Supplemental Material:

Anterior Insula Activity Predicts the Influence of Positively-Framed Messages on Decision Making

Adam Krawitz, Rena Fukunaga, and Joshua W. Brown

Supplemental Text

The following sections contain additional methods and results that complement, support, and expand on the main text.

Detecting Spike Artifacts

Functional volumes were checked for transient spike artifacts in individual slices using a custom algorithm implemented in MATLAB R2007a 7.4.0. Participants whose datasets contained unacceptable spike artifacts were excluded from further analysis. The algorithm analyzed a subset of voxels that lay outside of the brain in each slice of each volume. It computed the ratio of the median signal magnitude for the subset in each slice to the overall median for all the subsets in all slices of all volumes. A slice whose ratio was greater than or equal to 1.3 was classified as containing a spike. Participants were excluded from further analysis if spikes were detected in more than 20% of the total slices in all volumes in any single block. Spike artifacts were due to a technical problem with the scanner and were unrelated to participants' anatomy or performance.

Within-Subject Mediation Analysis

A linear regression with message effectiveness as the covariate and the heightened-risk effect as the contrast of interest is closely related to a within-subject mediation analysis with message as the predictor, decision quality as the outcome, and risk effect as the mediator. The linear regression is testing whether, across participants, message effectiveness is predictive of the heightened-risk effect. The critical relationship for within-subject mediation is that, across subjects, the difference in the value of the outcome, decision quality, is predicted by the difference in the value of the mediator, the risk effect (Judd, Kenny, & McClelland, 2001). But the difference in decision quality is simply message effectiveness, and the difference in the risk effect is simply the heightened-risk effect. Thus the mediation analysis is testing a linear regression where the heightened-risk effect is predictive of message effectiveness. This is the same as our original linear regression, but with the two variables trading places. Since the significance of the slope of x regressed on y is identical to that of y regressed on x, the two tests will yield the same results.

Averaging Across Messages: DLPFC ROI

Follow-up ROI analysis was performed for the left DLPFC region that showed a positive correlation between the risk effect and decision quality averaging across all three messages. Averaging across all of the blocks, the greater the risk effect the higher the decision quality, $r(25) = 0.83$. The correlation of risk effect and decision quality held when considering only the control message blocks, $r(25) = 0.40$, $p < 0.05$. In addition, the correlation of the heightened-risk effect and message effectiveness for the negatively-framed message was significant, $r(25) = 0.41$, $p < 0.05$. However, the correlation of the heightened-risk effect and message effectiveness with the positively-framed message was not significant, $r(25) = 0.27$, $p = 0.17$.

Main Effects Analyses

The whole-brain voxel-by-voxel analyses reported in the main text are linear regressions of behavioral measures of decision quality or message effectiveness on neural contrasts of the risk effect or the heightened-risk effect. For completeness, we report here whole-brain voxel-by-voxel analyses of the main effects of risk and heightened-risk. It should be noted that the main effects and the correlations of those effects with the behavioral measures are orthogonal. The regression analyses in the main text provide a more compelling link between neural activity in a particular brain region and behavior, because they show that the magnitude of activity for each participant is predictive of that individual's performance, whereas the main effects simply show that, on average across all of the participants, greater activity in a region is linked with higher mean performance.

For each of the contrasts for which a whole-brain voxel-by-voxel linear regression was performed, we report here whole-brain voxel-by-voxel tests for the main effect of that contrast. These analyses were done using *t*-tests and ReML estimation in SPM5. The statistical threshold for significance was $p < 0.05$, FWE. No significantly positive voxels were found for the risk effect with the control message. Likewise, no significantly positive voxels were found for the heightened-risk effect with the positively-framed message or the negatively-framed message. However, for the risk effect averaging across all messages, a number of significantly positive voxels were found (see Supplemental Table 2 and Supplementa1 Figure 1). No voxels were significantly negative for any of these contrasts.

Negative Correlation Analyses

Additional tests for negative slopes were performed for each of the linear regressions reported in the main text at $p < 0.05$, FWE. No voxels had significant negative slopes for any of these linear regressions.

Excluding Alternative Interpretations

We have ruled out a number of alternative explanations for the risk effects and heightened-risk effects. The order of presentation used here mirrors the main text.

1. The "prominent deck B" phenomenon. Lin et al. (2007) claimed that binning together the selections from the two bad decks and the two good decks when analyzing IGT is inappropriate because the net-loss/rare-loss deck (referred to by Lin as "deck B") is in fact preferred to the good decks. We evaluated this possibility in our results by comparing the proportion of selections from each deck. We found that the net-loss/rare-loss deck (mean choice probability $M = 0.16$) is slightly preferred to the net-loss/frequent-loss deck ($M = 0.13$), $t(26) = 3.02$, $p < 0.01$. However, it is preferred to neither the net-gain/frequent-loss deck ($M = 0.34$), $t(26) = -4.65$, $p < 0.0001$, nor the net-gain/rare-loss deck ($M = 0.37$), $t(26) = -6.50$, $p < 0.000001$ (see Supplemental Figure 3). Thus the "prominent deck B" phenomenon does not occur in this study and we are justified in grouping the two good decks and the two bad decks together to increase statistical power and simplify the analysis.

2. Time within block. Since choice behavior typically changed over the course of each block of trials as the participant learned about the decks, a potential confound exists between the time within a block and the deck selections being made. More of the bad selections were made earlier within each block and more of the good selections were made later within each block. As a result, a gradual decrease in brain activation during decision making over the course of each block, due, perhaps, to decreasing attention, could be misinterpreted as a difference between bad and good decisions (Dunn, Dalgleish, & Lawrence, 2006).

To test for this possibility, we ran a second GLM for each participant. In addition to all of the regressors in the original GLM, we added an additional decision-time event for each block. These events were aligned to the response and inserted for each trial where a response was given, providing three new event-related regressors: *ControlAll*, *PositiveAll*, and *NegativeAll*. Each of these events had a parametric modulator (*ControlSlope*, *PositiveSlope*, and *NegativeSlope)* that was mean-centered and increased linearly as a function of trial number within the block. These event-related slope regressors removed any overall linear trend in decision-time activation across the course of each block. In this analysis, the differences in activation between bad and good decisions are relative to any overall linear trend over the course of the block, and cannot be explained by the time within block.

For each of the regions identified as having a significant correlation in the original analysis, we analyzed the slopes of the event-related decision-time linear regressors. All of the slopes either trended negative or were significantly negative (see Supplemental Figure 4A). This indicates that there was a decrease in decision-time activation across each block.

For each of the regions identified as having a significant correlation in the original analysis, we retested the significance of the correlation with the results of the new analysis. In right IFO, the regression of decision quality on the risk effect for the control message was still significant, $r(25) = 0.79$, $p < 10^{-5}$. In right AI, the regression of message effectiveness on the heightened-risk effect for positively-framed messaged was still significant, $r(25) = 0.74$, $p < 10^{-4}$. In left ACC/MPFC, the regression of mean decision quality on the mean risk effect averaging across all messages was still significant, $r(25) = 0.80$, $p < 10^{-6}$. And in left DLPFC, the regression of mean decision quality on the mean risk effect averaging across all messages was still significant, $r(25) = 0.76$, $p < 10^{-5}$. In addition, the other correlations reported in the main text followed the same overall pattern as they had in the original analysis (see Supplemental Figure 4B and C).

The fact that the original pattern of results remains intact with the inclusion of the event-related decision-time linear regressors indicates that time-within-block cannot explain the findings reported in the main text. While activations during decision making did decline over the course of each block, these declines do not explain the differences in activation between bad and good decisions.

3. Response times. We defined response time as the time it took participants to select a deck on each trial, from the presentation of the decks and message until a response was selected by button press. If the difference in response time for bad versus good decisions varied as a function of message, then heightened-risk effects could be due to changes in time on task instead of changes in risk appraisal. Participants were significantly slower for selections from bad decks *(M* = 856 ms) than good decks *(M* = 791 ms), $t(26) = 3.12$, $p < 0.005$ (see Supplemental Figure 5A). However, a four-way ANOVA with decision type, message, block position within session, and epoch within block as factors showed that decision type did not interact with message, $F(2, 20) = 0.51$, $p = 0.61$ (see Supplemental Figure 5 and Supplemental Table 4). Thus, differences in the risk effect between messages cannot be accounted for by response time.

Since our main results look at correlations across participants between behavioral and neural measures, we also looked at whether response time correlated with these measures across participants. None of the behavioral measures were significantly correlated with response times (see Supplemental Figure 6). Decision quality was not correlated with the difference in response times for bad and good decisions, either across messages, $r(25) = -0.02$, $p = 0.92$, nor for the control message alone, $r(25) = 0.26$, $p = 0.19$. And message effectiveness was not correlated with the change in response time differences for bad versus good decisions for the positively-framed message compared to the control message, $r(25) = 0.04$, $p = 0.85$, nor for the negatively-framed message compared to the control message, $r(25) = -0.02$, $p = 0.94$.

None of the neural measures used to identify the ROIs were significantly correlated with response times (see Supplemental Figure 7). The risk effect in right IFO was not correlated with the difference in response times for bad and good decisions in the control message blocks, $r(25) = 0.26$, $p = 0.19$. The heightened-risk effect in right AI was not correlated with the change in response time differences for bad versus good decisions in the poisitively-framed message blocks compared to the control message blocks, $r(25) = -0.03$, $p = 0.90$. And the mean risk effect was not correlated with the difference in response times for bad and good decisions across all message blocks in the ACC/DMPFC ROI, $r(25) = -0.16$, $p = 0.43$, nor in the DLPFC, $r(25) = -0.02$, $p = 0.94$. These findings exclude the possibility that differences in neural activation found to correlate with decision quality or message effectiveness can be explained by differences in response times.

4. Order of message presentation. Since block order was counterbalanced across participants, individual differences in decision quality and the risk effect could have been driven by the order of the messages. More critically, individual differences in message effectiveness and the heightened-risk effect could have been driven by the relative ordering of the blocks and not by the relative effectiveness of the messages. In order to exclude block order as a confounding factor, we performed statistical model comparison with analysis of covariance (ANCOVA) using the mean parameter estimates within the ROIs described in the main text.

For the control message risk effect in right IFO, the position of the control-message block was entered as a categorical variable, and then decision quality was entered as a continuous variable. Decision quality explained significant additional variance, $F(1, 23) = 31.79$, $p < 10^{-4}$. For the positively-framed message heightened-risk effect in right AI, the position of the positively-framed-message block relative to the control-message block was entered as a categorical variable, and then message effectiveness was entered as a continuous variable. Message effectiveness explained significant additional variance, $F(1, 20) = 13.74$, $p < 0.005$. For the risk effect across all message types in left ACC/DMPFC and left DLPFC, the overall block sequence was entered as a categorical variable, and then the mean decision quality was entered as a continuous variable. Decision quality explained significant additional variance in both ACC/DMPFC, $F(1, 20) = 62.01$, $p < 10^{-5}$, and DLPFC, $F(1, 20) = 36.65$, $p < 10^{-4}$. In each case, the significant relationship between the behavioral measure and the neural measure was not explained by the order of message presentation.

5. Novelty. Considered alone, a correlation between decision quality and the risk effect might be explained as an effect of novelty, where the more infrequently that bad decks are selected, the greater the activation when they are selected. Indeed, such effects have been observed previously in the ACC (Braver, Barch, Gray, Molfese, & Snyder, 2001). However, this explanation would predict the same brain region to exhibit the novelty effect for rare bad decisions no matter their cause. But we found brain regions associated more strongly with particular messages, demonstrating that the effects are not due simply to the frequency of choices, but rather are due to the information driving those choices.

Supplemental References

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of frequency, inhibition and errors. *Cerebral Cortex, 11*(9), 825-836. doi:10.1093/cercor/11.9.825

Dunn, B. D., Dalgleish, T., & Lawrence, A. D. (2006). The somatic marker hypothesis: A critical evaluation. *Neuroscience and Biobehavioral Reviews, 30*(2), 239-271. doi:10.1016/j.neubiorev.2005.07.001

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods, 6*(2), 115-134. doi:10.1037/1082-989X.6.2.115

Lin, C.-H., Chiu, Y.-C., Lee, P.-L., & Hsieh, J.-C. (2007). Is deck B a disadvantageous deck in the Iowa Gambling Task? *Behavioral and Brain Functions, 3*(1), 16. doi:10.1186/1744-9081-3-16

Supplemental Table 1

Three-way ANOVA for Decision Quality

Term	DF	SS	DFE	SSE	*F*	*p*
Message (M)	2	5.65	46	10.16	12.79*	3.8×10^{-5}
Block position (B)	2	7.74	46	10.16	17.53*	2.2×10^{-6}
Epoch (E)	4	8.15	88	8.47	21.18*	2.9×10^{-12}
M x B	4	1.41	46	10.16	1.59	0.19
M x E	8	0.28	184	12.03	0.54	0.83
B x E	8	0.91	184	12.03	1.73	0.094
M x B x E	16	1.32	184	12.03	1.26	0.22

Note. 3x3x5 within-subject ANOVA for decision quality (calculated for each epoch of 20 trials within each block for each participant) with message presented (control, positive or negative), block position within session (1st, 2nd, or 3rd), and epoch within block (1, 2, 3, 4 or 5) as factors. DF = degrees of freedom; SS = sum of squares; DFE = degrees of freedom in error term; SSE = sum of squares for error term.

*$p < 0.05$.

Supplemental Table 2

Main Effect of Risk Appraisal (Bad – Good) Across All Messages

Brain region	BA	MNI coordinates x	y	z	*Z*-score	Cluster (voxels)
Putamen		-12	10	-10	5.77	54
Caudate		12	8	-10	5.58	32
Inferior frontal gyrus	47	-30	26	-8	5.31	18
Medial globus pallidus		-12	-4	0	5.26	6
Lateral globus pallidus		12	6	4	5.05	1
Medial frontal gyrus	9	12	40	26	5.03	5
Inferior frontal gyrus	13	38	22	10	5.02	3
Putamen		14	2	10	5.01	2
Cingulate gyrus	32	4	28	36	4.99	5
Insula	13	34	22	6	4.99	1
Cingulate gyrus	32	14	32	32	4.97	1
Caudate head		10	4	2	4.95	2

Note. Regions with a significant risk effect averaging across all messages, R_{all}, ordered by peak *Z*-score. Results are from a whole-brain voxel-by-voxel analysis thresholded at $p < 0.05$, FWE. BA, MNI coordinates, and *Z*-score are for the voxel of peak activation within each cluster. BA labels are only available for cortical regions. BA = Brodmann area; MNI = Montreal Neurological Institute.

Supplemental Table 3

Mediators of the Right AI Relationship to Decision Quality with the Positively-Framed Message

Direction of mediation / Region	BA	MNI coordinates x	y	z	Path a *Z*-score	Path b *Z*-score	Path ab *Z*-score	Cluster (voxels)
Positive								
Cingulate gyrus	32	6	14	40	3.43	3.55	3.58	22
Posterior cingulate	31	22	-66	14	3.58	3.58	3.57	37
Superior parietal lobule	7	34	-64	50	3.40	3.57	3.39	70
Inferior frontal gyrus	13	28	20	-10	3.54	3.30	3.23	3
Middle frontal gyrus	6	28	-10	46	3.57	3.14	3.12	3
Cingulate gyrus	32	2	4	44	3.58	2.92	2.96	3
Negative								
Insula	13	36	32	6	3.56	-3.52	-3.56	18

Note. For the positively-framed message, regions where the risk effect mediates the relationship between the risk effect in the right anterior insula region-of-interest and decision quality. Results are for a whole-brain voxel-by-voxel analysis, $p < 0.005$ for paths *a*, *b*, and *ab*, using bootstrapping with 1000 samples, 3 contiguous voxels. BA, MNI coordinates, and *Z*-scores are for the voxel of peak activation within each cluster. Results ordered by peak *Z*-score for path *ab*. BA labels are only available for cortical regions. BA = Brodmann area; MNI = Montreal Neurological Institute.

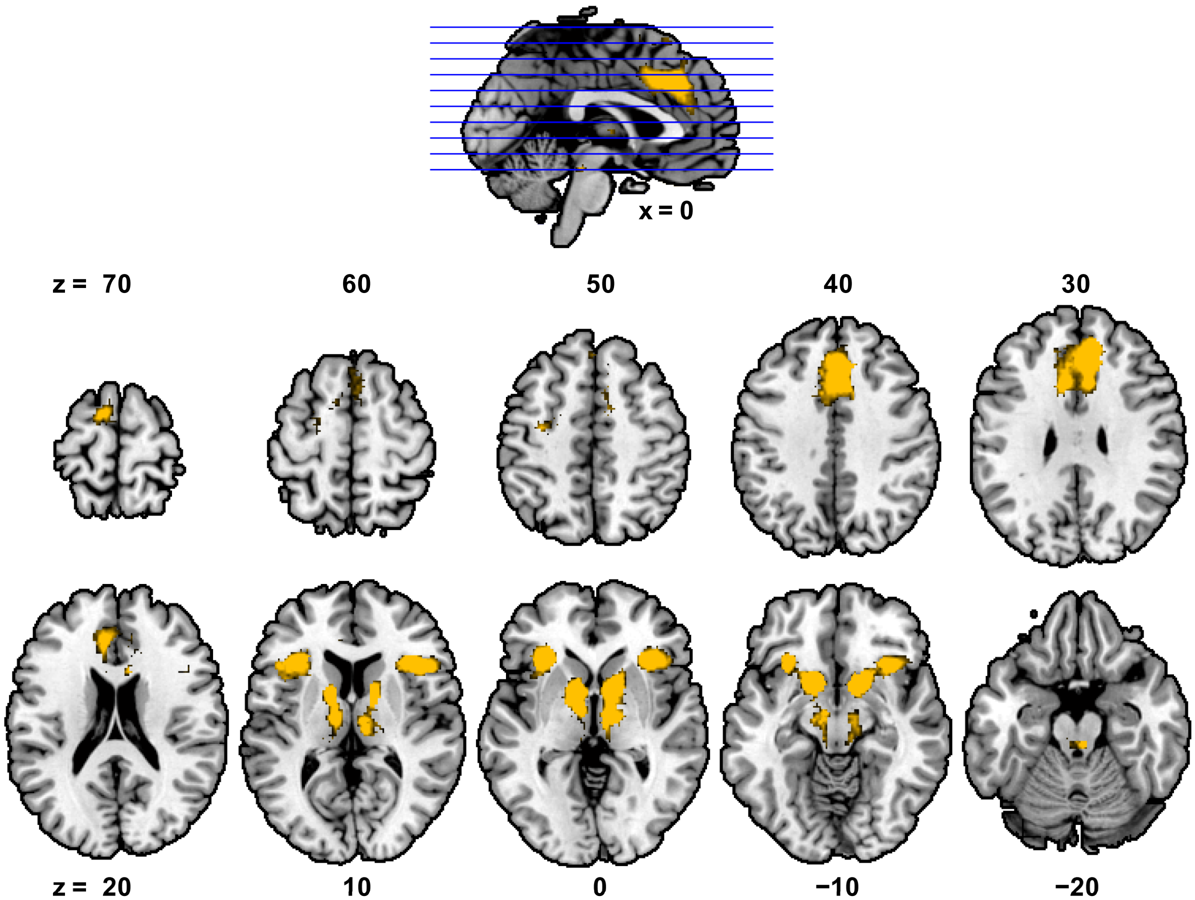
Supplemental Table 4

Four-way ANOVA for Response Time

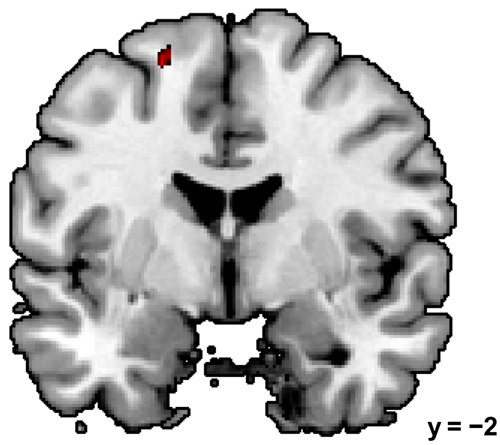
Term	DF	SS	DFE	SSE	*F*	*p*
Decision (D)	1	498398	5	40201	61.99*	0.00053
Message (M)	2	662390	21	3151283	2.21	0.13
Block position (B)	2	1528634	21	3151283	5.09*	0.016
Epoch (E)	4	2256884	56	3371018	9.37*	$7.2x10^{-6}$
D x M	2	40562	20	800715	0.51	0.61
D x B	2	65795	20	800715	0.82	0.45
D x E	4	41657	50	1843329	0.28	0.89
D x M x B	4	36453	20	800715	0.23	0.92
D x M x E	8	640470	109	4404284	1.98	0.056
D x B x E	8	188251	109	4404284	1.73	0.79
D x M x B x E	16	511834	109	4404284	0.79	0.69

Note. 2x3x3x5 within-subject ANOVA for response time with decision type (good or bad), message presented (control, positive or negative), block position within session (1st, 2nd, or 3rd), and epoch within block (1, 2, 3, 4 or 5) as factors. DF = degrees of freedom; SS = sum of squares; DFE = degrees of freedom in error term; SSE = sum of squares for error term.

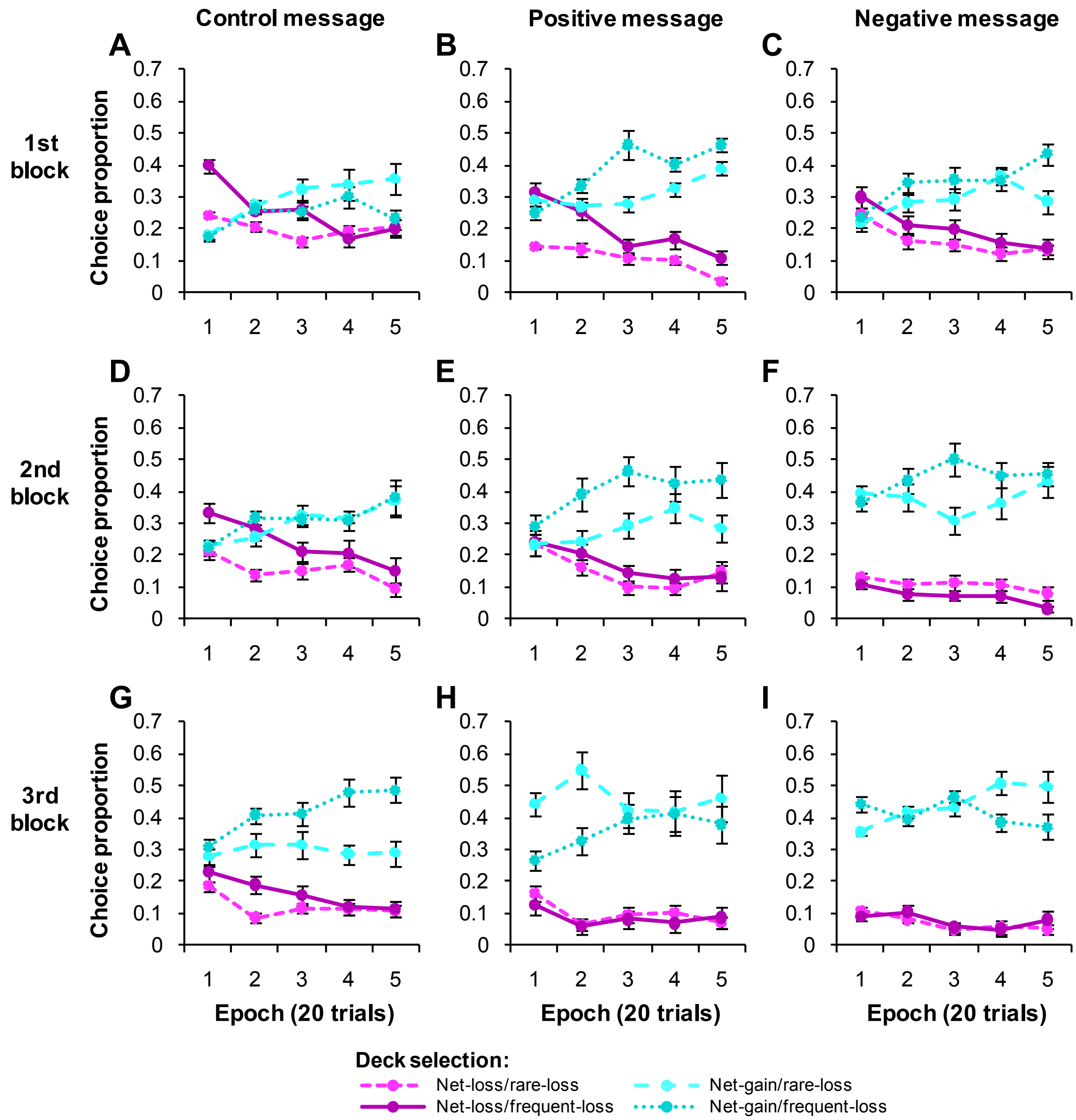
*$p < 0.05$.



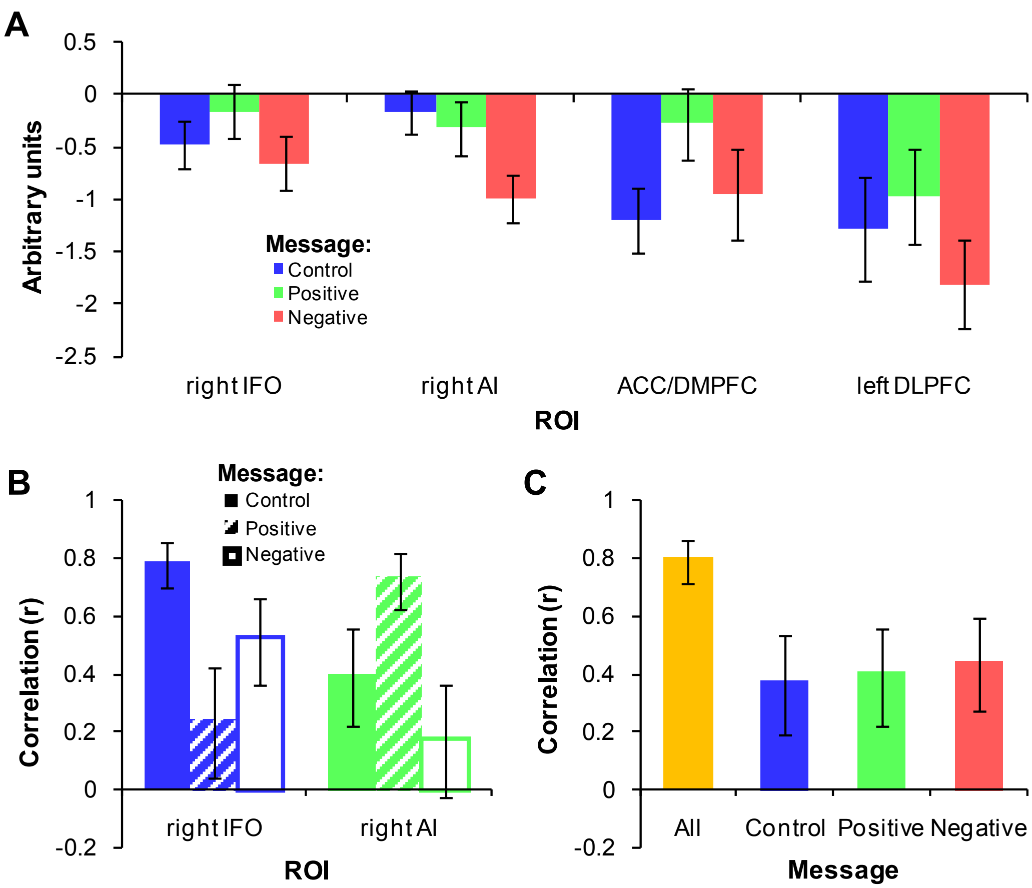
Supplemental Figure 1. Main effect of risk appraisal across all messages. Shown in orange are regions where significant activations were found for the risk effect. Refer to Supplemental Table 2 for a list of regions with significant activations at $p < 0.05$, FWE. Blue lines in the top sagittal section of the human brain (MNI x = 0), indicate locations of the ten transverse sections shown below ranging from MNI z = 70 to MNI z = -20. The display uses a threshold of $p < 0.0005$, uncorrected, with maximum color brightness indicating $p < 0.000005$, uncorrected.



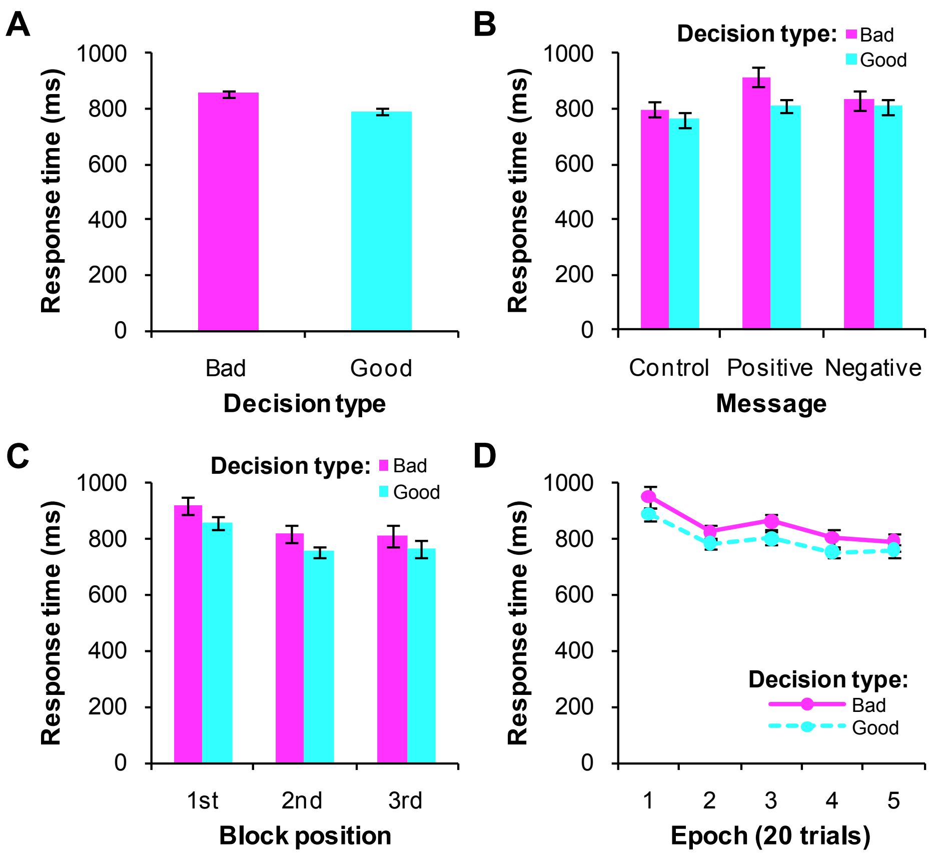
Supplemental Figure 2. Shown in red is a region in left middle frontal gyrus (BA 6, peak voxel: MNI −20, −2, 62) with a correlation between message effectiveness and the heightened-risk effect for the negatively-framed message. This region was significant at $p < .0001$, uncorrected, but failed to pass a more stringent FWE or FDR correction, and therefore should be considered as only suggestive. Coronal section of the human brain (MNI y = −2).



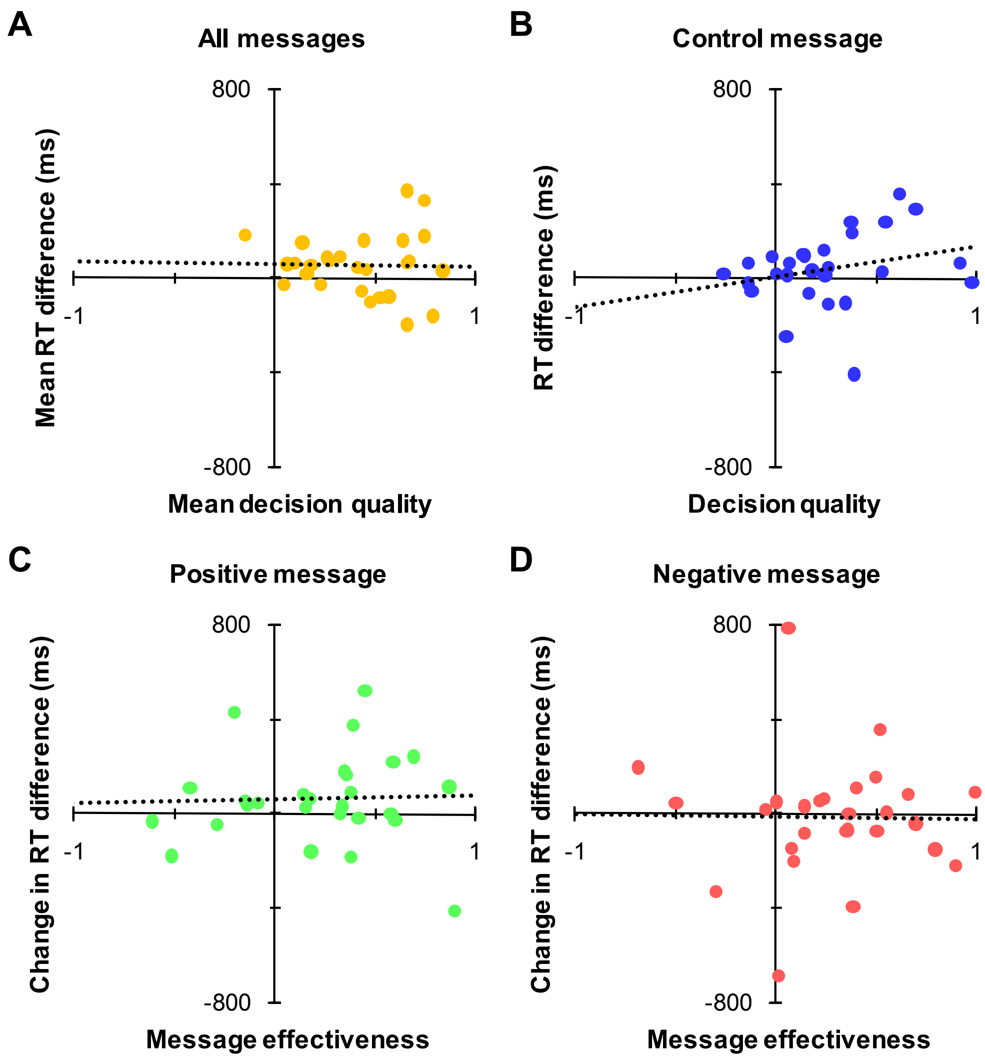
Supplemental Figure 3. Deck selections as a function of message, block position, and epoch. The choice proportion is the proportion of trials in which a particular deck was selected, excluding those trials where no selection was made. The bad decks are in magenta and the good decks in cyan. Each participant contributes to exactly one graph in each row and each column. Error bars indicate within-subject SEM.



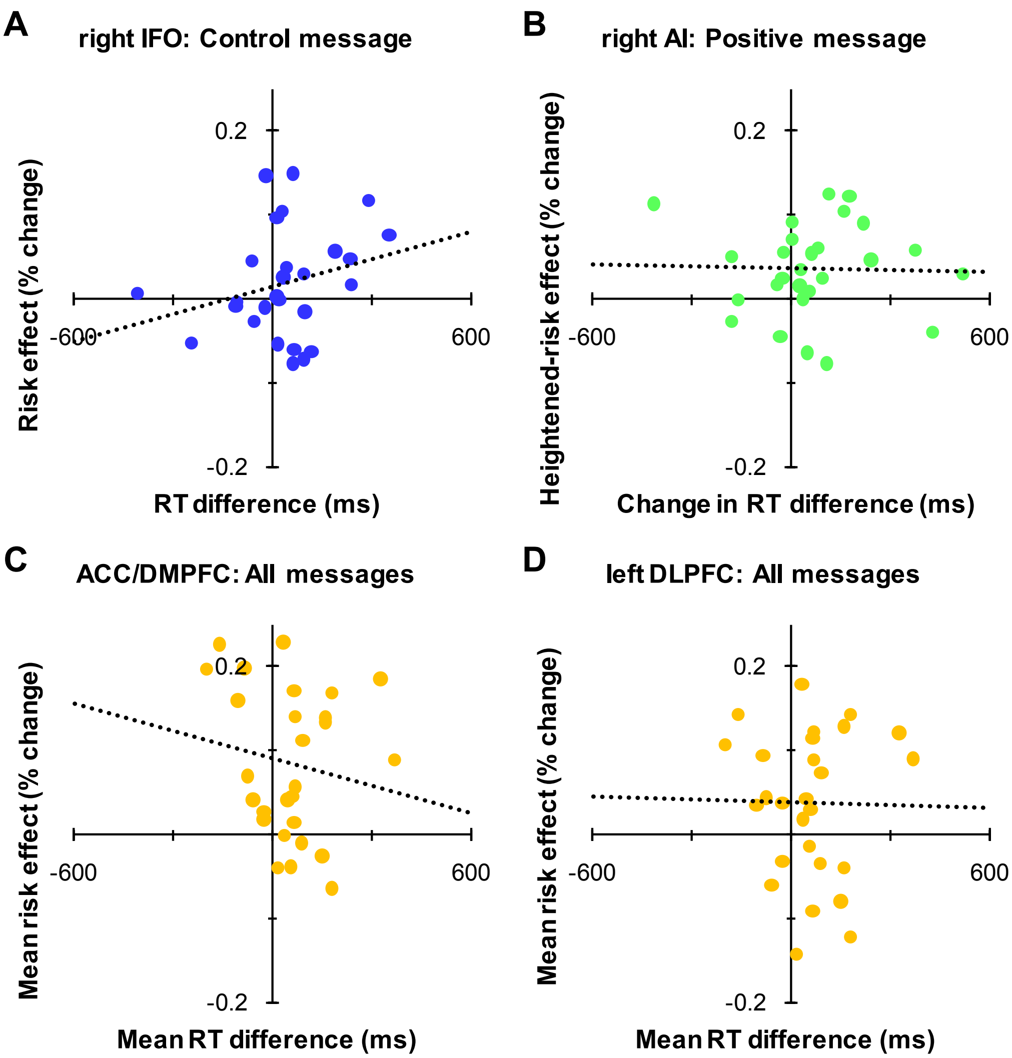
Supplemental Figure 4. Excluding time-within-block as a potential confound using new GLMs with event-related decision-time linear regressors. Error bars indicate SEM. (A) Slopes for event-related decision-time linear regressors for each message in each brain region where significant correlations are reported in the main text. All slopes either trend negative or are significantly negative, indicating a decrease in decision-time activation throughout each block. (B) Reanalysis of the data shown in Figure 3B duplicating the results with the new GLMs. (C) Reanalysis of the data shown in Figure 5B duplicating the results with the new GLMs.



Supplemental Figure 5. Response time (RT). There were significant main effects on RT of decision type, block position, and epoch, but not message. There were no significant interactions between decision type and any of the other factors. Error bars represent within-subject SEM. (A) RT as a function of decision type. (B) RT as a function of decision type and message. (C) RT as a function of decision type and block position. (D) RT as a function of decision type and epoch.



Supplemental Figure 6. Decision quality (DQ) versus differences in response time between bad and good decisions (ΔRT). Each point represents a participant. The dotted lines show the best-fit linear regressions. (A) The non-significant correlation of DQ and mean ΔRT across all messages. (B) The non-significant correlation of DQ and ΔRT with the control message. (C) The non-significant correlation between positively-framed message effectiveness (ME) and the increase in ΔRT with the positively-framed compared to the control message. (D) The non-significant correlation between negatively-framed ME and the increase in ΔRT with the negatively-framed message compared to the control message.



Supplemental Figure 7. Risk effect (RE) or heightened-risk effect (HRE) in ROIs versus differences in response time between bad and good decisions (ΔRT). Each point represents a participant. The dotted lines show the best-fit linear regressions. (A) The non-significant correlation of control message RE in right IFO and ΔRT with the control message. (C) The non-significant correlation between positively-framed message HRE in right AI and the increase in ΔRT with the positively-framed compared to the control message. (B) The non-significant correlation of mean RE in ACC/DMPFC and mean ΔRT across all messages. (D) The non-significant correlation of mean RE in left DLPFC and mean ΔRT across all messages.