

Table S1

	Coordinates			Cluster size		a path		ab path		Num. voxels	
Name	**x**	**y**	**z**	**Vox**	**Vol**	**Z**	**p**	**Z**	**p**	**P<.001**	**P<.005**
Mediated by nuc. accumbens											
VStr/basal forebrain	28	7	-18	25	1329	3.46	0.0005	3.48	0.0005	4	25
VMPFC	17	31	-14	303	16112	3.58	0.0003	3.58	0.0003	40	303
L VLPFC	-48	28	-14	31	1648	3.57	0.0004	3.55	0.0004	4	31
R VLPFC	41	48	0	13	691	3.35	0.0008	3.18	0.0015	0	13
Rostral MPFC	3	58	9	136	7232	3.58	0.0003	3.58	0.0003	24	136
R aPFC	21	55	18	12	638	3.56	0.0004	3.58	0.0003	7	12
L DLPFC	-31	24	45	15	798	3.55	0.0004	3.28	0.001	0	15
DMPFC	10	28	45	58	3084	3.58	0.0003	3.52	0.0004	8	58
Mediated by left amygdala											
R VLPFC	48	31	-9	17	904	3.51	0.0004	-3.448	0.0006	1	17
R mid-lateral OFC	28	31	-14	3	160	3.54	0.0004	-3.581	0.0003	3	3

Figure S1

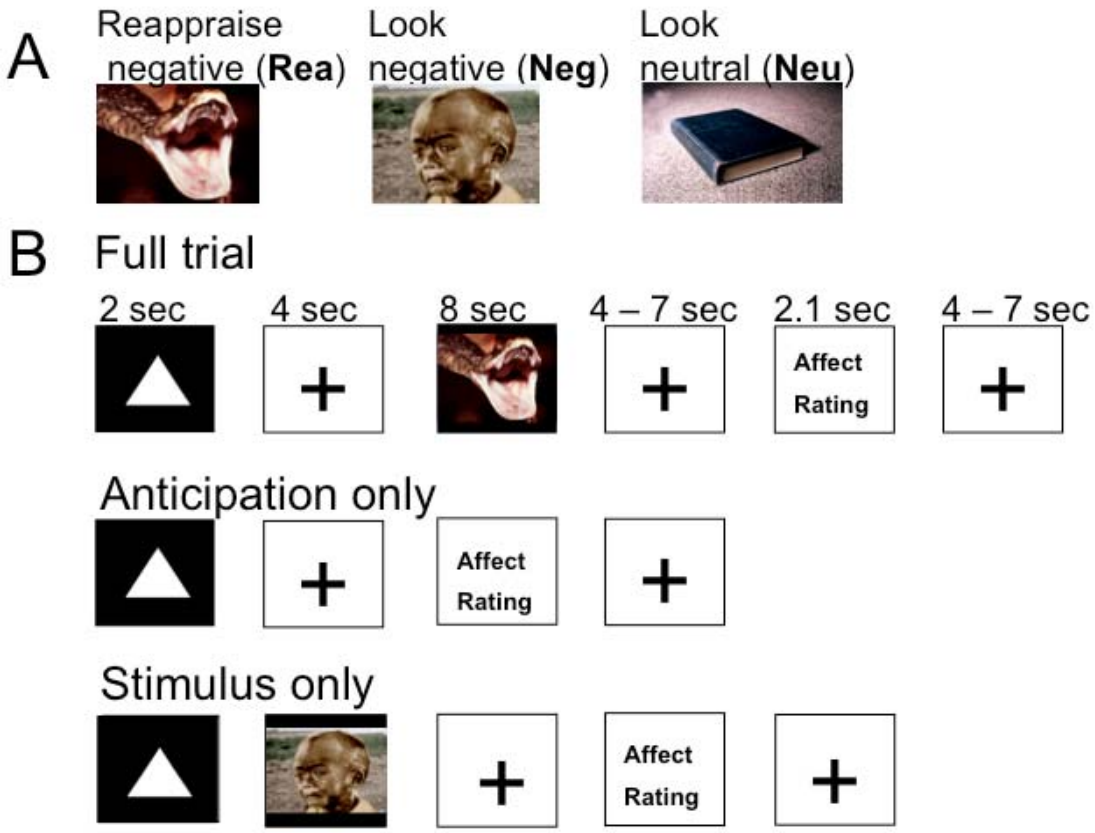


Figure S1. Task design. The design includes three trial types: Anticipation and stimulus trials, Anticipation only trials, and Stimulus only trials, for each of the 3 conditions (Reappraise Negative [ReappNeg], Experience Negative [LookNeg], and Experience Neutral [LookNeu]).

Figure S2

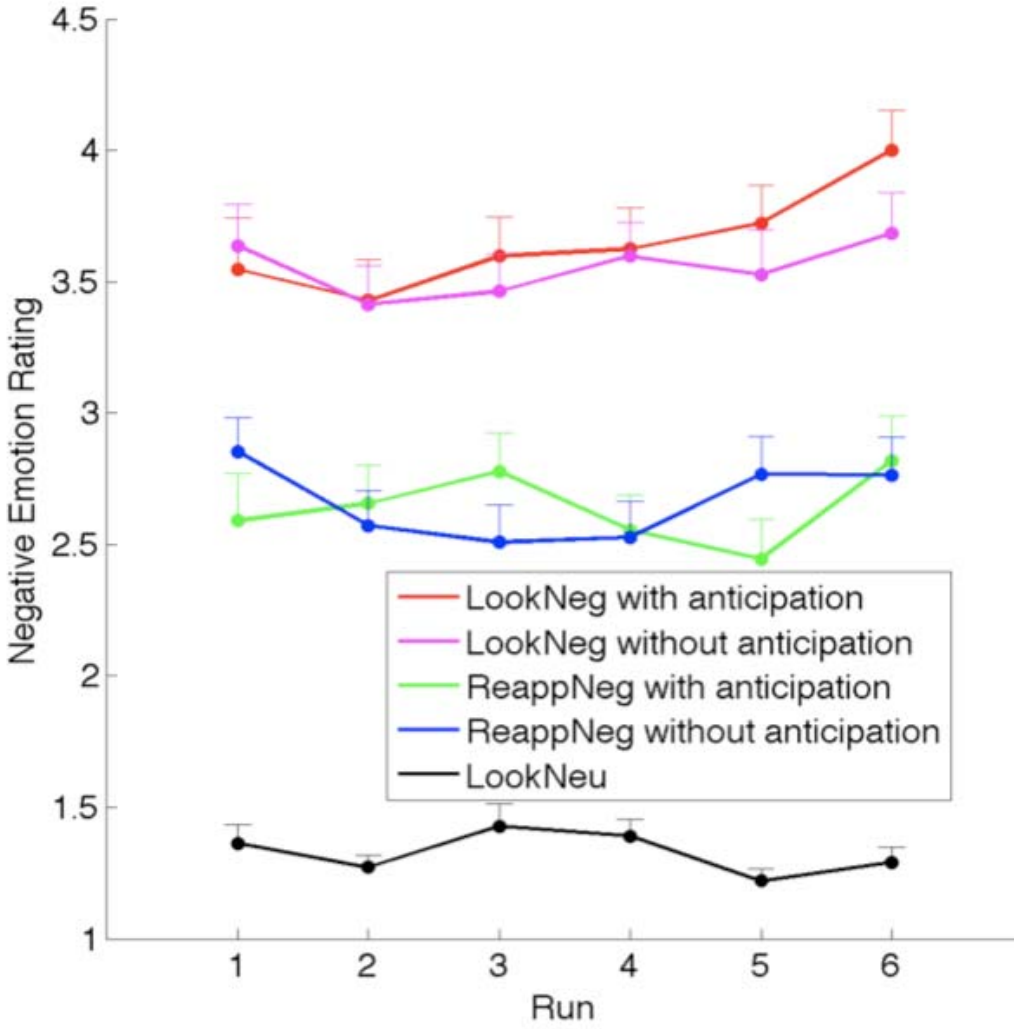


Figure S2. Emotion rating data. Group average ratings (y-axis) as a function of scanning run (x-axis) and task condition (lines). There was a robust effect of viewing negative vs. neutral images and a robust effect of reappraisal, both of which were sustained across runs. Neither interacted with whether the trial included a 4 sec anticipation period.

Figure S3

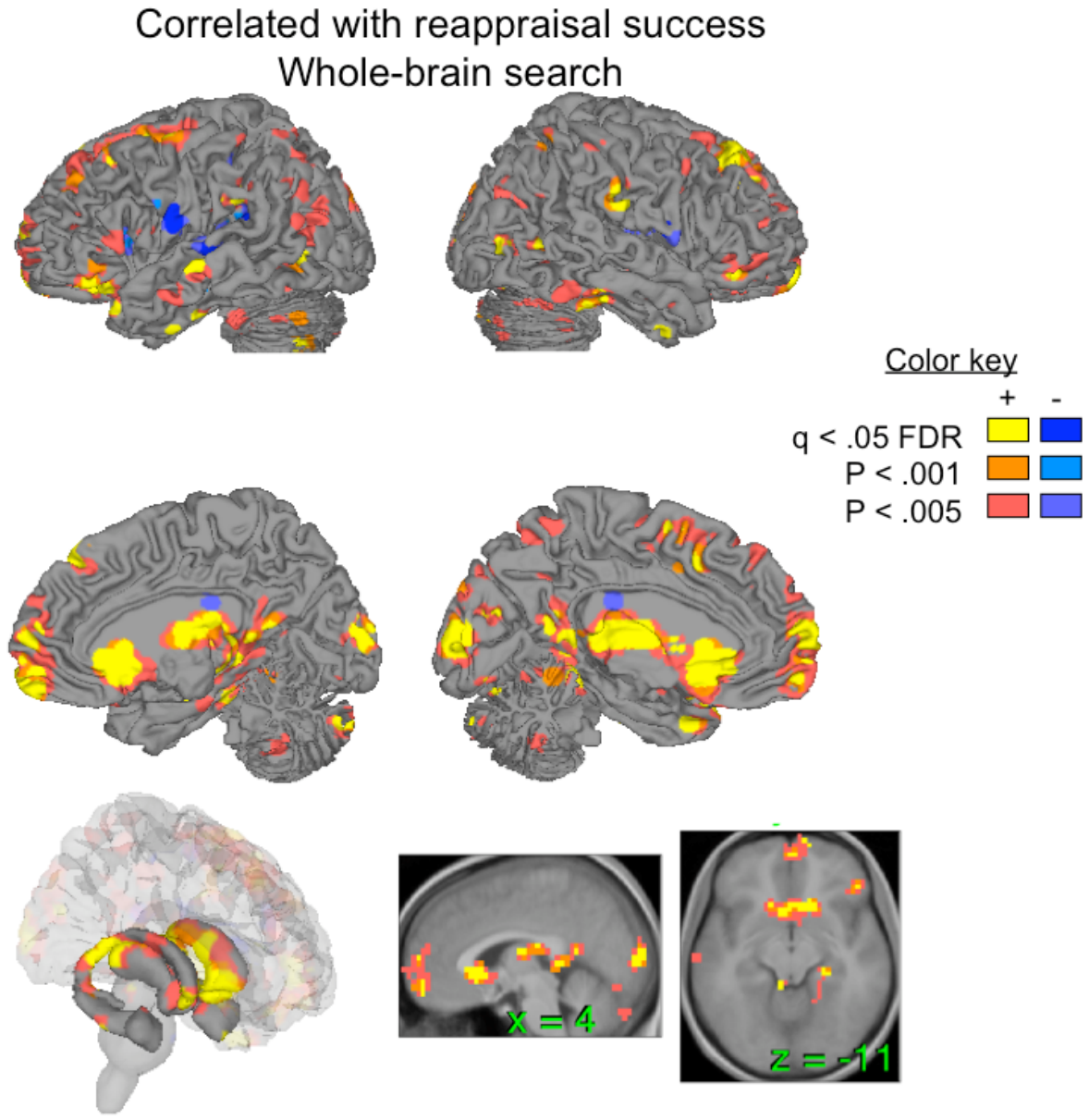


Figure S3. Whole-brain correlations between reappraisal activation [ReappNeg - LookNeg] and reappraisal success in reported experience [LookNeg – ReappNeg]. Positive correlations are shown in red/yellow, and negative correlations are shown in blue. Positive correlations indicate a greater relative increase in activity for participants who report more successful reappraisal , and negative correlations indicate a greater relative decrease in activity (LookNeg - ReappNeg) for participants who report more successful reappraisal. Thresholds are shown in the color key on the figure.

Figure S4

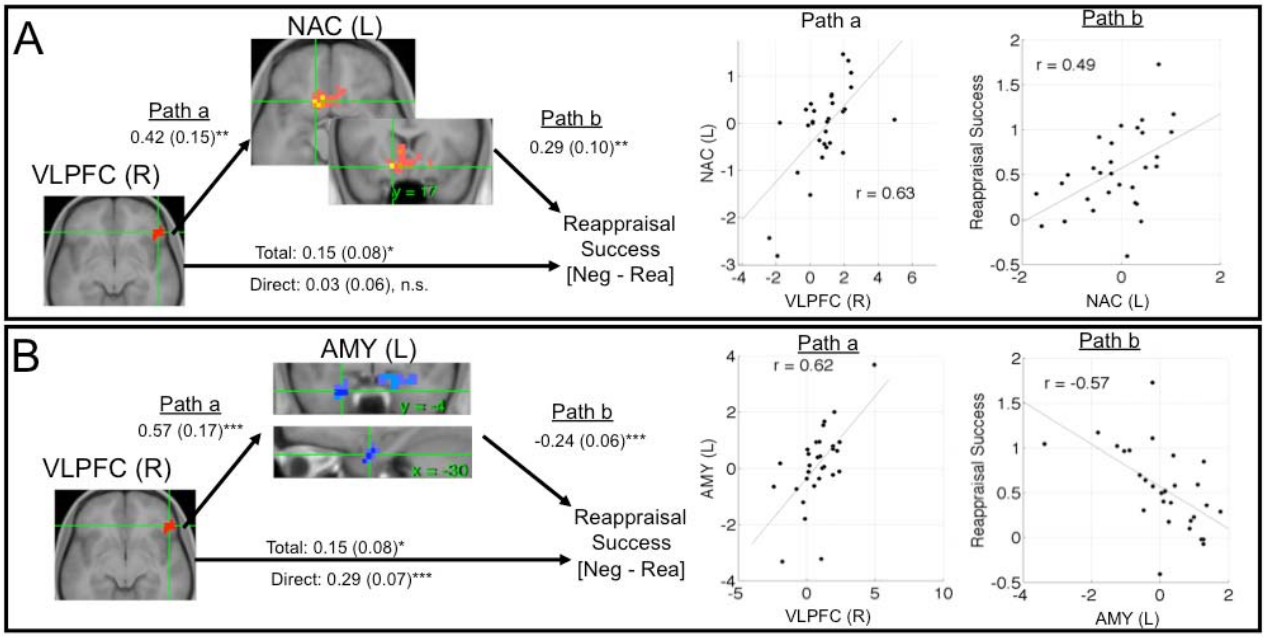


Figure S4. Mediation analyses for amygdala and nucleus accumbens regions of interest. The mediation effect parametric mapping identified regions that were mediators of the relationship between right ventrolateral prefrontal cortex (vlPFC) and reappraisal success by searching over voxels and performing mediation tests one voxel at a time. This figure shows statistics for the path models, averaging over voxels in each mediating region, for significant left nucleus accumbens/ventral striatal and left amygdala regions. These results differ slightly from the ones presented in Figure 3 because here each region is subjected to a separate mediation analysis, whereas in the analysis shown in Figure 3, they were included in the same mediation model.

Figure S5

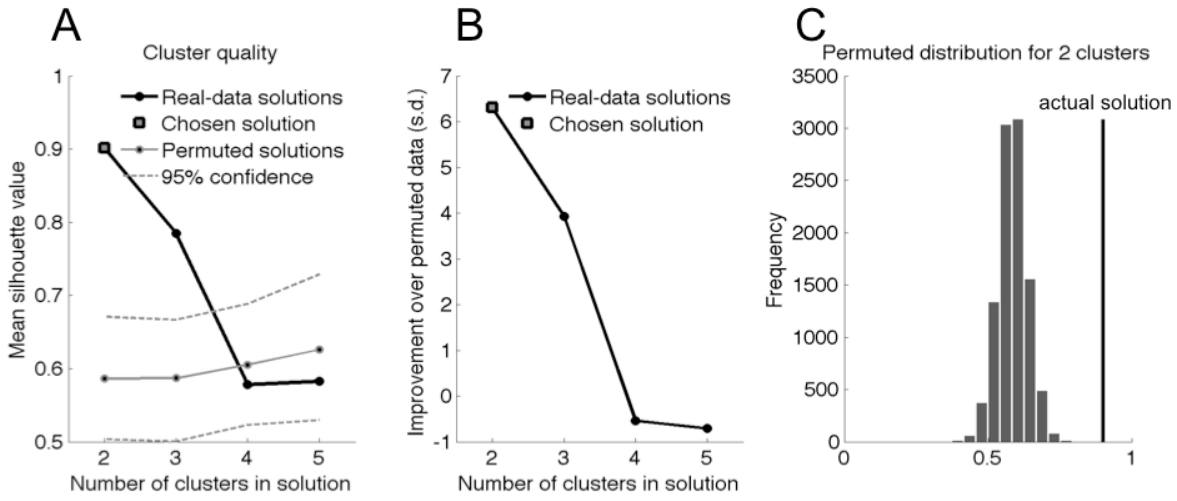


Figure S5. Results for cluster analysis examining grouping of mediators into functional networks. A) Cluster quality (y-axis, see Experimental Procedures) as a function of number of clusters in the candidate solution. Clusters in this context refer to interconnected networks of brain rgions. Higher values indicate tighter clustering of the data. The solid black line shows the actual data, with the chosen solution (2-cluster solution) marked with a gray square. The thin black line shows the average quality for permuted data, with 95% confidence intervals marked by dashed lines. Permutation disrupted the relationships among variables, so that there was no true grouping of the data into clusters, so that cluster quality under the null hypothesis of no grouping could be assessed. B) Z-scores (y-axis) for the real-data cluster quality relative to the cluster quality of permuted data. C) Distribution of null-hypothesis cluster quality (x-axis; frequency on y-axis) for the chosen 2-cluster solution. The real-data solution is marked with a vertical line.

Supplemental Discussion

Interpreting suppression effects in path models

A suppression effect occurs when the indirect pathway and the direct pathway have opposite signs—that is, when two variables independently influence an outcome variable (Y), but with effects of opposite signs. It might be best illustrated with an example from another domain. For example, increased use of technology may have both positive and negative effects on life satisfaction (Y), for different reasons. These effects tend to cancel each other out, resulting in a weak overall technology -> satisfaction relationship, unless the mediating variables with opposite signs can be identified: One might discover a positive pathway from technology -> connect with family and friends -> increased satisfaction, and a negative pathway from technology -> reachable by boss at home -> decreased satisfaction. Controlling for the suppressor variable (reachable by boss) would allow other positive relationships and any positive direct technology -> satisfaction effect to emerge.

There are many examples of suppression effects in published literature, and they can often uncover strong direct relationships that are masked by confounding suppressor variables or explain results that are initially counter-intuitive. As an example of the first case, Goldberg (Goldberg et al., 1996) found that a steroid use-prevention program reduced intentions to use steroids. They also found that greater reason for using steroids was related to increased intentions to use. Both of these effects are intuitive. However, as MacKinnon et al. (MacKinnon et al., 2000) describe in their re-analysis of these data, those who had greater reason to use steroids tended to enroll in the program, creating a positive link between reasons for using steroids and program enrollment. Thus, the indirect pathway connecting program enrollment, more reasons for using steroids, and increased intentions to use partially canceled out the beneficial direct effects of the program. After controlling for the confounding suppressor (reasons to use), the beneficial effects of the program grow stronger.

As an example of the second case, Guber (1999) describes a dataset drawn from the 1997 Digest of Education Statistics on the relationship between state spending on public schools and SAT scores. Surprisingly, the original relationship between spending and SAT scores across U.S. states was negative: More spending predicted significantly lower scores. However, a likely cause is that a third variable was acting as a confounding suppressor variable. High-spending

states tended to require all students to take the SAT, even the academically poor ones; thus, percentage of students taking the test is a strong negative predictor of average SAT. When percent taking the test was included as a mediator, the direct relationship between spending and SAT became significantly positive, in line with what one might expect.

In the present study, the amygdala may appear to be a complete mediator of the PFC-reappraisal success correlation (in the sense that the direct relationship is not significant) when it is the only mediator in the model, but that is partially because there is still much unexplained variance in this model that masks the direct effect. Including the nuc. accumbens (NAC) in the model explains additional variance, allowing the direct effect to reach statistical significance.

Additional implications of dual routes for reappraisal effects on emotion: how we measure changes in emotional responding

The finding that fMRI responses in both the amygdala and NAC/VS were meaningfully related to emotional experience has implications for both understanding amygdala and striatal function and the use of self-reported emotion in neuroscience research more generally. Consider first that although the amygdala has been broadly implicated in negative emotional responses, ranging from the acquisition of conditioned fear (Phelps, 2006) to hyperactivity in a range of anxiety disorders (Etkin and Wager, 2007), there has been reason to doubt that amygdala activation reflects experiential changes in all situations in which it has been activated (Anderson and Phelps, 2002; Barrett and Wager, 2006; Wager et al., in press). In part this may be due to the fact that the amgydala is a complex, heterogeneous group of structures, only some of which have been linked to emotional experience in some situations. Several meta-analyses from our group have shown that although the superior amygdala is activated both by stimuli that do (e.g. a small child crying hysterically) and stimuli that do not (e.g. a fearful facial expression) generate negative emotional experiences, it is more consistently activated by the stimuli that *do not* elicit experience (Phan et al., 2002; Phan et al., 2004; Wager et al., in press). This finding, and the individuals results on which it is based, have been taken to suggest that the amygdala may play a role in the perceptual detection and encoding of affectively relevant stimuli, which may or may not have a direct impact on affective experience (Anderson and Phelps, 2001, 2002; Barret et al., 2007; Wager et al., in press; Whalen, 1998). Therefore, the finding that amygdala activity mediates negative emotional experience *in this task context* is a significant finding, and it

provides support for the notion that prefrontal-amygdala correlations in other reappraisal studies that use the amygdala as an outcome variable – but did not measure emotional experience (or any other behavioral index of emotional response) (Johnstone et al., 2007; Urry et al., 2006; van Reekum et al., 2007) – may in fact be relevant to shaping emotional experience.

This argument leads to the broader implication of the amygdala-experience relationship observed here. One reason some of the aforementioned papers have not measured emotional experience is because of concerns that it may not be a wholly reliable or valid index of emotional responding. As noted at the outset, we recognized the necessity of behavioral measures for constraining functional interpretations of neural indicators (Poldrack, 2006), and selected self-reports as our behavioral measure because it can be reliably collected in the scanner environment, provides clear valence information (unlike most autonomic measures), and has been shown to predict mental and physical health outcomes (Brosschot et al., 2006; Gross and Munoz, 1995; Moskowitz, 2003; Scheier and Carver, 1992; Tugade et al., 2004). The fact that changes in self-reported negative emotion were mediated by activity in brain systems thought – on indepedent *a priori* grounds – to be important for emotional appraisals provides additional support for self-report as a useful measure for studies of emotion and its regulation.

Additional implications for the role of prefrontal cortex in emotional appraisal: the nature of correlations between PFC and amygdala

An interesting difference between this study and some previous studies of reappraisal is that we found positive correlations between vlPFC and amygdala, whereas several studies mentioned above found negative correlations. One reason for these differences may have to do with the structure of the reappraisal task as used by different investigators. In the Urry et al. and Johnstone et al. papers, the appraisal process was permitted to evolve for 4 s before presenting a reappraisal cue, perhaps separating in time an negative initial appraisal process that involves positive vlPFC-amgydala connectivity and a subsequent positive reappraisal process that involves negative vlPFC-activity. However, it is also possible that amygdala activity can be related in complex ways to appraisal. Kim et al., for example, found amygdala increases when participants increased positive emotional responses (Kim and Hamann, 2007), and studies of autobiographical memory retrieval have found co-activation of amygdala and vlPFC (Greenberg et al., 2005; Maguire and Frith, 2003). This raises the possibility that reappraisal cues presented

in advance of pictures in our paradigm may have strengthened participants' ability to engage in meaning-based reinterpretations, shifting frontal-amygdala connectivity positively. The present data cannot disentangle these and other potential explanations, but it opens the way for a more detailed mapping of effective connectivity as a function of specific strategies and stimulus types.

Additional implications of network analyses for understanding cognition-emotion interactions

The first network involved the NAC/VS (Figure 5B) and included three additional regions that have been implicated in action selection and memory. However, we also note that the NAC/VS region we identified extended into the subgenual cingulate, which has recently been shown to play an important role in the modulation of depression and mood (Drevets, 2001; Johansen-Berg et al., 2007; Mayberg et al., 1999); high-resolution studies are needed to dissociate the contributions of NAC and subgenual cingulate to the pathways reported here. One additional region in this network was the pre-SMA, which has been implicated in multiple cognitive control processes (van Snellenberg and Wager, in press; Wager et al., 2004a; Wager and Smith, 2003) and in the energization of internally driven processes that lead to effective intentional response selection (Alexander et al., 2007; Cunnington et al., 2005; Isoda and Hikosaka, 2007; Sumner et al., 2007) (Lau et al., 2004). A second region was the precuneus, which has been implicated in the control and switching of attention among objects and object features (Barber and Carter, 2005; Wager et al., 2004a) and in monitoring of the internal environment and "self" (Cavanna and Trimble, 2006; Gusnard and Raichle, 2001) and episodic memory retrieval (Lundstrom et al., 2005). Finally, a third region encompassed the retrosplenial cingulate cortex and cingulate isthmus, which have been implicated in autobiographical memory (Maddock, 1999) and may provide an anatomical bridge from the precuneus to medial temporal lobe structures that have well-known roles in the formation of declarative memories (Davachi, 2006). The positive association of vlPFC with activity in this network suggests that the selection of reappraisal-appropriate information from memory may be used to construct a positive construal of the situations depicted in negative images (Ochsner, 2004), and that this can mediate reappraisal success. Although speculative, this interpretation is consistent with what is known about the regions involved in this network and is consistent with the connectivity pattern we observed, which indicated a pathway from vlPFC to pre-SMA, which served as a network "hub" that is connected to the precuneus and retrosplenial cortex/cingulate isthmus, which in turn

connected to the NAC/VS. In this network context, one way of reappraising an aversive event can be seen to involve the interactions of vlPFC with a network of regions that together represent a positive, approach motivated appraisal of the event.

Here it should be noted that our search for frontal "mediated" regions suggested that the medial PFC (including dorsal, ventral, and rostral divisions) was functionally connected to this network. Though we found strong correlations between rostral MPFC and reappraisal success, it was not a primary focus here because we did not find overall reappraisal-related activation in our group of participants. However, its focus in previous papers (Johnstone et al., 2007; Urry et al., 2006) suggests that it may play an important role in reappraisal, and in our study it seemed to be related to the enhancement of NAC/VS activity that, in turn, predicted reappraisal success.

The second network involved the amgydala and included three additional anatomically interconnected areas that have been commonly associated with negative appraisals and negative emotion more generally (Figure 5B). The first region was the right lateral orbitofrontal cortex (lOFC), which has been associated with diverse forms of valenced affective experience (Berridge and Kringelbach, 2008; Wager et al., in press) and related motivational processes such as updating behavior on the basis of negative feedback (O'Doherty, 2003; O'Doherty et al., 2003) and negative (aversive) prediction errors (Seymour et al., 2005). The second was the anterior insula (AI), which is activated by aversive stimuli of various kinds and is most commonly associated with negative emotional experience in normal emotion (Wager et al., in press) and increased emotional activation across several kinds of anxiety disorders (Etkin and Wager, 2007). This region has been called primary "interoceptive" cortex (Craig, 2003), though it appears to be activated by aversive social experiences as well (Sanfey et al., 2003). The third was the rostral dorsal cingulate (rdACC), a region often activated when cognitive expectancies modulate affective processes. For example, hypnosis (Faymonville et al., 2000; Rainville et al., 1997) and placebo treatments (Lieberman et al., 2004; Price et al., 2007; Wager et al., 2004b) that lead to reduced pain have consistently reduced rdACC activity. A fourth region in the subthalamus was also included in this network, but we do not speculate on its interpretation here. The positive association of vlPFC activity with activity in this network is consistent with the idea that reappraisal success can also hinge on limiting activity in a network of structures important for representing negative affective states.

Supplemental Experimental Procedures

Reappraisal training procedures

Prior to scanning, participants completed a training session on the reappraisal strategy. Previous work on the reappraisal process has established that an effective means of down-regulating negative emotion is by re-interpreting the affects, dispositions, outcomes or contexts for the actors and actions shown in less negative ways. Prior to scanning, participants were trained in the use of this strategy, which is known as reinterpretation (Ochsner and Gross, 2008). For example, if participants viewed an image of a crying baby in a barren landscape, they might imagine that the mother has simply gone to do an errand and will return soon. Different specific reappraisals – all of the same basic kind – were generated for each image. As described previously (Ochsner et al., 2002; Ochsner et al., 2004), this method is meant to capture the fact that in everyday no single type of reinterpretation is expected to be applicable to all life events – or here, all photos.

The training session consisted of two parts. In the first part, participants were asked to memorize the cue-task condition associations (e.g., circle with LookNeu, etc.). Participants were subsequently quizzed until they achieved 100% accuracy on 10 consecutive trials of each type. In part two, participants completed 7 sample trials with the experimenter present (1 Look Neutral, 2 Look Negative, 4 Reappraise Negative). On reappraisal trials they were given feedback on the appropriateness of their reappraisals for each image. During this time participants also were reminded not to look away from images or distract themselves with irrelevant and/or positive thoughts unrelated to the context of the image. After appropriate coaching to ensure that participants could reinterpret images quickly and effectively, the training ended with the completion of 18 practice trials (6 Look Neutral, 6 Look Negative, 6 Reappraise Negative) on their own. Images used during training were different from those used during the subsequent test. Eye position was monitored during scanning by viewing an image of the right eye projected onto a screen in the scanner control room (I-SCAN, Inc.). No subjects closed their eyes or averted their gaze during image viewing.

fMRI task design

Previous studies of reappraisal have not separated brain activity related to anticipation

and instruction processing, stimulus viewing, and picture rating, and a goal of our task design was to provide the ability to estimate separately brain activation magnitude related to each of these three phases of the image viewing and rating process. To accomplish this, a partial trial design was employed (Ollinger et al., 2001; Stern et al., 2007). Within each task condition, LookNeu, LookNeg, and ReappNeg, three different trial types were used: full trials, anticipation-only trials, and stimulus-only trials (see Supplementary Figure 1).

On full trials, a 2 sec condition cue was followed by a 4 sec anticipatory interval during which a fixation cross was presented on the screen. The image was subsequently presented for 8-sec. Following image presentation, a fixation cross was presented during a 4 or 7 sec jittered inter-stimulus interval (ISI; uniform distribution of 4 and 7 sec intervals). Following the ISI period, the words "how negative do you feel?" appeared on-screen for 2.1 sec, and participants were asked to rate negative affect on a five-point scale by pressing a button with one of five fingers on a button-response unit (1 = "not at all negative," indicated by a thumb button press, up to 5 = "extremely negative," indicated by a fifth-finger button press). Following the rating, a 4 or 7 sec jittered ISI concluded the trial.

The trial was identical to the full trial through the anticipation interval. Instead of an image presentation, participants were asked to rate the negative feelings experienced during anticipation on the same five-point scale. The stimulus only trials were identical to the full trials, except that the 4 sec anticipation interval was omitted. This design allowed us to construct predictors for Cue-, Anticipation-, and Image-related brain activity related to each task condition in the General Linear Model (GLM) that were uncorrelated enough to provide efficient estimates of activation in each task condition x trial phase combination (see below). In addition, responses during picture or anticipation rating could be estimated separately, thus permitting the parsing of brain activity during trial phases of interest from brain activity related to the reporting and button-press processes.

Subjects completed 6 functional runs of 18 trials each for a total of 108 trials. Each run included 6 trials of each condition (Reapp Neg, Look Neg, Look Neutral) and trial type (full trial, ant only trial, stim only trial).

In this report, we compared brain responses during the viewing of aversive images in the ReappNeg vs. LookNeg conditions, as estimated controlling for activity during other phases of the trial (Cue, Anticipation, and Report). Negative emotion ratings did not differ significantly

based whether the anticipation period was included or not in either Look or Reappraise conditions (see Supplementary Figure 2). For LookNeg, the average rating difference for full trial vs. stimulus only trials was 0.10, paired $t(35) = 1.58$, $p > 0.10$. For ReappNeg, the difference was -0.02, paired $t(35) = -0.29$, $p > 0.7$. Subsequent analyses of behavioral and brain data were conducted on averages across full trials and stimulus-only trials. Likewise, reported reappraisal success (LookNeg – ReappNeg) was stable across time (Success x Run interaction $F(5, 172) = 1.02$, $p = .41$), and ratings for all image types were likewise stable across time. Thus, subsequent analyses were also conducted on averages across runs.

Image processing and data analysis

Preprocessing. Functional images were slice-time and motion corrected using FSL (FMRIB Centre, University of Oxford). Structural T1-weighted images were coregistered to the first functional image for each subject using an iterative procedure of automated registration using mutual information coregistration in SPM2 and manual adjustment of the automated algorithm's starting point by a trained analyst until the automated procedure provided satisfactory alignment. Structural images were normalized (spatially warped) to a standard template brain (the MNI avg152T1.img) using SPM2 software (Wellcome Department of Cognitive Neurology, UCL) using default options (7 x 8 x 7 nonlinear basis functions), and the warping parameters were applied to functional images for each subject. Normalized functional images were interpolated to 2 x 2 x 2 mm voxels and spatially smoothed with a 6-mm Gaussian filter.

First-level GLM model. First-level GLM analysis for each participant was performed using SPM2 (Friston, Jezzard, & Turner, 2004). Effects were modeled as a boxcar regressor convolved with a canonical hemodynamic response function (double-gamma) for the 2 sec cue period, 4 sec anticipation period, 8 sec stimulus viewing period during which subjects either attend or reappraise, and 2.1 sec rating period separately. Twelve regressors were specified, for cue-related responses, anticipation-related responses, image-viewing related responses, and rating-related responses in each of the three task conditions (LookNeu, LookNeg, and ReappNeg). Separate sets of regressors were specified for each of the 6 scanning runs for each subject. In addition, regressors specifiying a high-pass filter with a 120 sec cutoff were included

to model low-frequency drift. Since our primary concern was group statistics, no autoregressive (AR) model was used, which are unbiased and valid even without AR modeling.

Importantly, the trial design resulted in regressors that were essentially uncorrelated across conditions of interest. The critical feature of the design for the results we present here is that the regressors related to Image Viewing for each task condition are not collinear with combinations of other regressors, so that their estimates are stable and efficient. The average correlations between each Image-Viewing related regressor and other regressors (across other regressors and subjects) were r = 0.067, r = 0.063, and r = 0.062 for LookNeu, LookNeg, and ReappNeg. The correlations were very similar across both regressors and subjects as well: The maximum correlations with any other regressor for any subject were r = 0.074, r = 0.067, and r = 0.070, for each task condition, and the minimum correlations were r = -0.17, r = -0.18, and r = -0.17. Variance inflation factors reflect the overall multicolinearlity in a design matrix for each regressor and were low (< 2) for all twelve regressors and for all subjects, indicating that the task design was effective in providing stable and efficient estimates of brain activity related to each task phase (cue, anticipation, image viewing, and rating) and task condition.

Contrasts. Activation estimates were obtained for each subject using SPM2, and contrasts across conditions were estimated and analyzed in a second-level group analysis treating subject as a random effect. The contrast of interest in this report was [ReappNeg image viewing – LookNeg image viewing]. Positive contrast values indicate greater relative activity during reappraisal vs. natural experience of aversive images, and negative values indicate greater relative activity during experience (i.e., the reverse subtraction). Both effects were analyzed and reported here. As our main *a priori* hypotheses concerned changes during image viewing, contrasts related to other phases of the trial (Cue, Anticipation, and Report) will be addressed in subsequent papers. This choice serves to avoid undue complexity and preserve clarity in the present report.

Group analysis (Analysis steps 1-2). The second level random-effects analysis was performed using robust regression, a technique that both increases statistical power and decreases false positive rates in the presence of outliers (Wager, Keller, Lacey, & Jonides, 2005). Reported reappraisal success was calculated for each subject as the average difference in negative affect reports for [LookNeg – ReappNeg]. The second-level design matrix included two regressors: one corresponding to reappraisal success, and the other an intercept term.

Reappraisal success scores were centered by subtracting the mean, allowing the intercept term to be interpreted as the population estimate for reappraisal-induced activation ([ReappNeg – LookNeg]) for a subject who shows average reappraisal success. The interpretation of the reappraisal success regressor is the change in reappraisal-induced activation as a function of reappraisal success, i.e. the activation-reappraisal success relationship. The advantage of including reappraisal-induced activation ([ReappNeg – LookNeg]) in the model is that it accounts for known sources of individual variation when testing the significance of average activation contrast values. Thus, for voxels that do show a brain activity-reappraisal success relationship, this model has greater sensitivity to detect overall activation compared with an intercept-only model (which is what is typically performed, e.g., in SPM software).

An anatomically defined gray matter mask was created based on the Montreal Neurologic Institute (MNI) avg152T1 template (smoothed with an 8 mm FWHM filter and thresholded at a value of 0.5) and explicitly specified during analysis.

Localization of results. Normalized structural T1 images were averaged across participants to create an anatomical underlay for visualizing significant regions of activation, and for visually assessing normalization quality. In our experience, this is advantageous because the quality of nonlinear warping can vary across brain regions, resulting in greater differences between the standard brain and participants' actual T1 images in ventral subcortical regions in particular. We assessed activation locations based on identified structural landmarks in our participants (e.g., amygdala gray matter, striatal gray matter) using several brain atlases (Haines, 2000; Mai et al., 2004; Martin, 1996), rather than relying on standardized Talairach or MNI coordinates, which can be misleading in some brain regions depending on warping quality. However, we provide MNI coordinates for reference and use in future studies and meta-analyses.

Mediation Effect Parametric Maps (Analysis steps 3, 4a). The Mediation Effect Parametric Map (MEPM) analysis is based on a standard 3-variable path model (Baron and Kenny, 1986) with a bootstrap test for statistical significance (Efron and Tibshirani, 1993; Shrout and Bolger, 2002). A test for mediation tests whether a covariance between two variables (X and Y) can be explained by a third variable (M). A significant mediator is one whose inclusion as an intermediate variable in a path model of the effects of X on Y significantly affects the slope of the X – Y relationship; that is, the difference ($c - c'$) is statistically significant (see Results for nomenclature). The test of mediation is fundamentally different from a moderation or "psycho-

physiological interaction" (PPI) effect, in which the *level* of a moderating variable M predicts the strength of the X –Y relationship. A moderator interacts with the X –Y relationship, whereas a mediator explains it. Thus, the mediation test (not the moderation test) is critical for localizing and testing functional pathways that span more than two regions. Brain regions that are *mediators* are candidates for links in functional pathways that relate brain activity in multiple regions to behavior and other outcomes.

In the current application, we used right VLPFC activity in the [ReappNeg – LookNeg] contrast as the X variable and reappraisal success as the Y variable. Thus, the X-Y relationship (and the *c* path) is the linear association between prefrontal increases during reappraisal and reported emotion decreases. Thus, the association is the correlation in contrast estimates across subjects in this case, rather than within-subject time series values. The MEPM strategy is to search for voxels in the brain that mediate that relationship, so that X and Y are specified as in Figure 1, but the mediator (M) is unknown. In this case, we first searched for voxels whose [ReappNeg – LookNeg] contrast values mediate the VLPFC – reappraisal success relationship within amygdala and NAC, and then conducted a whole-brain search to test for additional mediating brain regions that may form functional networks with our *a priori* target regions.

More formally, the mediation test can be captured in a system of three equations:

$$y = cx + e_y$$

$$m = ax + e_m$$

$$y = bm + c'x + e'_y$$

where *y*, *x*, and *m* are *n* (participants) x 1 data vectors containing the outcome (y, reappraisal success), the predictor (x, VLPFC), and data from a candidate mediating voxel (m, [ReappNeg – LookNeg] contrast values). e_y , e_m, and $e`_y$ vectors denote residual error for the outcome and mediator controlling for *x* and the outcome controlling for *x* and *m*, respectively. The *a* path is the estimated linear change in *m* per unit change in *x* (e.g., the slope of the VLPFC-mediator relationship). The *b* path is the slope of the mediator-outcome relationship controlling for *x* (mediator to reappraisal success, controlling for VLPFC). The *c* and *c'* paths are as described above.

Statistical tests on *a* and *b* path coefficients assess the significance of each relationship. In addition, a statistical test of (*c* – *c'*) can be performed by testing the significance of the product of the path coefficients *a***b*. A positive mediator is one involved in a pathway that has a

net positive effect on the outcome (e.g., *a*b* is positive, greater reappraisal success in this case). A negative mediator has a net negative effect (lower reappraisal success). We test the significance of *a*b* using the accelerated, bias-corrected bootstrap test (Efron and Tibshirani, 1993) with 10,000 bootstrap samples to test each of the *a*, *b*, and *a*b* path coefficients at each voxel, saving maps of path coefficients and bootstrapped P-values for each effect. Custom Matlab code (Mathworks, Natick, MA, R2007b) was written to perform these tests and optimize them for speed, so that the whole-brain MEPM maps could be accomplished in ~12 hours on a 4-processor Intel Macintosh computer with 2 dual-core Intel Xeon processors (see Author Note for software download information).

If multiple regions are included as mediators (as in analyses with amygdala and NAacc as independent mediators), then additional equations specifying *a* paths for each additional mediator are added, e.g., $m_2 = a_2 x + e_{m2}$ for a second mediator, and the equation that assesses the direct effects of the mediators is modified to include the additional mediators, e.g., $y = bm + b_2 m_2 + c' \quad x + e' \quad _y$ for a 2-mediator model. Thus, the *b* path for each mediator tests whether it is significantly related to the outcome controlling for all other variables.

Though only a significant *a*b* product is required to provide evidence for a mediation effect, for significance in both ROIs and in the whole-brain analysis, we required that each of the paths *a*, *b*, and *a*b* be significant at $p < .005$, two-tailed, with three contiguous voxels. The conjunction across the three effects locates regions with the strongest evidence for both individual pathway links and the total mediation effect. Our choice of threshold was motivated by 1) the desire to use a standard threshold for comparability with other published results ($p < .005$ is one of the most common (Wager et al., 2007a)); 2) Control of false positives; and 3) Sensitivity, or the ability to detect many of the regions that show true mediation effects. The latter is important because using a threshold low enough to afford relatively high sensitivity enhances the ability to make inferences based on the pattern of results across the brain, and to meaningfully aggregate positive and negative findings across studies.

Mediation analysis is a tool for testing particular pathway relationships, but complementary tools are necessary for identifying functional networks of broadly interconnected regions. Principal components analysis (PCA), independent components analysis, and related tools provide complementary perspectives on neuroimaging data by identifying broadly interconnected networks. Though the MEPM analysis searches for mediating regions, we do not

conceptualize these regions as each constituting a separate functional path; rather, these individual networks are likely to be grouped into distributed functional networks that operate as a unit to mediate prefrontal-experience relationships. The analyses described below address this second point.

Region of interest (ROI) analyses and multiple comparisons correction. Four ROIs were drawn on the group-averaged, normalized T1-weighted images from our sample using custom software (SCAN Lab tools, T.D.W.), using several brain atlases for reference (Haines, 2000; Mai et al., 2004; Martin, 1996). The ROIs were in left and right amygdala (108 and 79 [3.5 x 3.5 x 4.5] mm voxels, respectivley) and left and right NAC/VS (52 and 53 voxels, respectively).

We assessed the family-wise error rate (FWER) for the mediation search analyses in each ROI by using a permutation test ((Nichols and Hayasaka, 2003)), which provides estimates of the chances of obtaining a false positive result anywhere in the ROI, accounting for the observed correlations among voxels. We permuted the rows of the 30 (subjects) x *k* (voxels) matrix of subject contrast values, preserving spatial structure, and re-computed the mediation test (including the significance of *a, b,* and *a*b* effects, and their conjunction) for each of 2000 permutations. The primary chosen threshold of $p < .005$ controlled the false positive rate below $p < .05$ in each ROI. That is, under the null hypothesis of no true relationships among X, M, and Y variables, the chances of a false positive conjunction result (significant *a, b,* and *a*b*) anywhere within a given ROI were less than 5%. Based on the permutation test, corrected p-values were $p < 0.03$ and 0.02 for the L/R amygdalae, and $p < .008$ and .002 for the L/R NAC/VS, respectively. This threshold also provided adequate correction for multiple comparisons for each of the *a* and *b* pathways independently. For the left and right amygdala, *a*: $p < .075$, *b*: $p < .06$ (one-tailed, corrected). For the left and right NAC/VS, *a*: $p < .055$, *b*: $p < .04$.

Clustering of mediators and functional networks (Analysis step 4b).

An advantage of the clustering procedure is that a nonparametric permutation test can be used to assess whether there is significant grouping of regions into clusters (networks), and the null hypothesis that regions are operating either independently or as a single large network can be rejected. Secondly, clustering (as opposed to PCA/ICA alone) provides a way of identifying networks that is stable with respect to rotation of the data in multidimensional space. Rotational indeterminacy and instability have plagued PCA/ICA-based and factor analytic approaches.

Briefly, we first defined mediating regions as sets of contiguous voxels that survived the conjunction test (*a*, *b*, and *a*b,* each at $p < .005$, with a 3-voxel extent threshold) as described above. We extended the regions to include contiguous voxels at $p < .05$ in the conjunction, and calculated the average contrast activity for each subject within each region. We then removed linear effects of right VLPFC activity from each region using robust regression, so that the networks captured relationships among regions with respect to their direct relationship with reappraisal success.

In order to achieve stable clustering in a high-dimensional dataset, it is typical to use a data-reduction algorithm (e.g. PCA) prior to clustering in order to remove dimensions with near-zero variance from the dataset. Here we used non-metric multidimensional scaling (NMDS), which makes fewer assumptions than PCA, as the dissimilarities are not assumed to reflect distances in a Euclidean space (Shepard, 1980). NMDS involves an initial decision about the number of dimensions, c, to retain. Thereafter the NMDS algorithm returns a set of component scores in c-dimensional space, which are then clustered. To choose the appropriate dimensionality, we performed PCA on the covariance matrix of contrast values for each of 12 mediating regions identified in the MEPM analysis and used a scree (eigenvalue) plot to determine the appropriate number of dimensions to retain. The plot showed that 86% of the covariance in the dataset was contained in the first four dimensions, and additional dimensions explained little additional variance. We then calculated the 12 x 12 matrix of correlations among these regions and converted the correlations to dissimilarity values using the formula $(1-r)/2$, so that 0 indicated zero distance and 2 indicated the maximum possible difference. We applied nonmetric multidimensional scaling (NMDS) analysis to the resulting dissimilarity matrix using stress as the error metric (details in (Wager et al., 2007b)), retaining four dimensions. The result was a 12 region x 4 component matrix of scores, which were subjected to cluster analysis.

We used hierarchical clustering with average linkage (clusterdata.m in Matlab R2007b) to identify networks, as in (Kober et al., in press; Wager et al., 2007b). We used a permutation test to choose the number of clusters and to provide inferences on whether the distances between regions were distributed multimodally (as opposed to a single-mode, single-cluster distribution expected if there were no systematic groups of interconnected regions). For each possible solution between 2 and 5 clusters (networks), we first computed a measure of clustering quality, as defined in (Struyf et al., 1996):

$$q = \sum_k \sum_i \frac{d_{i_o} - d_{i_{nn}}}{\max(d_{i_o}, d_{i_{nn}})}$$

where d denotes Euclidean distance, and d_{i_o} is the distance from region i to the center of its own class, $d_{i_{nn}}$ is the distance to the nearest neighboring class, and k indexes over clusters. We then permuted the columns of the dimension scores, re-applied the clustering algorithm, and calculated *q* based on the permuted data. The permutation procedure disrupts clusters of nearby regions by exchanging their locations in each dimension with those of other regions, while conditionalizing on the marginal distribution of regions in each dimension. This process was repeated 10,000 times to develop a null-hypothesis distribution of *q*. Estimating the distribution of q for each candidate number of clusters k allowed us to assess Z-scores observed-data clustering solution, defined as:

$$Z_k = \frac{q_{obs} - \overline{q}_{null}}{\sqrt{\frac{\sum_1^I (q_{null} - \overline{q}_{null})^2}{I}}}$$

where q_{obs} is the quality for the observed solution, q_{null} is the quality for the permuted-data solution for one iteration, and I is the number of iterations (10,000). Fig. S5B shows Z_k on the y-axis plotted against candidate choices for k on the x-axis. The highest Z-value was found for 2 clusters, which we used as our estimate of k. The permuted-data distribution of q is shown for the 2-cluster solution in Fig. S5C, and q for the observed-data solution is shown by the vertical black line. The significance of the results ($p < .0001$, $Z = 6.24$) indicates that there were at least two separable networks of interconnected mediators. However, we did not test the significance of the difference between 2 and other numbers of clusters; other choices of k (e.g., 3 networks) may be reasonable candidates as well, though we note that the 4 and 5 cluster solutions provided poor fits.

Identification of additional mediated regions in the frontal cortex (Analysis step 5).

Though the choice of right VLPFC as a predictor was motivated by theoretical considerations, it is not likely to be unique in its relationship with subcortical affective pathways, given that activity in multiple frontal regions predicted reappraisal success. In a final analysis, we used the MEPM approach to localize frontal regions whose relationship with reappraisal success was mediated by subcortical activity. Two separate analyses were conducted with reappraisal

success as the outcome. In this analysis, both left NAC and amygdala were specified as mediators, and reappraisal success was specified as the outcome. Average activity values in each region from the previous analysis were used for the mediators. We performed a voxel-wise search for frontal regions whose [ReappNeg – LookNeg] contrast values were mediated (significant *a*b* effect) by each subcortical region.

To constrain the search to frontal regions, we created a mask of frontal regions labeled as "frontal" in the International Consortium for Brain Imaging (ICBM) single-subject atlas ((Mazziotta et al., 2001); http://www.loni.ucla.edu/ICBM/; an average of 27 T1-weighted images of a single subject registered to MNI space and manually labeled) and included the gyrus rectus and cingulate gyrus as labeled for more complete coverage. We smoothed the resulting mask with a 3 mm kernel thresholded at a value of 0, and searched for regions with a significant mediation (*a*b*) effect within this mask. For this exploratory analysis, we report results significant at $p < .001$ with a 3-contiguous-voxel extent or $p < .005$ with a 10-voxel extent.

References

Alexander, M.P., Stuss, D.T., Picton, T., Shallice, T., and Gillingham, S. (2007). Regional frontal injuries cause distinct impairments in cognitive control. Neurology *68*, 1515-1523.

Anderson, A.K., and Phelps, E.A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. Nature *411*, 305-309.

Anderson, A.K., and Phelps, E.A. (2002). Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. Journal of Cognitive Neuroscience *14*, 709-720.

Barber, A.D., and Carter, C.S. (2005). Cognitive control involved in overcoming prepotent response tendencies and switching between tasks. Cereb Cortex *15*, 899-912.

Baron, R.M., and Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol *51*, 1173-1182.

Barret, L.F., Mesquita, B., Ochsner, K.N., and Gross, J.J. (2007). The experience of emotion. Annual Review of Psychology *58*, 387-403.

Barrett, L.F., and Wager, T.D. (2006). The Structure of Emotion: Evidence From Neuroimaging Studies. Current Directions in Psychological Science *15*, 79-83.

Berridge, K.C., and Kringelbach, M.L. (2008). Affective neuroscience of pleasure: reward in humans and animals. Psychopharmacology (Berl).

Brosschot, J.F., Gerin, W., and Thayer, J.F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. Journal of Psychosomatic Research *60*, 113-124.

Cavanna, A.E., and Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. Brain *129*, 564-583.

Craig, A.D. (2003). Interoception: the sense of the physiological condition of the body. Curr Opin Neurobiol *13*, 500-505.

Cunnington, R., Windischberger, C., and Moser, E. (2005). Premovement activity of the pre-supplementary motor area and the readiness for action: studies of time-resolved event-related functional MRI. Hum Mov Sci *24*, 644-656.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. Curr Opin Neurobiol *16*, 693-700.

Drevets, W.C. (2001). Neuroimaging and neuropathological studies of depression: implications for the cognitive-emotional features of mood disorders. Curr Opin Neurobiol *11*, 240-249.

Efron, B., and Tibshirani, R. (1993). An Introduction to the Bootstrap (Chapman & Hall/CRC).

Etkin, A., and Wager, T.D. (2007). Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry *164*, 1476-1488.

Faymonville, M.E., Laureys, S., Degueldre, C., DelFiore, G., Luxen, A., Franck, G., Lamy, M., and Maquet, P. (2000). Neural mechanisms of antinociceptive effects of hypnosis. Anesthesiology *92*, 1257-1267.

Goldberg, L., Elliot, D., Clarke, G.N., MacKinnon, D.P., Moe, E., Zoref, L., Green, C., Wolf, S.L., Greffrath, E., Miller, D.J., and Lapin, A. (1996). Effects of a multidimensional anabolic steroid prevention intervention. The Adolescents Training and Learning to Avoid Steroids (ATLAS) Program. Jama *276*, 1555-1562.

Greenberg, D.L., Rice, H.J., Cooper, J.J., Cabeza, R., Rubin, D.C., and Labar, K.S. (2005). Co-activation of the amygdala, hippocampus and inferior frontal gyrus during autobiographical memory retrieval. Neuropsychologia *43*, 659-674.

Gross, J.J., and Munoz, R.F. (1995). Emotion Regulation and Mental Health. Clinical Psychology: Science and Practice *2*, 151-164.

Gusnard, D.A., and Raichle, M.E. (2001). Searching for a baseline: functional imaging and the resting human brain. Nat Rev Neurosci *2*, 685-694.

Haines, D.E. (2000). Neuroanatomy: An Atlas of Structures, Sections, and Systems (Philadelphia: Lippincott Williams & Wilkins).

Isoda, M., and Hikosaka, O. (2007). Switching from automatic to controlled action by monkey medial frontal cortex. Nat Neurosci *10*, 240-248.

Johansen-Berg, H., Gutman, D.A., Behrens, T.E.J., Matthews, P.M., Rushworth, M.F.S., Katz, E., Lozano, A.M., and Mayberg, H.S. (2007). Anatomical Connectivity of the Subgenual Cingulate Region Targeted with Deep Brain Stimulation for Treatment-Resistant Depression. Cerebral Cortex.

Johnstone, T., van Reekum, C.M., Urry, H.L., Kalin, N.H., and Davidson, R.J. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. J Neurosci *27*, 8877-8884.

Kim, S.H., and Hamann, S. (2007). Neural correlates of positive and negative emotion regulation. J Cogn Neurosci *19*, 776-798.

Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Duncan, S., Mize, J., and Wager, T.D. (in press). Networks of Emotion: Functional parcelation and pathway analysis from meta-analytic data.

Lau, H.C., Rogers, R.D., Haggard, P., and Passingham, R.E. (2004). Attention to intention. Science *303*, 1208-1210.

Lieberman, M.D., Jarcho, J.M., Berman, S., Naliboff, B.D., Suyenobu, B.Y., Mandelkern, M., and Mayer, E.A. (2004). The neural correlates of placebo effects: a disruption account. Neuroimage *22*, 447-455.

Lundstrom, B.N., Ingvar, M., and Petersson, K.M. (2005). The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. Neuroimage *27*, 824-834.

MacKinnon, D.P., Krull, J.L., and Lockwood, C.M. (2000). Equivalence of the mediation, confounding and suppression effect. Prev Sci *1*, 173-181.

Maddock, R.J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. Trends Neurosci *22*, 310-316.

Maguire, E.A., and Frith, C.D. (2003). Lateral asymmetry in the hippocampal response to the remoteness of autobiographical memories. J Neurosci *23*, 5302-5307.

Mai, J.K., Assheuer, J., and Paxinos, G. (2004). Atlas of the human brain, 2nd edn (San Diego, Calif.: Elsevier Academic Press).

Martin, J.H. (1996). Neuroanatomy: Text and Atlas, 2nd edn (Stamford, CT: Appleton & Lange).

Mayberg, H.S., Liotti, M., Brannan, S.K., McGinnis, S., Mahurin, R.K., Jerabek, P.A., Silva, J.A., Tekell, J.L., Martin, C.C., Lancaster, J.L., and Fox, P.T. (1999). Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. Am J Psychiatry *156*, 675-682.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B.*, et al.* (2001). A probabilistic atlas and reference system for the human

brain: International Consortium for Brain Mapping (ICBM). Philosophical Transactions of the Royal Society B: Biological Sciences *356*, 1293-1322.

Moskowitz, J.T. (2003). Positive Affect Predicts Lower Risk of AIDS Mortality. (Am Psychosomatic Soc), pp. 620-626.

Nichols, T., and Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res *12*, 419-446.

O'Doherty, J. (2003). Can't learn without you: predictive value coding in orbitofrontal cortex requires the basolateral amygdala. Neuron *39*, 731-733.

O'Doherty, J., Critchley, H., Deichmann, R., and Dolan, R.J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J Neurosci *23*, 7931-7939.

Ochsner, K.N. (2004). Current directions in social cognitive neuroscience. Curr Opin Neurobiol *14*, 254-258.

Ochsner, K.N., Bunge, S.A., Gross, J.J., and Gabrieli, J.D. (2002). Rethinking feelings: an FMRI study of the cognitive regulation of emotion. J Cogn Neurosci *14*, 1215-1229.

Ochsner, K.N., and Gross, J.J. (2008). Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. Currents Directions in Psychological Science *17*, 153-158.

Ochsner, K.N., Ray, R.D., Cooper, J.C., Robertson, E.R., Chopra, S., Gabrieli, J.D., and Gross, J.J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. Neuroimage *23*, 483-499.

Ollinger, J.M., Corbetta, M., and Shulman, G.L. (2001). Separating processes within a trial in event-related functional MRI. Neuroimage *13*, 218-229.

Phan, K.L., Wager, T., Taylor, S.F., and Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. Neuroimage *16*, 331-348.

Phan, K.L., Wager, T.D., Taylor, S.F., and Liberzon, I. (2004). Functional neuroimaging studies of human emotions. CNS Spectr *9*, 258-266.

Phelps, E.A. (2006). Emotion and cognition: insights from studies of the human amygdala. Annu Rev Psychol *57*, 27-53.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci *10*, 59-63.

Price, D.D., Craggs, J., Verne, G.N., Perlstein, W.M., and Robinson, M.E. (2007). Placebo analgesia is accompanied by large reductions in pain-related brain activity in irritable bowel syndrome patients. Pain *127*, 63-72.

Rainville, P., Duncan, G.H., Price, D.D., Carrier, B., and Bushnell, M.C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science *277*, 968-971.

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2003). The neural basis of economic decision-making in the Ultimatum Game. Science *300*, 1755-1758.

Scheier, M.F., and Carver, C.S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. Cognitive Therapy and Research *16*, 201-228.

Seymour, B., O'Doherty, J.P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K., and Dolan, R. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. Nat Neurosci *8*, 1234-1240.

Shepard, R.N. (1980). Multidimensional Scaling, Tree-Fitting, and Clustering. Science *210*, 390-398.

Shrout, P.E., and Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. Psychol Methods *7*, 422-445.

Stern, E.R., Wager, T.D., Egner, T., Hirsch, J., and Mangels, J.A. (2007). Preparatory neural activity predicts performance on a conflict task. Brain Res *1176*, 92-102.

Struyf, A., Hubert, M., and Rousseeuw, P. (1996). Clustering in an Object-Oriented Environment. Journal of Statistical Software *1*, 1–30.

Sumner, P., Nachev, P., Morris, P., Peters, A.M., Jackson, S.R., Kennard, C., and Husain, M. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. Neuron *54*, 697-711.

Tugade, M.M., Fredrickson, B.L., and Feldman Barrett, L. (2004). Psychological Resilience and Positive Emotional Granularity: Examining the Benefits of Positive Emotions on Coping and Health. Journal of Personality *72*, 1161-1190.

Urry, H.L., van Reekum, C.M., Johnstone, T., Kalin, N.H., Thurow, M.E., Schaefer, H.S., Jackson, C.A., Frye, C.J., Greischar, L.L., Alexander, A.L., and Davidson, R.J. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. J Neurosci *26*, 4415-4425.

van Reekum, C.M., Johnstone, T., Urry, H.L., Thurow, M.E., Schaefer, H.S., Alexander, A.L., and Davidson, R.J. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. Neuroimage *36*, 1041-1055.

van Snellenberg, J.X., and Wager, T.D. (in press). Cognitive and Motivational Functions of the Human Prefrontal Cortex. In A tribute to Alexander Luria, E. Goldberg, and D. Bougakov, eds.

Wager, T.D., Barrett, L.F., Bliss-Moreau, E., Lindquist, K., Duncan, S., Kober, H., Joseph, J., Davidson, M., and Mize, J. (in press). The Neuroimaging of Emotion. In Handbook of Emotion, M. Lewis, ed.

Wager, T.D., Lindquist, M., and Kaplan, L. (2007a). Meta-analysis of functional neuroimaging data: Current and future directions. Social, Cognitive, and Affective Neuroscience *2*, 150-158.

Wager, T.D., Reading, S., and Jonides, J. (2004a). Neuroimaging studies of shifting attention: a meta-analysis. Neuroimage *22*, 1679-1693.

Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M., and Cohen, J.D. (2004b). Placebo-induced changes in FMRI in the anticipation and experience of pain. Science *303*, 1162-1167.

Wager, T.D., Scott, D.J., and Zubieta, J.K. (2007b). Placebo effects on human mu-opioid activity during pain. Proceedings of the National Academy of Sciences *104*, 11056-11061.

Wager, T.D., and Smith, E.E. (2003). Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci *3*, 255-274.

Whalen, P.J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. Current Directions in Psychological Science *7*, 177-188.

Neuron, volume *59*

Supplemental Data

Prefrontal-Subcortical Pathways Mediating Successful Emotion Regulation

Tor D. Wager, Matthew L. Davidson, Brent L. Hughes, Martin A. Lindquist, and Kevin N. Ochsner

List of Supplemental Tables and Figures

Supplemental Text

Table S1

	Coordinates			Cluster size		a path		ab path		Num. voxels	
Name	**x**	**y**	**z**	**Vox**	**Vol**	**Z**	**p**	**Z**	**p**	**P<.001**	**P<.005**
Mediated by nuc. accumbens											
VStr/basal forebrain	28	7	-18	25	1329	3.46	0.0005	3.48	0.0005	4	25
VMPFC	17	31	-14	303	16112	3.58	0.0003	3.58	0.0003	40	303
L VLPFC	-48	28	-14	31	1648	3.57	0.0004	3.55	0.0004	4	31
R VLPFC	41	48	0	13	691	3.35	0.0008	3.18	0.0015	0	13
Rostral MPFC	3	58	9	136	7232	3.58	0.0003	3.58	0.0003	24	136
R aPFC	21	55	18	12	638	3.56	0.0004	3.58	0.0003	7	12
L DLPFC	-31	24	45	15	798	3.55	0.0004	3.28	0.001	0	15
DMPFC	10	28	45	58	3084	3.58	0.0003	3.52	0.0004	8	58
Mediated by left amygdala											
R VLPFC	48	31	-9	17	904	3.51	0.0004	-3.448	0.0006	1	17
R mid-lateral OFC	28	31	-14	3	160	3.54	0.0004	-3.581	0.0003	3	3

Figure S1

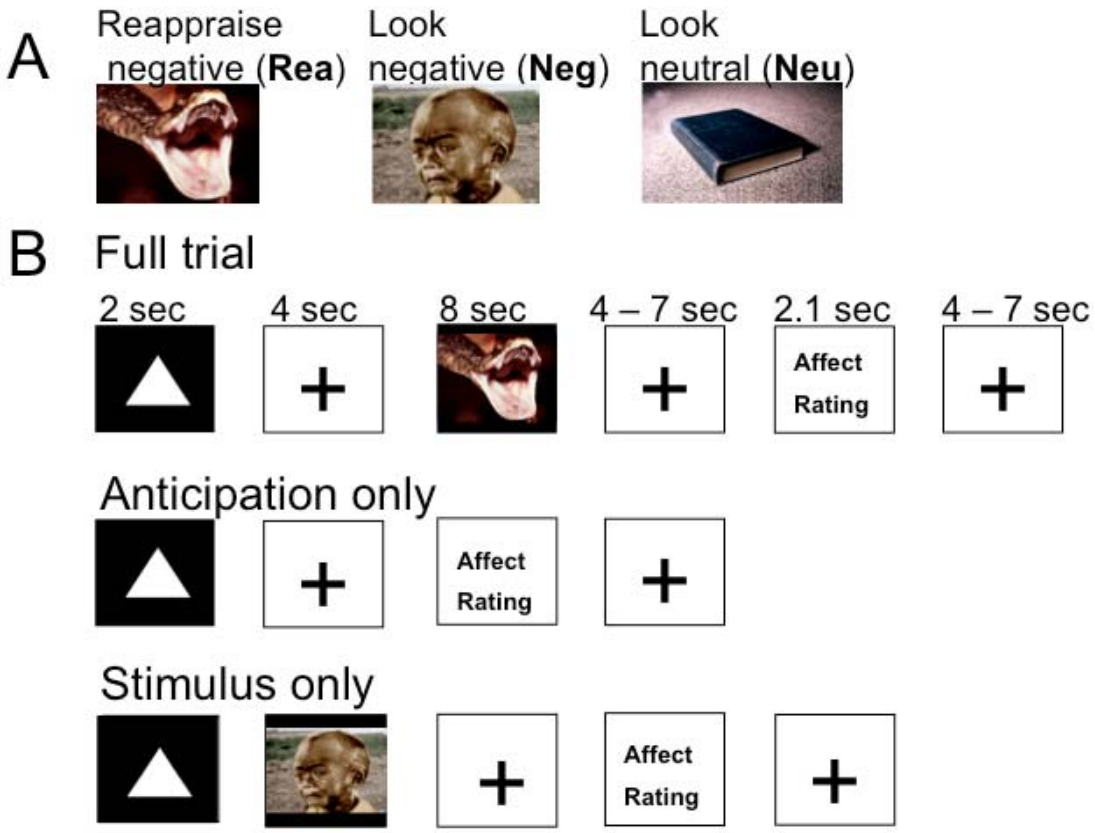


Figure S1. Task design. The design includes three trial types: Anticipation and stimulus trials, Anticipation only trials, and Stimulus only trials, for each of the 3 conditions (Reappraise Negative [ReappNeg], Experience Negative [LookNeg], and Experience Neutral [LookNeu]).

Figure S2

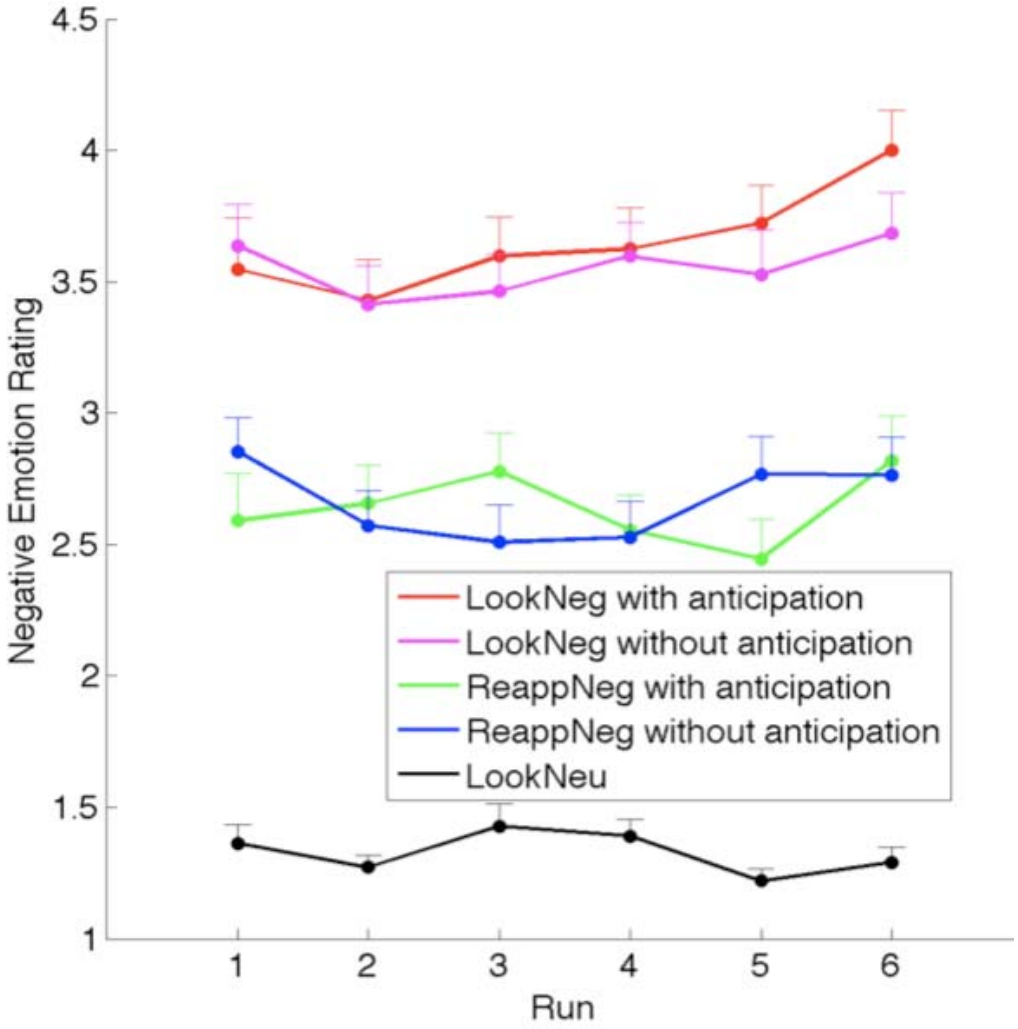


Figure S2. Emotion rating data. Group average ratings (y-axis) as a function of scanning run (x-axis) and task condition (lines). There was a robust effect of viewing negative vs. neutral images and a robust effect of reappraisal, both of which were sustained across runs. Neither interacted with whether the trial included a 4 sec anticipation period.

Figure S3

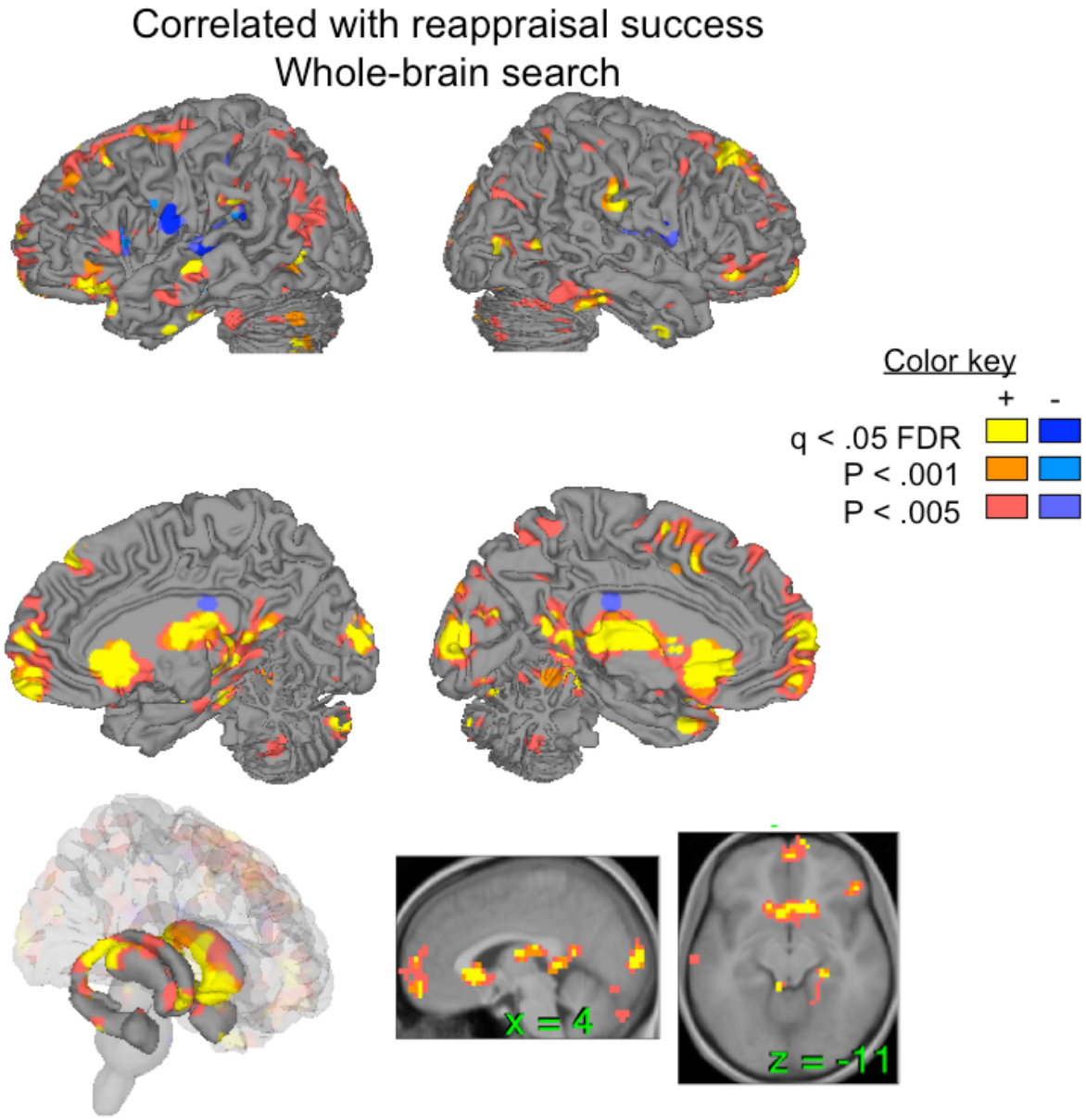


Figure S3. Whole-brain correlations between reappraisal activation [ReappNeg - LookNeg] and reappraisal success in reported experience [LookNeg – ReappNeg]. Positive correlations are shown in red/yellow, and negative correlations are shown in blue. Positive correlations indicate a greater relative increase in activity for participants who report more successful reappraisal , and negative correlations indicate a greater relative decrease in activity (LookNeg - ReappNeg) for participants who report more successful reappraisal. Thresholds are shown in the color key on the figure.

Figure S4

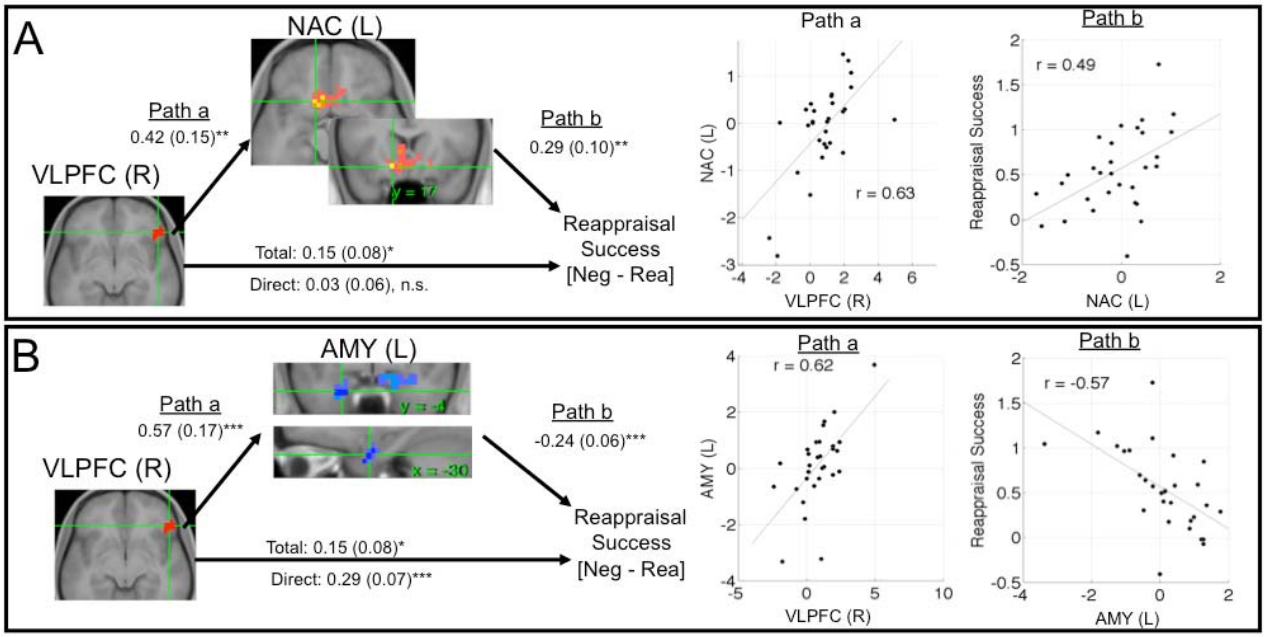


Figure S4. Mediation analyses for amygdala and nucleus accumbens regions of interest. The mediation effect parametric mapping identified regions that were mediators of the relationship between right ventrolateral prefrontal cortex (vlPFC) and reappraisal success by searching over voxels and performing mediation tests one voxel at a time. This figure shows statistics for the path models, averaging over voxels in each mediating region, for significant left nucleus accumbens/ventral striatal and left amygdala regions. These results differ slightly from the ones presented in Figure 3 because here each region is subjected to a separate mediation analysis, whereas in the analysis shown in Figure 3, they were included in the same mediation model.

Figure S5

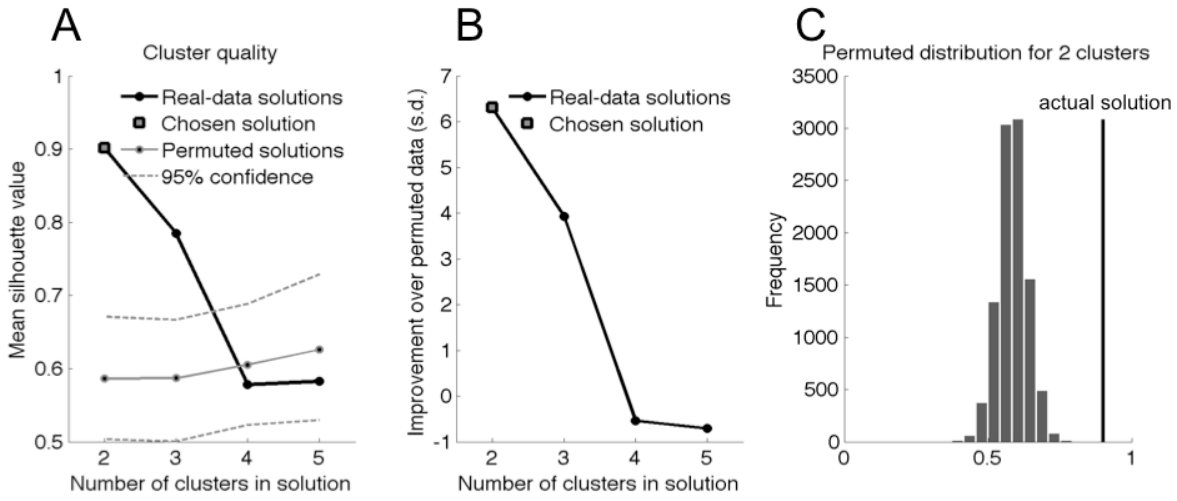


Figure S5. Results for cluster analysis examining grouping of mediators into functional networks. A) Cluster quality (y-axis, see Experimental Procedures) as a function of number of clusters in the candidate solution. Clusters in this context refer to interconnected networks of brain rgions. Higher values indicate tighter clustering of the data. The solid black line shows the actual data, with the chosen solution (2-cluster solution) marked with a gray square. The thin black line shows the average quality for permuted data, with 95% confidence intervals marked by dashed lines. Permutation disrupted the relationships among variables, so that there was no true grouping of the data into clusters, so that cluster quality under the null hypothesis of no grouping could be assessed. B) Z-scores (y-axis) for the real-data cluster quality relative to the cluster quality of permuted data. C) Distribution of null-hypothesis cluster quality (x-axis; frequency on y-axis) for the chosen 2-cluster solution. The real-data solution is marked with a vertical line.

Supplemental Discussion

Interpreting suppression effects in path models

A suppression effect occurs when the indirect pathway and the direct pathway have opposite signs—that is, when two variables independently influence an outcome variable (Y), but with effects of opposite signs. It might be best illustrated with an example from another domain. For example, increased use of technology may have both positive and negative effects on life satisfaction (Y), for different reasons. These effects tend to cancel each other out, resulting in a weak overall technology -> satisfaction relationship, unless the mediating variables with opposite signs can be identified: One might discover a positive pathway from technology -> connect with family and friends -> increased satisfaction, and a negative pathway from technology -> reachable by boss at home -> decreased satisfaction. Controlling for the suppressor variable (reachable by boss) would allow other positive relationships and any positive direct technology -> satisfaction effect to emerge.

There are many examples of suppression effects in published literature, and they can often uncover strong direct relationships that are masked by confounding suppressor variables or explain results that are initially counter-intuitive. As an example of the first case, Goldberg (Goldberg et al., 1996) found that a steroid use-prevention program reduced intentions to use steroids. They also found that greater reason for using steroids was related to increased intentions to use. Both of these effects are intuitive. However, as MacKinnon et al. (MacKinnon et al., 2000) describe in their re-analysis of these data, those who had greater reason to use steroids tended to enroll in the program, creating a positive link between reasons for using steroids and program enrollment. Thus, the indirect pathway connecting program enrollment, more reasons for using steroids, and increased intentions to use partially canceled out the beneficial direct effects of the program. After controlling for the confounding suppressor (reasons to use), the beneficial effects of the program grow stronger.

As an example of the second case, Guber (1999) describes a dataset drawn from the 1997 Digest of Education Statistics on the relationship between state spending on public schools and SAT scores. Surprisingly, the original relationship between spending and SAT scores across U.S. states was negative: More spending predicted significantly lower scores. However, a likely cause is that a third variable was acting as a confounding suppressor variable. High-spending

states tended to require all students to take the SAT, even the academically poor ones; thus, percentage of students taking the test is a strong negative predictor of average SAT. When percent taking the test was included as a mediator, the direct relationship between spending and SAT became significantly positive, in line with what one might expect.

In the present study, the amygdala may appear to be a complete mediator of the PFC-reappraisal success correlation (in the sense that the direct relationship is not significant) when it is the only mediator in the model, but that is partially because there is still much unexplained variance in this model that masks the direct effect. Including the nuc. accumbens (NAC) in the model explains additional variance, allowing the direct effect to reach statistical significance.

Additional implications of dual routes for reappraisal effects on emotion: how we measure changes in emotional responding

The finding that fMRI responses in both the amygdala and NAC/VS were meaningfully related to emotional experience has implications for both understanding amygdala and striatal function and the use of self-reported emotion in neuroscience research more generally. Consider first that although the amygdala has been broadly implicated in negative emotional responses, ranging from the acquisition of conditioned fear (Phelps, 2006) to hyperactivity in a range of anxiety disorders (Etkin and Wager, 2007), there has been reason to doubt that amygdala activation reflects experiential changes in all situations in which it has been activated (Anderson and Phelps, 2002; Barrett and Wager, 2006; Wager et al., in press). In part this may be due to the fact that the amgydala is a complex, heterogeneous group of structures, only some of which have been linked to emotional experience in some situations. Several meta-analyses from our group have shown that although the superior amygdala is activated both by stimuli that do (e.g. a small child crying hysterically) and stimuli that do not (e.g. a fearful facial expression) generate negative emotional experiences, it is more consistently activated by the stimuli that *do not* elicit experience (Phan et al., 2002; Phan et al., 2004; Wager et al., in press). This finding, and the individuals results on which it is based, have been taken to suggest that the amygdala may play a role in the perceptual detection and encoding of affectively relevant stimuli, which may or may not have a direct impact on affective experience (Anderson and Phelps, 2001, 2002; Barret et al., 2007; Wager et al., in press; Whalen, 1998). Therefore, the finding that amygdala activity mediates negative emotional experience *in this task context* is a significant finding, and it

provides support for the notion that prefrontal-amygdala correlations in other reappraisal studies that use the amygdala as an outcome variable – but did not measure emotional experience (or any other behavioral index of emotional response) (Johnstone et al., 2007; Urry et al., 2006; van Reekum et al., 2007) – may in fact be relevant to shaping emotional experience.

This argument leads to the broader implication of the amygdala-experience relationship observed here. One reason some of the aforementioned papers have not measured emotional experience is because of concerns that it may not be a wholly reliable or valid index of emotional responding. As noted at the outset, we recognized the necessity of behavioral measures for constraining functional interpretations of neural indicators (Poldrack, 2006), and selected self-reports as our behavioral measure because it can be reliably collected in the scanner environment, provides clear valence information (unlike most autonomic measures), and has been shown to predict mental and physical health outcomes (Brosschot et al., 2006; Gross and Munoz, 1995; Moskowitz, 2003; Scheier and Carver, 1992; Tugade et al., 2004). The fact that changes in self-reported negative emotion were mediated by activity in brain systems thought – on indepedent *a priori* grounds – to be important for emotional appraisals provides additional support for self-report as a useful measure for studies of emotion and its regulation.

Additional implications for the role of prefrontal cortex in emotional appraisal: the nature of correlations between PFC and amygdala

An interesting difference between this study and some previous studies of reappraisal is that we found positive correlations between vlPFC and amygdala, whereas several studies mentioned above found negative correlations. One reason for these differences may have to do with the structure of the reappraisal task as used by different investigators. In the Urry et al. and Johnstone et al. papers, the appraisal process was permitted to evolve for 4 s before presenting a reappraisal cue, perhaps separating in time an negative initial appraisal process that involves positive vlPFC-amgydala connectivity and a subsequent positive reappraisal process that involves negative vlPFC-activity. However, it is also possible that amygdala activity can be related in complex ways to appraisal. Kim et al., for example, found amygdala increases when participants increased positive emotional responses (Kim and Hamann, 2007), and studies of autobiographical memory retrieval have found co-activation of amygdala and vlPFC (Greenberg et al., 2005; Maguire and Frith, 2003). This raises the possibility that reappraisal cues presented

in advance of pictures in our paradigm may have strengthened participants' ability to engage in meaning-based reinterpretations, shifting frontal-amygdala connectivity positively. The present data cannot disentangle these and other potential explanations, but it opens the way for a more detailed mapping of effective connectivity as a function of specific strategies and stimulus types.

Additional implications of network analyses for understanding cognition-emotion interactions

The first network involved the NAC/VS (Figure 5B) and included three additional regions that have been implicated in action selection and memory. However, we also note that the NAC/VS region we identified extended into the subgenual cingulate, which has recently been shown to play an important role in the modulation of depression and mood (Drevets, 2001; Johansen-Berg et al., 2007; Mayberg et al., 1999); high-resolution studies are needed to dissociate the contributions of NAC and subgenual cingulate to the pathways reported here. One additional region in this network was the pre-SMA, which has been implicated in multiple cognitive control processes (van Snellenberg and Wager, in press; Wager et al., 2004a; Wager and Smith, 2003) and in the energization of internally driven processes that lead to effective intentional response selection (Alexander et al., 2007; Cunnington et al., 2005; Isoda and Hikosaka, 2007; Sumner et al., 2007) (Lau et al., 2004). A second region was the precuneus, which has been implicated in the control and switching of attention among objects and object features (Barber and Carter, 2005; Wager et al., 2004a) and in monitoring of the internal environment and "self" (Cavanna and Trimble, 2006; Gusnard and Raichle, 2001) and episodic memory retrieval (Lundstrom et al., 2005). Finally, a third region encompassed the retrosplenial cingulate cortex and cingulate isthmus, which have been implicated in autobiographical memory (Maddock, 1999) and may provide an anatomical bridge from the precuneus to medial temporal lobe structures that have well-known roles in the formation of declarative memories (Davachi, 2006). The positive association of vlPFC with activity in this network suggests that the selection of reappraisal-appropriate information from memory may be used to construct a positive construal of the situations depicted in negative images (Ochsner, 2004), and that this can mediate reappraisal success. Although speculative, this interpretation is consistent with what is known about the regions involved in this network and is consistent with the connectivity pattern we observed, which indicated a pathway from vlPFC to pre-SMA, which served as a network "hub" that is connected to the precuneus and retrosplenial cortex/cingulate isthmus, which in turn

connected to the NAC/VS. In this network context, one way of reappraising an aversive event can be seen to involve the interactions of vlPFC with a network of regions that together represent a positive, approach motivated appraisal of the event.

Here it should be noted that our search for frontal "mediated" regions suggested that the medial PFC (including dorsal, ventral, and rostral divisions) was functionally connected to this network. Though we found strong correlations between rostral MPFC and reappraisal success, it was not a primary focus here because we did not find overall reappraisal-related activation in our group of participants. However, its focus in previous papers (Johnstone et al., 2007; Urry et al., 2006) suggests that it may play an important role in reappraisal, and in our study it seemed to be related to the enhancement of NAC/VS activity that, in turn, predicted reappraisal success.

The second network involved the amgydala and included three additional anatomically interconnected areas that have been commonly associated with negative appraisals and negative emotion more generally (Figure 5B). The first region was the right lateral orbitofrontal cortex (lOFC), which has been associated with diverse forms of valenced affective experience (Berridge and Kringelbach, 2008; Wager et al., in press) and related motivational processes such as updating behavior on the basis of negative feedback (O'Doherty, 2003; O'Doherty et al., 2003) and negative (aversive) prediction errors (Seymour et al., 2005). The second was the anterior insula (AI), which is activated by aversive stimuli of various kinds and is most commonly associated with negative emotional experience in normal emotion (Wager et al., in press) and increased emotional activation across several kinds of anxiety disorders (Etkin and Wager, 2007). This region has been called primary "interoceptive" cortex (Craig, 2003), though it appears to be activated by aversive social experiences as well (Sanfey et al., 2003). The third was the rostral dorsal cingulate (rdACC), a region often activated when cognitive expectancies modulate affective processes. For example, hypnosis (Faymonville et al., 2000; Rainville et al., 1997) and placebo treatments (Lieberman et al., 2004; Price et al., 2007; Wager et al., 2004b) that lead to reduced pain have consistently reduced rdACC activity. A fourth region in the subthalamus was also included in this network, but we do not speculate on its interpretation here. The positive association of vlPFC activity with activity in this network is consistent with the idea that reappraisal success can also hinge on limiting activity in a network of structures important for representing negative affective states.

Supplemental Experimental Procedures

Reappraisal training procedures

Prior to scanning, participants completed a training session on the reappraisal strategy. Previous work on the reappraisal process has established that an effective means of down-regulating negative emotion is by re-interpreting the affects, dispositions, outcomes or contexts for the actors and actions shown in less negative ways. Prior to scanning, participants were trained in the use of this strategy, which is known as reinterpretation (Ochsner and Gross, 2008). For example, if participants viewed an image of a crying baby in a barren landscape, they might imagine that the mother has simply gone to do an errand and will return soon. Different specific reappraisals – all of the same basic kind – were generated for each image. As described previously (Ochsner et al., 2002; Ochsner et al., 2004), this method is meant to capture the fact that in everyday no single type of reinterpretation is expected to be applicable to all life events – or here, all photos.

The training session consisted of two parts. In the first part, participants were asked to memorize the cue-task condition associations (e.g., circle with LookNeu, etc.). Participants were subsequently quizzed until they achieved 100% accuracy on 10 consecutive trials of each type. In part two, participants completed 7 sample trials with the experimenter present (1 Look Neutral, 2 Look Negative, 4 Reappraise Negative). On reappraisal trials they were given feedback on the appropriateness of their reappraisals for each image. During this time participants also were reminded not to look away from images or distract themselves with irrelevant and/or positive thoughts unrelated to the context of the image. After appropriate coaching to ensure that participants could reinterpret images quickly and effectively, the training ended with the completion of 18 practice trials (6 Look Neutral, 6 Look Negative, 6 Reappraise Negative) on their own. Images used during training were different from those used during the subsequent test. Eye position was monitored during scanning by viewing an image of the right eye projected onto a screen in the scanner control room (I-SCAN, Inc.). No subjects closed their eyes or averted their gaze during image viewing.

fMRI task design

Previous studies of reappraisal have not separated brain activity related to anticipation

and instruction processing, stimulus viewing, and picture rating, and a goal of our task design was to provide the ability to estimate separately brain activation magnitude related to each of these three phases of the image viewing and rating process. To accomplish this, a partial trial design was employed (Ollinger et al., 2001; Stern et al., 2007). Within each task condition, LookNeu, LookNeg, and ReappNeg, three different trial types were used: full trials, anticipation-only trials, and stimulus-only trials (see Supplementary Figure 1).

On full trials, a 2 sec condition cue was followed by a 4 sec anticipatory interval during which a fixation cross was presented on the screen. The image was subsequently presented for 8-sec. Following image presentation, a fixation cross was presented during a 4 or 7 sec jittered inter-stimulus interval (ISI; uniform distribution of 4 and 7 sec intervals). Following the ISI period, the words "how negative do you feel?" appeared on-screen for 2.1 sec, and participants were asked to rate negative affect on a five-point scale by pressing a button with one of five fingers on a button-response unit (1 = "not at all negative," indicated by a thumb button press, up to 5 = "extremely negative," indicated by a fifth-finger button press). Following the rating, a 4 or 7 sec jittered ISI concluded the trial.

The trial was identical to the full trial through the anticipation interval. Instead of an image presentation, participants were asked to rate the negative feelings experienced during anticipation on the same five-point scale. The stimulus only trials were identical to the full trials, except that the 4 sec anticipation interval was omitted. This design allowed us to construct predictors for Cue-, Anticipation-, and Image-related brain activity related to each task condition in the General Linear Model (GLM) that were uncorrelated enough to provide efficient estimates of activation in each task condition x trial phase combination (see below). In addition, responses during picture or anticipation rating could be estimated separately, thus permitting the parsing of brain activity during trial phases of interest from brain activity related to the reporting and button-press processes.

Subjects completed 6 functional runs of 18 trials each for a total of 108 trials. Each run included 6 trials of each condition (Reapp Neg, Look Neg, Look Neutral) and trial type (full trial, ant only trial, stim only trial).

In this report, we compared brain responses during the viewing of aversive images in the ReappNeg vs. LookNeg conditions, as estimated controlling for activity during other phases of the trial (Cue, Anticipation, and Report). Negative emotion ratings did not differ significantly

based whether the anticipation period was included or not in either Look or Reappraise conditions (see Supplementary Figure 2). For LookNeg, the average rating difference for full trial vs. stimulus only trials was 0.10, paired $t(35) = 1.58$, $p > 0.10$. For ReappNeg, the difference was -0.02, paired $t(35) = -0.29$, $p > 0.7$. Subsequent analyses of behavioral and brain data were conducted on averages across full trials and stimulus-only trials. Likewise, reported reappraisal success (LookNeg – ReappNeg) was stable across time (Success x Run interaction $F(5, 172) = 1.02$, $p = .41$), and ratings for all image types were likewise stable across time. Thus, subsequent analyses were also conducted on averages across runs.

Image processing and data analysis

Preprocessing. Functional images were slice-time and motion corrected using FSL (FMRIB Centre, University of Oxford). Structural T1-weighted images were coregistered to the first functional image for each subject using an iterative procedure of automated registration using mutual information coregistration in SPM2 and manual adjustment of the automated algorithm's starting point by a trained analyst until the automated procedure provided satisfactory alignment. Structural images were normalized (spatially warped) to a standard template brain (the MNI avg152T1.img) using SPM2 software (Wellcome Department of Cognitive Neurology, UCL) using default options (7 x 8 x 7 nonlinear basis functions), and the warping parameters were applied to functional images for each subject. Normalized functional images were interpolated to 2 x 2 x 2 mm voxels and spatially smoothed with a 6-mm Gaussian filter.

First-level GLM model. First-level GLM analysis for each participant was performed using SPM2 (Friston, Jezzard, & Turner, 2004). Effects were modeled as a boxcar regressor convolved with a canonical hemodynamic response function (double-gamma) for the 2 sec cue period, 4 sec anticipation period, 8 sec stimulus viewing period during which subjects either attend or reappraise, and 2.1 sec rating period separately. Twelve regressors were specified, for cue-related responses, anticipation-related responses, image-viewing related responses, and rating-related responses in each of the three task conditions (LookNeu, LookNeg, and ReappNeg). Separate sets of regressors were specified for each of the 6 scanning runs for each subject. In addition, regressors specifiying a high-pass filter with a 120 sec cutoff were included

to model low-frequency drift. Since our primary concern was group statistics, no autoregressive (AR) model was used, which are unbiased and valid even without AR modeling.

Importantly, the trial design resulted in regressors that were essentially uncorrelated across conditions of interest. The critical feature of the design for the results we present here is that the regressors related to Image Viewing for each task condition are not collinear with combinations of other regressors, so that their estimates are stable and efficient. The average correlations between each Image-Viewing related regressor and other regressors (across other regressors and subjects) were r = 0.067, r = 0.063, and r = 0.062 for LookNeu, LookNeg, and ReappNeg. The correlations were very similar across both regressors and subjects as well: The maximum correlations with any other regressor for any subject were r = 0.074, r = 0.067, and r = 0.070, for each task condition, and the minimum correlations were r = -0.17, r = -0.18, and r = -0.17. Variance inflation factors reflect the overall multicolinearlity in a design matrix for each regressor and were low (< 2) for all twelve regressors and for all subjects, indicating that the task design was effective in providing stable and efficient estimates of brain activity related to each task phase (cue, anticipation, image viewing, and rating) and task condition.

Contrasts. Activation estimates were obtained for each subject using SPM2, and contrasts across conditions were estimated and analyzed in a second-level group analysis treating subject as a random effect. The contrast of interest in this report was [ReappNeg image viewing – LookNeg image viewing]. Positive contrast values indicate greater relative activity during reappraisal vs. natural experience of aversive images, and negative values indicate greater relative activity during experience (i.e., the reverse subtraction). Both effects were analyzed and reported here. As our main *a priori* hypotheses concerned changes during image viewing, contrasts related to other phases of the trial (Cue, Anticipation, and Report) will be addressed in subsequent papers. This choice serves to avoid undue complexity and preserve clarity in the present report.

Group analysis (Analysis steps 1-2). The second level random-effects analysis was performed using robust regression, a technique that both increases statistical power and decreases false positive rates in the presence of outliers (Wager, Keller, Lacey, & Jonides, 2005). Reported reappraisal success was calculated for each subject as the average difference in negative affect reports for [LookNeg – ReappNeg]. The second-level design matrix included two regressors: one corresponding to reappraisal success, and the other an intercept term.

Reappraisal success scores were centered by subtracting the mean, allowing the intercept term to be interpreted as the population estimate for reappraisal-induced activation ([ReappNeg – LookNeg]) for a subject who shows average reappraisal success. The interpretation of the reappraisal success regressor is the change in reappraisal-induced activation as a function of reappraisal success, i.e. the activation-reappraisal success relationship. The advantage of including reappraisal-induced activation ([ReappNeg – LookNeg]) in the model is that it accounts for known sources of individual variation when testing the significance of average activation contrast values. Thus, for voxels that do show a brain activity-reappraisal success relationship, this model has greater sensitivity to detect overall activation compared with an intercept-only model (which is what is typically performed, e.g., in SPM software).

An anatomically defined gray matter mask was created based on the Montreal Neurologic Institute (MNI) avg152T1 template (smoothed with an 8 mm FWHM filter and thresholded at a value of 0.5) and explicitly specified during analysis.

Localization of results. Normalized structural T1 images were averaged across participants to create an anatomical underlay for visualizing significant regions of activation, and for visually assessing normalization quality. In our experience, this is advantageous because the quality of nonlinear warping can vary across brain regions, resulting in greater differences between the standard brain and participants' actual T1 images in ventral subcortical regions in particular. We assessed activation locations based on identified structural landmarks in our participants (e.g., amygdala gray matter, striatal gray matter) using several brain atlases (Haines, 2000; Mai et al., 2004; Martin, 1996), rather than relying on standardized Talairach or MNI coordinates, which can be misleading in some brain regions depending on warping quality. However, we provide MNI coordinates for reference and use in future studies and meta-analyses.

Mediation Effect Parametric Maps (Analysis steps 3, 4a). The Mediation Effect Parametric Map (MEPM) analysis is based on a standard 3-variable path model (Baron and Kenny, 1986) with a bootstrap test for statistical significance (Efron and Tibshirani, 1993; Shrout and Bolger, 2002). A test for mediation tests whether a covariance between two variables (X and Y) can be explained by a third variable (M). A significant mediator is one whose inclusion as an intermediate variable in a path model of the effects of X on Y significantly affects the slope of the X – Y relationship; that is, the difference ($c - c'$) is statistically significant (see Results for nomenclature). The test of mediation is fundamentally different from a moderation or "psycho-

physiological interaction" (PPI) effect, in which the *level* of a moderating variable M predicts the strength of the X –Y relationship. A moderator interacts with the X –Y relationship, whereas a mediator explains it. Thus, the mediation test (not the moderation test) is critical for localizing and testing functional pathways that span more than two regions. Brain regions that are *mediators* are candidates for links in functional pathways that relate brain activity in multiple regions to behavior and other outcomes.

In the current application, we used right VLPFC activity in the [ReappNeg – LookNeg] contrast as the X variable and reappraisal success as the Y variable. Thus, the X-Y relationship (and the *c* path) is the linear association between prefrontal increases during reappraisal and reported emotion decreases. Thus, the association is the correlation in contrast estimates across subjects in this case, rather than within-subject time series values. The MEPM strategy is to search for voxels in the brain that mediate that relationship, so that X and Y are specified as in Figure 1, but the mediator (M) is unknown. In this case, we first searched for voxels whose [ReappNeg – LookNeg] contrast values mediate the VLPFC – reappraisal success relationship within amygdala and NAC, and then conducted a whole-brain search to test for additional mediating brain regions that may form functional networks with our *a priori* target regions.

More formally, the mediation test can be captured in a system of three equations:

$$y = cx + e_y$$

$$m = ax + e_m$$

$$y = bm + c'x + e'_y$$

where *y*, *x*, and *m* are *n* (participants) x 1 data vectors containing the outcome (y, reappraisal success), the predictor (x, VLPFC), and data from a candidate mediating voxel (m, [ReappNeg – LookNeg] contrast values). e_y , e_m, and $e`_y$ vectors denote residual error for the outcome and mediator controlling for *x* and the outcome controlling for *x* and *m*, respectively. The *a* path is the estimated linear change in *m* per unit change in *x* (e.g., the slope of the VLPFC-mediator relationship). The *b* path is the slope of the mediator-outcome relationship controlling for *x* (mediator to reappraisal success, controlling for VLPFC). The *c* and *c'* paths are as described above.

Statistical tests on *a* and *b* path coefficients assess the significance of each relationship. In addition, a statistical test of $(c - c')$ can be performed by testing the significance of the product of the path coefficients *a***b*. A positive mediator is one involved in a pathway that has a

net positive effect on the outcome (e.g., *a*b* is positive, greater reappraisal success in this case). A negative mediator has a net negative effect (lower reappraisal success). We test the significance of *a*b* using the accelerated, bias-corrected bootstrap test (Efron and Tibshirani, 1993) with 10,000 bootstrap samples to test each of the *a*, *b*, and *a*b* path coefficients at each voxel, saving maps of path coefficients and bootstrapped P-values for each effect. Custom Matlab code (Mathworks, Natick, MA, R2007b) was written to perform these tests and optimize them for speed, so that the whole-brain MEPM maps could be accomplished in ~12 hours on a 4-processor Intel Macintosh computer with 2 dual-core Intel Xeon processors (see Author Note for software download information).

If multiple regions are included as mediators (as in analyses with amygdala and NAacc as independent mediators), then additional equations specifying *a* paths for each additional mediator are added, e.g., $m_2 = a_2 x + e_{m2}$ for a second mediator, and the equation that assesses the direct effects of the mediators is modified to include the additional mediators, e.g., $y = bm + b_2 m_2 + c' \quad x + e' \quad _y$ for a 2-mediator model. Thus, the *b* path for each mediator tests whether it is significantly related to the outcome controlling for all other variables.

Though only a significant *a*b* product is required to provide evidence for a mediation effect, for significance in both ROIs and in the whole-brain analysis, we required that each of the paths *a*, *b*, and *a*b* be significant at $p < .005$, two-tailed, with three contiguous voxels. The conjunction across the three effects locates regions with the strongest evidence for both individual pathway links and the total mediation effect. Our choice of threshold was motivated by 1) the desire to use a standard threshold for comparability with other published results ($p < .005$ is one of the most common (Wager et al., 2007a)); 2) Control of false positives; and 3) Sensitivity, or the ability to detect many of the regions that show true mediation effects. The latter is important because using a threshold low enough to afford relatively high sensitivity enhances the ability to make inferences based on the pattern of results across the brain, and to meaningfully aggregate positive and negative findings across studies.

Mediation analysis is a tool for testing particular pathway relationships, but complementary tools are necessary for identifying functional networks of broadly interconnected regions. Principal components analysis (PCA), independent components analysis, and related tools provide complementary perspectives on neuroimaging data by identifying broadly interconnected networks. Though the MEPM analysis searches for mediating regions, we do not

conceptualize these regions as each constituting a separate functional path; rather, these individual networks are likely to be grouped into distributed functional networks that operate as a unit to mediate prefrontal-experience relationships. The analyses described below address this second point.

Region of interest (ROI) analyses and multiple comparisons correction. Four ROIs were drawn on the group-averaged, normalized T1-weighted images from our sample using custom software (SCAN Lab tools, T.D.W.), using several brain atlases for reference (Haines, 2000; Mai et al., 2004; Martin, 1996). The ROIs were in left and right amygdala (108 and 79 [3.5 x 3.5 x 4.5] mm voxels, respectivley) and left and right NAC/VS (52 and 53 voxels, respectively).

We assessed the family-wise error rate (FWER) for the mediation search analyses in each ROI by using a permutation test ((Nichols and Hayasaka, 2003)), which provides estimates of the chances of obtaining a false positive result anywhere in the ROI, accounting for the observed correlations among voxels. We permuted the rows of the 30 (subjects) x *k* (voxels) matrix of subject contrast values, preserving spatial structure, and re-computed the mediation test (including the significance of *a, b,* and *a*b* effects, and their conjunction) for each of 2000 permutations. The primary chosen threshold of $p < .005$ controlled the false positive rate below $p < .05$ in each ROI. That is, under the null hypothesis of no true relationships among X, M, and Y variables, the chances of a false positive conjunction result (significant *a, b,* and *a*b*) anywhere within a given ROI were less than 5%. Based on the permutation test, corrected p-values were $p < 0.03$ and 0.02 for the L/R amygdalae, and $p < .008$ and .002 for the L/R NAC/VS, respectively. This threshold also provided adequate correction for multiple comparisons for each of the *a* and *b* pathways independently. For the left and right amygdala, *a*: $p < .075$, *b*: $p < .06$ (one-tailed, corrected). For the left and right NAC/VS, *a*: $p < .055$, *b*: $p < .04$.

Clustering of mediators and functional networks (Analysis step 4b).

An advantage of the clustering procedure is that a nonparametric permutation test can be used to assess whether there is significant grouping of regions into clusters (networks), and the null hypothesis that regions are operating either independently or as a single large network can be rejected. Secondly, clustering (as opposed to PCA/ICA alone) provides a way of identifying networks that is stable with respect to rotation of the data in multidimensional space. Rotational indeterminacy and instability have plagued PCA/ICA-based and factor analytic approaches.

Briefly, we first defined mediating regions as sets of contiguous voxels that survived the conjunction test (*a*, *b*, and *a*b,* each at $p < .005$, with a 3-voxel extent threshold) as described above. We extended the regions to include contiguous voxels at $p < .05$ in the conjunction, and calculated the average contrast activity for each subject within each region. We then removed linear effects of right VLPFC activity from each region using robust regression, so that the networks captured relationships among regions with respect to their direct relationship with reappraisal success.

In order to achieve stable clustering in a high-dimensional dataset, it is typical to use a data-reduction algorithm (e.g. PCA) prior to clustering in order to remove dimensions with near-zero variance from the dataset. Here we used non-metric multidimensional scaling (NMDS), which makes fewer assumptions than PCA, as the dissimilarities are not assumed to reflect distances in a Euclidean space (Shepard, 1980). NMDS involves an initial decision about the number of dimensions, c, to retain. Thereafter the NMDS algorithm returns a set of component scores in c-dimensional space, which are then clustered. To choose the appropriate dimensionality, we performed PCA on the covariance matrix of contrast values for each of 12 mediating regions identified in the MEPM analysis and used a scree (eigenvalue) plot to determine the appropriate number of dimensions to retain. The plot showed that 86% of the covariance in the dataset was contained in the first four dimensions, and additional dimensions explained little additional variance. We then calculated the 12 x 12 matrix of correlations among these regions and converted the correlations to dissimilarity values using the formula $(1-r)/2$, so that 0 indicated zero distance and 2 indicated the maximum possible difference. We applied nonmetric multidimensional scaling (NMDS) analysis to the resulting dissimilarity matrix using stress as the error metric (details in (Wager et al., 2007b)), retaining four dimensions. The result was a 12 region x 4 component matrix of scores, which were subjected to cluster analysis.

We used hierarchical clustering with average linkage (clusterdata.m in Matlab R2007b) to identify networks, as in (Kober et al., in press; Wager et al., 2007b). We used a permutation test to choose the number of clusters and to provide inferences on whether the distances between regions were distributed multimodally (as opposed to a single-mode, single-cluster distribution expected if there were no systematic groups of interconnected regions). For each possible solution between 2 and 5 clusters (networks), we first computed a measure of clustering quality, as defined in (Struyf et al., 1996):

$$q = \sum_k \sum_i \frac{d_{i_o} - d_{i_{nn}}}{\max(d_{i_o}, d_{i_{nn}})}$$

where d denotes Euclidean distance, and d_{i_o} is the distance from region i to the center of its own class, $d_{i_{nn}}$ is the distance to the nearest neighboring class, and k indexes over clusters. We then permuted the columns of the dimension scores, re-applied the clustering algorithm, and calculated *q* based on the permuted data. The permutation procedure disrupts clusters of nearby regions by exchanging their locations in each dimension with those of other regions, while conditionalizing on the marginal distribution of regions in each dimension. This process was repeated 10,000 times to develop a null-hypothesis distribution of *q*. Estimating the distribution of q for each candidate number of clusters k allowed us to assess Z-scores observed-data clustering solution, defined as:

$$Z_k = \frac{q_{obs} - \overline{q}_{null}}{\sqrt{\frac{\sum_1^I (q_{null} - \overline{q}_{null})^2}{I}}}$$

where q_{obs} is the quality for the observed solution, q_{null} is the quality for the permuted-data solution for one iteration, and I is the number of iterations (10,000). Fig. S5B shows Z_k on the y-axis plotted against candidate choices for k on the x-axis. The highest Z-value was found for 2 clusters, which we used as our estimate of k. The permuted-data distribution of q is shown for the 2-cluster solution in Fig. S5C, and q for the observed-data solution is shown by the vertical black line. The significance of the results ($p < .0001$, $Z = 6.24$) indicates that there were at least two separable networks of interconnected mediators. However, we did not test the significance of the difference between 2 and other numbers of clusters; other choices of k (e.g., 3 networks) may be reasonable candidates as well, though we note that the 4 and 5 cluster solutions provided poor fits.

Identification of additional mediated regions in the frontal cortex (Analysis step 5).

Though the choice of right VLPFC as a predictor was motivated by theoretical considerations, it is not likely to be unique in its relationship with subcortical affective pathways, given that activity in multiple frontal regions predicted reappraisal success. In a final analysis, we used the MEPM approach to localize frontal regions whose relationship with reappraisal success was mediated by subcortical activity. Two separate analyses were conducted with reappraisal

success as the outcome. In this analysis, both left NAC and amygdala were specified as mediators, and reappraisal success was specified as the outcome. Average activity values in each region from the previous analysis were used for the mediators. We performed a voxel-wise search for frontal regions whose [ReappNeg – LookNeg] contrast values were mediated (significant *a*b* effect) by each subcortical region.

To constrain the search to frontal regions, we created a mask of frontal regions labeled as "frontal" in the International Consortium for Brain Imaging (ICBM) single-subject atlas ((Mazziotta et al., 2001); http://www.loni.ucla.edu/ICBM/; an average of 27 T1-weighted images of a single subject registered to MNI space and manually labeled) and included the gyrus rectus and cingulate gyrus as labeled for more complete coverage. We smoothed the resulting mask with a 3 mm kernel thresholded at a value of 0, and searched for regions with a significant mediation (*a*b*) effect within this mask. For this exploratory analysis, we report results significant at $p < .001$ with a 3-contiguous-voxel extent or $p < .005$ with a 10-voxel extent.

References

Alexander, M.P., Stuss, D.T., Picton, T., Shallice, T., and Gillingham, S. (2007). Regional frontal injuries cause distinct impairments in cognitive control. Neurology *68*, 1515-1523.

Anderson, A.K., and Phelps, E.A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. Nature *411*, 305-309.

Anderson, A.K., and Phelps, E.A. (2002). Is the human amygdala critical for the subjective experience of emotion? Evidence of intact dispositional affect in patients with amygdala lesions. Journal of Cognitive Neuroscience *14*, 709-720.

Barber, A.D., and Carter, C.S. (2005). Cognitive control involved in overcoming prepotent response tendencies and switching between tasks. Cereb Cortex *15*, 899-912.

Baron, R.M., and Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol *51*, 1173-1182.

Barret, L.F., Mesquita, B., Ochsner, K.N., and Gross, J.J. (2007). The experience of emotion. Annual Review of Psychology *58*, 387-403.

Barrett, L.F., and Wager, T.D. (2006). The Structure of Emotion: Evidence From Neuroimaging Studies. Current Directions in Psychological Science *15*, 79-83.

Berridge, K.C., and Kringelbach, M.L. (2008). Affective neuroscience of pleasure: reward in humans and animals. Psychopharmacology (Berl).

Brosschot, J.F., Gerin, W., and Thayer, J.F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. Journal of Psychosomatic Research *60*, 113-124.

Cavanna, A.E., and Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. Brain *129*, 564-583.

Craig, A.D. (2003). Interoception: the sense of the physiological condition of the body. Curr Opin Neurobiol *13*, 500-505.

Cunnington, R., Windischberger, C., and Moser, E. (2005). Premovement activity of the pre-supplementary motor area and the readiness for action: studies of time-resolved event-related functional MRI. Hum Mov Sci *24*, 644-656.

Davachi, L. (2006). Item, context and relational episodic encoding in humans. Curr Opin Neurobiol *16*, 693-700.

Drevets, W.C. (2001). Neuroimaging and neuropathological studies of depression: implications for the cognitive-emotional features of mood disorders. Curr Opin Neurobiol *11*, 240-249.

Efron, B., and Tibshirani, R. (1993). An Introduction to the Bootstrap (Chapman & Hall/CRC).

Etkin, A., and Wager, T.D. (2007). Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. Am J Psychiatry *164*, 1476-1488.

Faymonville, M.E., Laureys, S., Degueldre, C., DelFiore, G., Luxen, A., Franck, G., Lamy, M., and Maquet, P. (2000). Neural mechanisms of antinociceptive effects of hypnosis. Anesthesiology *92*, 1257-1267.

Goldberg, L., Elliot, D., Clarke, G.N., MacKinnon, D.P., Moe, E., Zoref, L., Green, C., Wolf, S.L., Greffrath, E., Miller, D.J., and Lapin, A. (1996). Effects of a multidimensional anabolic steroid prevention intervention. The Adolescents Training and Learning to Avoid Steroids (ATLAS) Program. Jama *276*, 1555-1562.

Greenberg, D.L., Rice, H.J., Cooper, J.J., Cabeza, R., Rubin, D.C., and Labar, K.S. (2005). Co-activation of the amygdala, hippocampus and inferior frontal gyrus during autobiographical memory retrieval. Neuropsychologia *43*, 659-674.

Gross, J.J., and Munoz, R.F. (1995). Emotion Regulation and Mental Health. Clinical Psychology: Science and Practice *2*, 151-164.

Gusnard, D.A., and Raichle, M.E. (2001). Searching for a baseline: functional imaging and the resting human brain. Nat Rev Neurosci *2*, 685-694.

Haines, D.E. (2000). Neuroanatomy: An Atlas of Structures, Sections, and Systems (Philadelphia: Lippincott Williams & Wilkins).

Isoda, M., and Hikosaka, O. (2007). Switching from automatic to controlled action by monkey medial frontal cortex. Nat Neurosci *10*, 240-248.

Johansen-Berg, H., Gutman, D.A., Behrens, T.E.J., Matthews, P.M., Rushworth, M.F.S., Katz, E., Lozano, A.M., and Mayberg, H.S. (2007). Anatomical Connectivity of the Subgenual Cingulate Region Targeted with Deep Brain Stimulation for Treatment-Resistant Depression. Cerebral Cortex.

Johnstone, T., van Reekum, C.M., Urry, H.L., Kalin, N.H., and Davidson, R.J. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. J Neurosci *27*, 8877-8884.

Kim, S.H., and Hamann, S. (2007). Neural correlates of positive and negative emotion regulation. J Cogn Neurosci *19*, 776-798.

Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Duncan, S., Mize, J., and Wager, T.D. (in press). Networks of Emotion: Functional parcelation and pathway analysis from meta-analytic data.

Lau, H.C., Rogers, R.D., Haggard, P., and Passingham, R.E. (2004). Attention to intention. Science *303*, 1208-1210.

Lieberman, M.D., Jarcho, J.M., Berman, S., Naliboff, B.D., Suyenobu, B.Y., Mandelkern, M., and Mayer, E.A. (2004). The neural correlates of placebo effects: a disruption account. Neuroimage *22*, 447-455.

Lundstrom, B.N., Ingvar, M., and Petersson, K.M. (2005). The role of precuneus and left inferior frontal cortex during source memory episodic retrieval. Neuroimage *27*, 824-834.

MacKinnon, D.P., Krull, J.L., and Lockwood, C.M. (2000). Equivalence of the mediation, confounding and suppression effect. Prev Sci *1*, 173-181.

Maddock, R.J. (1999). The retrosplenial cortex and emotion: new insights from functional neuroimaging of the human brain. Trends Neurosci *22*, 310-316.

Maguire, E.A., and Frith, C.D. (2003). Lateral asymmetry in the hippocampal response to the remoteness of autobiographical memories. J Neurosci *23*, 5302-5307.

Mai, J.K., Assheuer, J., and Paxinos, G. (2004). Atlas of the human brain, 2nd edn (San Diego, Calif.: Elsevier Academic Press).

Martin, J.H. (1996). Neuroanatomy: Text and Atlas, 2nd edn (Stamford, CT: Appleton & Lange).

Mayberg, H.S., Liotti, M., Brannan, S.K., McGinnis, S., Mahurin, R.K., Jerabek, P.A., Silva, J.A., Tekell, J.L., Martin, C.C., Lancaster, J.L., and Fox, P.T. (1999). Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. Am J Psychiatry *156*, 675-682.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B.*, et al.* (2001). A probabilistic atlas and reference system for the human

brain: International Consortium for Brain Mapping (ICBM). Philosophical Transactions of the Royal Society B: Biological Sciences *356*, 1293-1322.

Moskowitz, J.T. (2003). Positive Affect Predicts Lower Risk of AIDS Mortality. (Am Psychosomatic Soc), pp. 620-626.

Nichols, T., and Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. Stat Methods Med Res *12*, 419-446.

O'Doherty, J. (2003). Can't learn without you: predictive value coding in orbitofrontal cortex requires the basolateral amygdala. Neuron *39*, 731-733.

O'Doherty, J., Critchley, H., Deichmann, R., and Dolan, R.J. (2003). Dissociating valence of outcome from behavioral control in human orbital and ventral prefrontal cortices. J Neurosci *23*, 7931-7939.

Ochsner, K.N. (2004). Current directions in social cognitive neuroscience. Curr Opin Neurobiol *14*, 254-258.

Ochsner, K.N., Bunge, S.A., Gross, J.J., and Gabrieli, J.D. (2002). Rethinking feelings: an FMRI study of the cognitive regulation of emotion. J Cogn Neurosci *14*, 1215-1229.

Ochsner, K.N., and Gross, J.J. (2008). Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. Currents Directions in Psychological Science *17*, 153-158.

Ochsner, K.N., Ray, R.D., Cooper, J.C., Robertson, E.R., Chopra, S., Gabrieli, J.D., and Gross, J.J. (2004). For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. Neuroimage *23*, 483-499.

Ollinger, J.M., Corbetta, M., and Shulman, G.L. (2001). Separating processes within a trial in event-related functional MRI. Neuroimage *13*, 218-229.

Phan, K.L., Wager, T., Taylor, S.F., and Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. Neuroimage *16*, 331-348.

Phan, K.L., Wager, T.D., Taylor, S.F., and Liberzon, I. (2004). Functional neuroimaging studies of human emotions. CNS Spectr *9*, 258-266.

Phelps, E.A. (2006). Emotion and cognition: insights from studies of the human amygdala. Annu Rev Psychol *57*, 27-53.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci *10*, 59-63.

Price, D.D., Craggs, J., Verne, G.N., Perlstein, W.M., and Robinson, M.E. (2007). Placebo analgesia is accompanied by large reductions in pain-related brain activity in irritable bowel syndrome patients. Pain *127*, 63-72.

Rainville, P., Duncan, G.H., Price, D.D., Carrier, B., and Bushnell, M.C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science *277*, 968-971.

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2003). The neural basis of economic decision-making in the Ultimatum Game. Science *300*, 1755-1758.

Scheier, M.F., and Carver, C.S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. Cognitive Therapy and Research *16*, 201-228.

Seymour, B., O'Doherty, J.P., Koltzenburg, M., Wiech, K., Frackowiak, R., Friston, K., and Dolan, R. (2005). Opponent appetitive-aversive neural processes underlie predictive learning of pain relief. Nat Neurosci *8*, 1234-1240.

Shepard, R.N. (1980). Multidimensional Scaling, Tree-Fitting, and Clustering. Science *210*, 390-398.

Shrout, P.E., and Bolger, N. (2002). Mediation in experimental and nonexperimental studies: new procedures and recommendations. Psychol Methods *7*, 422-445.

Stern, E.R., Wager, T.D., Egner, T., Hirsch, J., and Mangels, J.A. (2007). Preparatory neural activity predicts performance on a conflict task. Brain Res *1176*, 92-102.

Struyf, A., Hubert, M., and Rousseeuw, P. (1996). Clustering in an Object-Oriented Environment. Journal of Statistical Software *1*, 1–30.

Sumner, P., Nachev, P., Morris, P., Peters, A.M., Jackson, S.R., Kennard, C., and Husain, M. (2007). Human medial frontal cortex mediates unconscious inhibition of voluntary action. Neuron *54*, 697-711.

Tugade, M.M., Fredrickson, B.L., and Feldman Barrett, L. (2004). Psychological Resilience and Positive Emotional Granularity: Examining the Benefits of Positive Emotions on Coping and Health. Journal of Personality *72*, 1161-1190.

Urry, H.L., van Reekum, C.M., Johnstone, T., Kalin, N.H., Thurow, M.E., Schaefer, H.S., Jackson, C.A., Frye, C.J., Greischar, L.L., Alexander, A.L., and Davidson, R.J. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. J Neurosci *26*, 4415-4425.

van Reekum, C.M., Johnstone, T., Urry, H.L., Thurow, M.E., Schaefer, H.S., Alexander, A.L., and Davidson, R.J. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. Neuroimage *36*, 1041-1055.

van Snellenberg, J.X., and Wager, T.D. (in press). Cognitive and Motivational Functions of the Human Prefrontal Cortex. In A tribute to Alexander Luria, E. Goldberg, and D. Bougakov, eds.

Wager, T.D., Barrett, L.F., Bliss-Moreau, E., Lindquist, K., Duncan, S., Kober, H., Joseph, J., Davidson, M., and Mize, J. (in press). The Neuroimaging of Emotion. In Handbook of Emotion, M. Lewis, ed.

Wager, T.D., Lindquist, M., and Kaplan, L. (2007a). Meta-analysis of functional neuroimaging data: Current and future directions. Social, Cognitive, and Affective Neuroscience *2*, 150-158.

Wager, T.D., Reading, S., and Jonides, J. (2004a). Neuroimaging studies of shifting attention: a meta-analysis. Neuroimage *22*, 1679-1693.

Wager, T.D., Rilling, J.K., Smith, E.E., Sokolik, A., Casey, K.L., Davidson, R.J., Kosslyn, S.M., Rose, R.M., and Cohen, J.D. (2004b). Placebo-induced changes in FMRI in the anticipation and experience of pain. Science *303*, 1162-1167.

Wager, T.D., Scott, D.J., and Zubieta, J.K. (2007b). Placebo effects on human mu-opioid activity during pain. Proceedings of the National Academy of Sciences *104*, 11056-11061.

Wager, T.D., and Smith, E.E. (2003). Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci *3*, 255-274.

Whalen, P.J. (1998). Fear, vigilance, and ambiguity: Initial neuroimaging studies of the human amygdala. Current Directions in Psychological Science *7*, 177-188.