Supporting Information

SI Methods

Participants. Thirty right-handed subjects (20–34 years old; 15 males) participated in the present study. Two additional subjects were recruited but were excluded from data analysis due to problems during scanning. Of these, one participant with cold symptoms reported some discomfort and needing to hold her breath frequently to avoid coughing; she also reported difficulty focusing on the task. A second participant confused the scene category (house/building) and/or confused certain kind of stimuli (e.g., responding "building" for two-story house stimuli). All participants were in good health with no history of neurological or psychiatric disorders. Participants had normal or corrected-to-normal vision.

Behavioral Experiment and Stimuli Presentation. Before the fMRI sessions, subjects performed a behavioral session in a "mock" scanner that lasted ≈1.5 h. As in fMRI sessions (see below), stimuli were presented through an LCD projector at a resolution of 1024 × 768 pixels and a refresh rate of 60 Hz. All stimuli were presented at fixation, subtended 4.5° × 4.5° of visual angle, and were shown in black-and-white on a gray background. The schedule of stimulus presentation and data collection were controlled by Presentation software (Neurobehavioral Systems). Three neutral face stimuli were selected from the Karolinska Directed Emotional Faces (KDEF) (1). For scene stimuli, 84 house images and 84 building images were selected. A grid of thin black lines was superimposed on all images (see Fig. 2 for examples). The T1 face image was displayed 200, 300, 400, 500, or 800 ms before the T2 scene (the T1-T2 lag was randomized). A T1 stimulus (face) was presented in every trial, and a T2 stimulus (house or building) was presented in 78% of the trials. For scene-absent trials, an additional distractor image was used to replace the scene image. The response button mapping for houses and buildings was counterbalanced across participants. Participants were instructed to give priority to the T1 task over the T2 task during the dual-task (i.e., T1 and T2) condition. Each trial began with a display of a green fixation cross for 400 ms, followed by the stimulus stream. After the stream, participants were asked to make a T1-related decision (Andy, Bill, or Chad?) during a 2-s response period and a subsequent T2-related decision (house, building, or no-scene?) during a 2-s response period. Dual-task trials occurred every 8 s. For the dual-task condition, T2 responses were analyzed only when the T1 response was correct to ensure proper attentional engagement on the T1 process (2). Single-task (i.e., T2 only) trials occurred every 6 s. Single- and dual-task conditions were performed in separate blocks with randomized order. A total of 960 dual-task trials were presented over eight blocks and a total of 96 single-task trials were presented over four blocks.

To familiarize participants with the three face identities labeled Andy, Bill, and Chad, hard copies of printed face images with their names were provided. At the beginning of the behavioral session, subjects performed a short block of a face identity task and a short block of a scene categorization task, during which feedback (correct/incorrect) was provided after each response.

fMRI Experiment and Fear Conditioning Procedure. After the behavioral session, participants underwent two fMRI sessions (three sessions were administered for two participants who exhibited a low number of miss trials; see below). The two fMRI sessions were completed within 1 week of each other (typically in subsequent days).

For each fMRI session, participants performed 10 conditioning training trials (including two that involved electrical stimulation) and 40 dual-task training trials during the initial anatomical scan. The main experimental phase was subdivided into two phases, learning and AB. During the affective learning phase, differential fear conditioning was used as a manipulation of the affective significance of scene stimuli. During this phase, participants performed a two-choice scene categorization task (house or building). House or building images were designated as the CS^+ category, and mild electrical stimulation served as the US (the CS^+ category was counterbalanced across participants). During conditioning trials, images of CS^+ scenes were followed by a US according to a 50% partial reinforcement schedule. The US shock was administered to the distal phalange of the third and fourth fingers of the nondominant (left) hand through a stimulator (E13–22; Coulbourn Instruments), which included a grounded RF filter, and MR-compatible leads and electrodes (BIOPAC Systems). At the beginning of the MRI experiment, subjects were explicitly instructed of the contingency rule (i.e., the CS^+ scene category), but were not informed about the probability of US delivery. The intensity of the "highly unpleasant but not painful" electric shock (range: 0.6–4.0 mA) was set for each participant individually while he/she was prepared for MRI scanning. The subjective aversiveness of US shock was monitored between the functional runs via the intercom system, and the intensity of stimulation was readjusted if needed. On each conditioning trial, images of scenes (without the superimposed grid) were displayed for 2 s and followed by an 8-s fixation cross. Thirty-six scene images were displayed at the center of the display in random order, with the constraint that no more than two buildings or two houses were presented successively: 12 CS^+ scenes without shock (CS^+), 12 CS^+ scenes paired with shock (CS^+ with US; thus CS^+ scenes were paired with shock 50% of the time), and 12 CS^- scenes. The US was delivered 1,500 ms after the onset of a CS^+ stimulus and co-terminated with the CS^+, following a delayed conditioning paradigm. All trials involving electrical stimulation were discarded from further analyses.

Across two fMRI sessions, participants finished a total of 24 dual-task runs (36 runs for two participants with three fMRI sessions). Unless otherwise noted, all stimulation parameters were identical to those used during the behavioral session administered in the mock scanner (including projector type). A total of 168 CS^+ dual trials and 168 CS^- dual trials were performed (252 for participants with three fMRI sessions). In addition, 96 no-scene trials (T1-only) were collected for each participant (144 for participants with three fMRI sessions). To minimize the extinction of conditioned responses, three additional CS^+ trials with US were present in each dual-task run (these were discarded from further analyses, too). Subjects were discouraged to guess "house" or "building" when unsure and were encouraged instead to use a third option ("no-scene") in such cases. Accordingly, very few false alarms (i.e., "house" or "building" responses during T1-only trials) or incorrect scene categorization responses (i.e., "house" response in trials containing a building, or "building" response in trials containing a house) were observed, and these trials were excluded from further analyses.

Skin Conductance Responses. During MRI data acquisition, SCRs were simultaneously recorded with the MP-150 system (BIOPAC Systems) and MRI-compatible Ag/AgCl electrodes placed on the distal phalange of the index and middle finger of the nondominant (left) hand. SCR was amplified and sampled at 250 Hz, and the analysis of SCR waveforms was conducted using AcqKnowledge software (BIOPAC Systems). As in our previous studies (3, 4), recorded SCR waveforms were resampled at 100

Hz, detrended, and smoothed with a median filter over 50 samples to filter out MRI-induced noise. On each trial, the SCR was calculated by subtracting a baseline (average signal between 0 and 1 s) from the peak amplitude during the 4–6 s time window following stimulus onset (5).

Functional Localizer. At the end of the first fMRI session, participants performed one functional "localizer" run. During this run, subjects performed a simple one-back working memory task that was administered in a blocked fashion with alternating face and scene blocks. Participants were explicitly informed that no shocks would be administered during these trials. A total of six face blocks and six scene blocks were performed, each of which contained 10 trials. Different colored face and scene were used. Each trial begun with a 500-ms white fixation cross, followed by a 1,500-ms presentation of a face or scene.

MRI Data Acquisition and Basic Analysis. Anatomical and functional scans were acquired using a 3T TRIO scanner (Siemens Medical Systems) with an eight-channel phased-array head coil. Structural images were acquired first with a high-resolution MPRAGE anatomical sequence (TR = 1,900 ms; TE = 4.15 ms; TI = 1,100 ms; 1-mm isotropic voxel; 256-mm field of view). Next, blood oxygenation level-dependent (BOLD) contrast functional images were acquired with gradient-echo echo-planar T2*-weighted imaging. Each functional volume consisted of 34 axial slices (TR = 2,000 ms; TE = 25 ms; FA = 70 °; field of view, 24 cm; 64 × 64 matrix; 3.8 mm thickness; interleaved acquisition order).

Analysis of fMRI data were performed using AFNI tools (6), custom software written by the authors, or otherwise indicated. The first six functional volumes of each run were removed to account for equilibration effects of magnetization. The following processing steps were applied: Slice-time correction, motion correction, normalization to Talairach space (7), Gaussian spatial smoothing (full width at half maximum: 6 mm), and intensity normalization (each voxel's mean was set to 100). Functional images acquired during the second (and third when needed) scanning session were aligned to those collected during the first scanning session by applying a transformation matrix determined by registering the second (and third when needed) anatomical data set to the first session anatomical data set.

ROI Analysis. As our main goal was to assess the relationship between responses in key brain regions and behavior, our analyses were performed based on a set of target ROIs. To determine ROIs, voxel-based analyses were carried out by analyzing each individual's functional data via multiple linear regression.

Three criteria were used to determine ROIs. First, we capitalized on the existence of relatively category-specific responses in visual cortex to faces and scenes, an approach that was successfully applied to the investigation of the AB in the past (8, 9). For each individual, to determine PHG ROIs, scene-containing blocks were contrasted to face-containing blocks based on the separate localizer run ($P < 0.01$, uncorrected) via a paired t-test. Voxels so activated (scenes > faces) that were within an anatomically defined mask of the PHG (as obtained in AFNI) and that were also activated during any condition of the main experimental phase ($P < 0.001$, uncorrected) were averaged together to determine a representative time series for the region. All participants exhibited an ROI in the right hemisphere, but only 17 participants exhibited a left-hemisphere ROI. For control analyses, we also defined ROIs in the FG by applying an analogous procedure but with the opposite object selectivity (faces > scenes). Again, all participants yielded ROIs in the right hemisphere, but only 23 yielded ROIs in the left hemisphere.

Second, we defined ROIs for the amygdala. A two-stage mixed-effects analysis was performed in which regression coefficients for each condition of interest [fixed factor; CS^+ hit, CS^+ miss, CS^- hit, CS^- miss, and correct reject during no-scene trials (cr)] were tested across subjects (random factor) via a repeated-measures ANOVA. Error trials including false alarms and incorrectly categorized T2 responses (see above) were modeled as a separate regressor in the model. Considering the typically lower signal-to-noise ratio in the amygdala, its small size, and the fact that it was an a priori ROI, voxels exhibiting a significant main effect in the one-way repeated-measures ANOVA (conditions: CS^+ hit, CS^+ miss, CS^- hit, CS^- miss, and cr) at a more lenient threshold ($P < 0.05$, uncorrected) were averaged together to determine a representative time series if they fell within a standard anatomical mask of the amygdala (as provided by AFNI). Because at the group level, only the right amygdala exhibited a significant main effect, we did not carry out analyses with the left amygdala.

Third, ROIs for frontoparietal regions were defined in a manner that was similar to the procedure used for the amygdala. Voxels that exhibited a main effect according to a one-way repeated-measures ANOVA (conditions: CS^+ hit, CS^+ miss, CS^- hit, CS^- miss, and cr) were initially considered. In this case, a stricter statistical criterion was adopted ($P < 0.05$, corrected according to the false discovery rate; reference 10). The centers of the ROIs were located at the local-maximum voxel (at the group level) and a 6-mm radius sphere was used to average the time series of voxels to generate a representative time series. Although we were interested a priori in regions of frontoparietal cortex that have been implicated in attentional processes, for completeness, we report other sites that exhibited robust main effects as outlined above (see Table S1). As an index of single-trial response amplitude, the mean response at 4 and 6 s post T2 stimulus onset was used (i.e., 6 and 8 s post-trial onset).

Although reaction times (RTs) are not of primary interest in AB studies, in general, variability in RT and time-on-task effects are known to modulate fMRI responses (11); see also De Martino et al. (9) and Anderson (12) for RT effects in the AB. In the present study, we formally assessed the relationship between RT and fMRI responses via, for example, a voxelwise correlational analysis for miss trials during the CS^- condition (we chose miss trials because they exhibited the longest RT; the CS^- condition was chosen because it was affectively neutral). Across subjects, RTs were correlated with increases in evoked responses in several frontoparietal regions, including the cingulate gyrus, bilateral MFG, and bilateral IPL, in addition to the right PHG in visual cortex (all Ps < 0.05) [see also section "Behavioral Performance of Dual-Task Trials (fMRI Session)" below for further RT effects]. Accordingly, for the ROI analyses further described below, before averaging the time-series in each ROI, the variance explained by RT was subtracted out of the original time series of each voxel (slow-varying drifts in MR signal were likewise removed). A convenient way to implement this is available via the 3dSynthesize program in AFNI. Critically, note that a very similar pattern of results was obtained when no RT correction was applied (the main difference concerned frontoparietal regions whose miss evoked responses were largest and were associated with the longest RTs).

Path Analysis

Hits: CS^+ vs. CS^-. To probe network interactions, standard path analysis was used. The first path analysis investigated the relationships between mean fMRI and behavioral responses. Because we were interested in the effect of aversive conditioning on brain responses and behavior, differential scores (CS^+ vs. CS^-) were used. In the case of amygdala and PHG responses, differential responses for hits were determined for each subject (as noted in the main text, no differences were observed for miss trials in the amygdala or PHG). In the case of behavior, accuracy

differences were used for each subject. All scores were then entered into a three-variable mediation analysis.

To evaluate the significance of potential mediating effects, a standard approach was adopted, which involved evaluating the following components (Fig. S5) (13–15): (*i*) total effect *c* (initial variable → outcome, which can also be written in terms of the indirect effect *ab* + direct effect *c′*); (*ii*) indirect path *a* (initial variable → intervening variable); (*iii*) indirect path *b* (intervening variable → outcome after controlling for the initial variable); and (*iv*) direct effect *c′* (initial variable → outcome after controlling for the intervening variable). The final mediation effect was tested by assessing the product *ab* (for further details, see reference16). Statistical tests for path coefficients were performed by employing bootstrapping methods with 10,000 samples as implemented in the Mediation Toolbox by T. Wager and colleagues (27) (http://www.columbia.edu/cu/psychology/tor/).

CS$^+$: Hit vs. Miss. We performed a second, trial-by-trial-based path analysis. To do so, single-trial responses from the amygdala and the PHG were used together with single-trial outcome (hit or miss) in logistic regression analyses. Specifically, for each subject, trial-by-trial behavioral outcome (hit or miss) was predicted via simple logistic regression as a function of single-trial amygdala responses. In addition, trial-by-trial behavioral outcome (hit or miss) was predicted via multiple logistic regression as a function of single-trial responses in the amygdala and the PHG. To test the significance of the path coefficients, one sample *t*-tests across participants were performed (the mediation was again tested by assessing the product *ab*).

Note that an additional trial-by-trial path analysis involving frontoparietal regions was not performed. This is because a trial-by-trial analysis involved predicting hit/miss outcomes based on single-trial responses. Given that frontoparietal regions are clearly sensitive to attentional fluctuations in both T1- and T2-related processes, we believe that it is problematic to employ a single index of response strength (e.g., average response 6–8 s post-trial onset), because it confounds both T1 and T2 contributions. In other words, the low-pass nature of the BOLD response blurs together the contributions of the two processes. Critically, T1 processing likely differs between hit and miss trials (see also section "Responses in Frontoparietal Regions" below). On the other hand, the path analysis based on mean responses above did not suffer from the same problems, because it focused on parsing the contributions of the amygdala and frontoparietal regions to PHG responses for the case of two types of hit trial (CS$^+$ and CS$^-$).

Relationship Between Amygdala Response Magnitude and the Strength of the Visual Cortex to Behavior Relationship. We investigated the relationship between trial-by-trial fluctuations in amygdala responses and the strength of the association between visual cortex and behavior. To do so, we probed whether the trial-by-trial relationship between PHG and behavior depended on the magnitude of amygdala signals. To increase statistical power, we pooled data from all participants and binned trials based on the strength of amygdala activation. For each bin, all trials were used to determine the logistic regression slope between PHG response strength and behavior (hit vs. miss). The regression slopes were then correlated with the median amplitude of amygdala responses in each bin. A total of 14 bins were used to partition the range of amygdala responses into approximately equal number of trials per bin (Fig. S7). Note that the results were robust with respect to the specific number of bins used when we partitioned the data into 5 to 20 bins.

Supplemental Results and Discussion

Behavioral Performance of Dual- and Single-Task Trials ("Mock Session"). Behavioral sessions (without conditioning) conducted before the fMRI ones revealed typical AB behavioral results (Fig. S2*A*). T1 detection accuracy was 97.0% (SE = 0.7). Subjects showed impairment at detecting T2 scenes for short T1-T2 lags and better performance with longer lags [significant linear trend: F (1, 29) = 58.77, $P < 0.001$], indicating that T2 detection was lag-dependent. T2 detection accuracy when an 800-ms lag was used was not different from T2 detection accuracy during the single-task condition (i.e., T2-only trials) [t (29) = −1.55, n.s.], indicating full recovery from the AB. As expected, T2 detection during dual-task conditions did not differ for houses or buildings [t (29) = −0.37, n.s.], which demonstrates that subsequent differences observed following conditioning were not due to potential differences in low-level physical features.

Behavioral Performance of Dual-Task Trials (fMRI Session). T2 behavioral accuracy is reported in the main text. Mean T1 face identification accuracy was 96.0% (SE = 0.8), and did not differ as a function of the subsequent T2 scene conditions (CS$^+$, CS$^-$, or no-scene) [F (2, 58) = 1.14, n.s.].

Subjects were encouraged to avoid guessing and were instructed to choose the no-scene option when uncertain. For trials that physically did not contain a scene, subjects correctly chose "no-scene" (i.e., correct reject trials) 97.4% of the time. For trials that actually contained a scene, subjects incorrectly chose "no-scene" (i.e., miss trials) 27.1% of the time.

To check for the presence of a potential response bias during T2 responses, the pattern of false alarms was investigated. Specifically, we were interested in comparing the proportion of no-scene trials for which the subject produced the CS$^+$ scene response (M = 1.3%, SE = 0.5%) to the proportion of no-scene trials for which the subject produced the CS$^-$ scene response (M = 1.0%, SE = 0.3%). No significant differences were detected [t (29) = 0.71, n.s.], indicating that subjects did not show a response bias to report either CS$^+$ or CS$^-$ scenes. A similar comparison with incorrectly categorized scene trials (CS$^+$ scene reported in the presence of CS$^-$ scenes, M = 6.1%, SE = 1.1%; CS$^-$ scene reported in the presence of CS$^+$ scenes, M = 6.0%, SE = 0.1%) also did not reveal significant differences [t (29) = 0.94, n.s.]. Overall, both of these error trial types were rather infrequent and were excluded from fMRI analyses.

AB data are typically probed in terms of accuracy. However, given that fMRI responses may also depend on RTs as discussed above, we also probed the RT data, which were analyzed separately for T1 and T2 responses (Figs. S2 *B* and *C*). A repeated-measures ANOVA on T1 RT-based on T2 trial type (CS$^+$ hit, CS$^+$ miss, CS$^-$ hit, CS$^-$ miss, and cr) revealed a significant main effect [F (4, 116) = 5.04, $P < 0.001$] (note that T1 accuracy itself was near ceiling). Subsequent posthoc tests revealed that miss trials were significantly slower than hit and cr trials [t (29) = 2.82, $P < 0.05$; t (29) = 3.29, $P < 0.01$, respectively]. No significant RT differences were observed between hit trials (CS$^+$ vs. CS$^-$) or between miss trials (CS$^+$ vs. CS$^-$) (Ps > 0.38).

A similar analysis of T2 RT data also revealed a significant main effect [F (4, 116) = 13.25, $P < 0.001$]. Subsequent posthoc tests revealed that miss trials were significantly slower than both hit and cr trials [t (29) = 5.76, $P < 0.001$; t (29) = 3.51, $P < 0.005$, respectively]. Related findings were reported by De Martino et al. (9). No significant RT differences were observed between hit trials (CS$^+$ vs. CS$^-$) or between miss trials (CS$^+$ vs. CS$^-$) (Ps > 0.33).

SCRs and fMRI Responses During the Conditioning Phase. To explore the relationship between SCR and fMRI responses during the conditioning phase, we performed additional correlation analyses. The difference between CS^+ and CS^- fMRI responses in the right amygdala (Fig. S1) did not show a significant correlation with the differential SCRs, although a mild trend in that direction was observed ($r = 0.27, P = 0.165$). However, responses in the right PHG exhibited significant correlations with both SCRs ($r = 0.50$, $P < 0.01$) and right amygdala fMRI responses ($r = 0.39$, $P < 0.05$).

Responses in the FG. In the main text, T2-related responses in the right PHG were presented (Fig. S3*A* shows the results for the left PHG). To evaluate the specificity of the results in visual cortex, control analyses were performed in the FG ROI, which was defined based on separate localizer runs (like the PHG ROI). Because the FG responds strongly to faces, but weakly to scenes, we expected that responses in this region would mainly reflect T1-related processing. One-way repeated-measures ANOVAs (conditions: CS^+ hit, CS^+ miss, CS^- hit, CS^- miss, and cr) did not reveal significant main effects in the left [F (4, 88) = 0.30, n.s.] or right [F (4, 116) = 1.54, n.s.] hemisphere, consistent with the time series plots shown in Fig. S3 *C* and *D*. These results reveal that T2-related evoked responses were specific to the PHG and suggest that T1-related face processing in the FG did not vary as a function of subsequent T2-related processing.

Responses in the Amygdala: Conditioning by Decision Interaction. As described in the main text, a two conditioning (CS^+, CS^-) by two decision (hit, miss) repeated-measures ANOVA showed a significant main effect of decision, a trend of main effect of conditioning, and no significant interaction. As seen in Fig. S1*D*, although the difference (hit - miss) was larger for CS^+ vs. CS^- on average, the fair amount of variability observed in the CS^- case was likely responsible for the interaction not being detected via ANOVA. However, given that we used a trial-by-trial design, the comparison of the logistic regression slopes (CS^+ vs. CS^-) provided a different way in which an interaction-type result could be assessed—because each slope itself assessed the relationship between hits vs. misses. As stated in the main text, a significant difference was observed when contrasting the CS^+ and CS^- slopes [t (29) = 2.23, $P < 0.05$]. Plotting the results in Fig. S1*E* reveals that much less variability was observed in the estimated logistic regressions slopes. This was the case with the possible exception of one observation shown in green [removal of this data point affected the contrast of CS^+ and CS^- slopes only modestly: t (28) = 1.93, $P = 0.06$]. We thus believe that our results do indeed reveal that an interaction-type pattern was present in our data.

It is important to emphasize that the main goals of our investigation involved assessing hypotheses that were tested in a way that was independent from the interaction effects. Specifically, we were interested in differences observed during hit trials (CS^+ vs. CS^-) for the mediation analysis based on mean responses and between hit vs. miss trials for the mediation analysis based on trial-by-trial fluctuations. In other words, no effects depended heavily on the reliability of the interaction.

Responses in the Amygdala: Why Deactivations? An important study by Drevets and Raichle (17) showed a push-pull relationship between amygdala responses and those in "cognitive" brain regions, leading to the idea that emotion and cognition may be mutually suppressive. Accordingly, when subjects engage in highly demanding cognitive tasks, the amygdala is deactivated relative to a low-level, fixation baseline. We explicitly investigated this type of cognitive modulation of the amygdala in studies in which the cognitive load was parametrically varied (18, 19). In the present study, because the attentional blink is a very demanding cognitive task, the deactivation of the amygdala is in line with these earlier studies. Based on the considerations above, we would expect that even though amygdala responses decreased during the AB, differential responses in the amygdala would exhibit a pattern consistent with that observed in the PHG. Indeed, this was the case (e.g., hit responses during CS^+ trials evoked the largest responses), as further illustrated by considering that the scatterplots of Fig. 3 *B* and *D* are very similar (both used differential responses).

Responses in Frontoparietal Regions. In our study, we contrasted responses to hit and miss trials in visual cortex and the amygdala. In the main text, these responses were not discussed in the context of frontoparietal regions for the following reason. Because these regions are more strongly sensitive to attentional demands, their responses would be expected to fluctuate as a function of both T1- and T2-related processes. Fig. S4 summarizes the results obtained for these regions. For some regions, such as right MFG (Fig. S4*A*), comparable responses were obtained for hit and miss trials. In other cases, stronger responses were actually observed for miss relative to hit trials, such as the right IPL (Fig. S4*B*) and the ACC (Fig. S4*D*)—note that evoked responses shown here covaried out the effect of RT (as described in the Methods section above), such that the actual "raw" responses exhibited even larger responses during miss trials. Thus, our findings do not replicate the results by Marois and colleagues who reported stronger evoked responses for hit relative to miss trials in some of these regions (8). One possibility for this discrepancy is that in this previous study, subjects were asked to choose the response "unknown scene" when their confidence was low, and these trials were grouped with other hit trials. In our study, participants were instructed to choose "no-scene" when they were uncertain. Furthermore, although RTs in the AB need to be interpreted with care because responses take place after the presentation of the visual stream, our analyses above revealed that T1-related RT differed for trials in which T2 was associated with hit or miss. It is thus conceivable that such differential RTs reflected differential T1-related processing.

Why Choose Scenes and Not Faces as T2 Stimuli? To separate T1 and T2-related responses, we used faces (known to show preferential responses in the FG) and scenes (known to show preferential responses in the PHG). Because scenes such as houses and buildings do not typically have intrinsic emotional content, our goal was to investigate their behavioral and neural fate when they were imbued with value (i.e., by initially pairing them with mild electrical stimulation). Faces, on the other hand, are biologically relevant stimuli that may always carry some value, even when their expression is neutral. Indeed, neutral faces are known to evoke responses in the amygdala, as we observed in previous studies (e.g., 20).

Directionality of Network Interactions. One of the main goals of the current study was to characterize the network interactions that underlie the AB when affectively potent items are involved. In the Discussion section of the main text, we have described some of the reasons for hypothesizing the directionality of the interactions outlined in Fig. 1. Although the main motivation for proposing these specific interactions are conceptual, it is instructive to perform additional analyses in which the regions in Fig. 1 are switched around (i.e., their positions are exchanged). Fig. S8*A* illustrates additional analyses when the PHG and the amygdala were switched around and path analysis based on mean responses was attempted. Although PHG responses were correlated with behavior, no mediation through the amygdala was observed. In addition, the path between the amygdala and behavior was not significant after accounting for PHG responses.

No mediation was observed through the MFG either, as the path between the MFG and the amygdala was not significant (Fig. S8*B*). A mediation analysis was also attempted based on trial-by-trial data (Fig. S8*C*). As in the mean response-based analysis, although PHG responses were correlated with behavior, no mediation through the amygdala was observed, as the path between the amygdala and behavior was not significant after accounting for PHG responses.

For the mediation analysis based on mean responses, in addition to switching the amygdala and PHG positions (as done in Fig. S8*A*), we also evaluated four additional arrangements, as shown in Fig. S8 *D–G* (exhausting all remaining possible combinations). Only one model revealed evidence of a mediation relationship (Fig. S8*G*), although the interactions were not readily interpretable (i.e., "the influence of the MFG on PHG responses was partially mediated by the amygdala").

Fig. 6*B* in the main text illustrated the relationship between trial-by-trial fluctuations in amygdala responses and the strength of the association between visual cortex and behavior. Specifically, the higher the single-trial amygdala response, the steeper the slope of the logistic regression linking PHG responses and hit/miss trials. We performed a control analysis that attempted to investigate this issue by interchanging the functional roles of the amygdala and the PHG. In other words, we probed whether the trial-by-trial relationship between amygdala and behavior depended on the magnitude of PHG signals. As done in the context of the analysis of the main text, we pooled the data from all participants and binned trials based on the strength of PHG activation. For each bin, all trials were used to determine the logistic regression slope between amygdala response strength and behavior (hit vs. miss). The regression slopes were then correlated with the median amplitude of PHG responses in each bin. No significant correlations were observed for a series of analysis that partitioned the range of PHG responses in 5 to 20 bins (for each of the 16 distinct binnings, approximately equal number of trials were used in each bin). All correlation values were less than 0.2 (Ps > 0.6).

Minimal Functional Interactions vs. Comprehensive Network Models. In the present study, we opted to investigate "minimal functional circuits" that may help elucidate how network interactions are related to behavioral performance in the AB when affectively potent information is involved. An alternative strategy would have been to consider more complete models, involving a larger set of regions (e.g., 6–10 ROIs; see, e.g., reference 21), and to possibly test their interactions based on the framework of dynamic causal modeling (22). Our goal, however, was not to compare different models representing different hypotheses. For one, model comparison with more than, say, three regions presents formidable challenges due to the combinatorial nature of the possible models, and because criteria for choosing the best model often exhibit competing biases (e.g., Bayesian information criterion vs. Akaike's information criterion; see reference 23). More importantly, while bypassing such issues, our goal was to investigate simple interactions that have considerable explanatory power in linking brain and behavior during the AB. Although these interactions naturally should not be interpreted in causal terms, they provide testable hypothesis that should be further investigated with the use of more "causal methods," including TMS and lesion work; see also reference 24 for a related discussion in the domain of clinical trials.

Relationship to a Recent AB Study. A recent fMRI study by De Martino et al. (9) investigated the AB by using neutral and emotional faces as T2 targets (scenes were used as T1 targets). For each trial, participants were asked to identify the T2 face among a set of three faces of the same sex and expression (one of which was shown in the trial). Consistent with previous studies, behaviorally, participants were better at detecting an emotional T2 (58.8%) than a neutral one (46.8%). Unexpectedly, however, no differences were observed between evoked responses between emotional vs. neutral targets during hit trials in the two key regions described in the present paper, namely, the amygdala and visual cortex (i.e., fusiform cortex in the case of the De Martino study). On the other hand, De Martino et al. (9) reported differential activation in the FG between emotional and neutral targets during incorrect T2 detection. It is less straightforward to compare this result to those obtained in our study because of differences in experimental design. In particular, their paradigm did not include a no-face option, which would have allowed for the comparison of hit vs. miss trials in a way that was done in the present study. It is also important to emphasize that, in our study, across participants, neutral and affectively significant T2 stimuli differed only in terms of their learning histories and that the use of a slow event-related design allowed us to link trial-by-trial fluctuations in amygdala and visual cortex responses to observed behavior. Finally, our study also characterized the interactions between regions and how these were related to the behavioral enhancement of T2 detection when the target was affectively potent.

The AB and the Roles of the Right and Left Amygdala. A study of the AB with affectively significant stimuli reported that lesions of the left amygdala, but not the right, eliminated the behavioral advantage of emotional stimuli (25). As word stimuli were used, the authors suggested that such laterality may have been due to interactions between the left amygdala and structures in the left hemisphere that are important for language. In the present experiment, we observed instead, a stronger role for the right amygdala in the modulation of the AB with affective scene stimuli. Together, these findings suggest that the type of stimulus used may be an important factor determining the laterality of amygdala involvement.

1. Lundqvist D, Flykt A, Öhman A (1998) *The Karolinska Directed Emotional Faces* (Karolinska Institute, Stockholm, Sweden).
2. Raymond JE, Shapiro KL, Arnell KM (1992) Temporary suppression of visual processing in an RSVP task: An attentional blink? *J Exp Psychol Hum Percept Perform* 18:849–860.
3. Padmala S, Pessoa L (2008) Affective learning enhances visual detection and responses in primary visual cortex. *J Neurosci* 28:6202–6210.
4. Lim SL, Padmala S, Pessoa L (2008) Affective learning modulates spatial competition during low-load attentional conditions. *Neuropsychologia* 46:1267–1278.
5. Prokasy WF, Raskin DC (1974) *Electrodermal Activity in Psychological Research* (Academic, New York, NY).
6. Cox RW (1996) AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29:162–173.
7. Talairach J, Tournoux P (1988) *Co-Planar Stereotaxis Atlas of the Human Brain* (Thieme Medical, New York, NY).
8. Marois R, Yi D-J, Chun MM (2004) The neural fate of consciously perceived and missed events in the attentional blink. *Neuron* 41:465–472.
9. De Martino B, Kalisch R, Rees G, Dolan RJ (2009) Enhanced processing of threat stimuli under limited attentional resources. *Cereb Cortex* 19:127–133.
10. Genovese CR, Lazar NA, Nichols T (2002) Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15:870–878.
11. Yarkoni T, Barch DM, Gray JR, Conturo TE, Braver TS (2009) BOLD correlates of trial-by-trial reaction time variability in gray and white matter: A multi-study fMRI analysis. *PLoS ONE* 4:e4257.
12. Anderson AK (2005) Affective influences on the attentional dynamics supporting awareness. *J Exp Psychol Gen* 134:258–281.
13. Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 51:1173–1182.
14. Shrout PE, Bolger N (2002) Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychol Methods* 7:422–445.
15. MacKinnon DP, Fairchild AJ, Fritz MS (2007) Mediation analysis. *Annu Rev Psychol* 58:593–614.
16. MacKinnon DP (2008) *Introduction to Statistical Mediation Analysis* (Taylor & Francis Group, New York, NY).

17. Drevets WC, Raichle ME (1998) Reciprocal suppression of regional cerebral blood flow during emotional versus higher cognitive processes: Implications for interactions between emotion and cognition. *Cogn Emot* 12:353–385.
18. Pessoa L, Padmala S, Morland T (2005) Fate of unattended fearful faces in the amygdala is determined by both attentional resources and cognitive modulation. *Neuroimage* 28:249–255.
19. Hsu SM, Pessoa L (2007) Dissociable effects of bottom-up and top-down factors on the processing of unattended fearful faces. *Neuropsychologia* 45:3075–3086.
20. Pessoa L, McKenna M, Gutierrez E, Ungerleider LG (2002) Neural processing of emotional faces requires attention. *Proc Natl Acad Sci USA* 99:11458–11463.
21. Stein JL, et al. (2007) A validated network of effective amygdala connectivity. *Neuroimage* 36:736–745.
22. Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19:1273–1302.
23. Grunwald P (2007) *The Minimum Description Length Principle* (MIT Press, Cambridge, MA).
24. Kraemer HC, Wilson GT, Fairburn CG, Agras WS (2002) Mediators and moderators of treatment effects in randomized clinical trials. *Arch Gen Psychiatry* 59:877–883.
25. Anderson AK, Phelps EA (2001) Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature* 411:305–309.
26. Loftus GR, Masson ME (1994) Using confidence intervals in within-subject designs. *Psychon Bull Rev* 1:476–490.
27. Wager TD, Davidson ML, Hughes BL, Lindquist MA, Ochsner KN (2008) Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* 59:1037–1050.

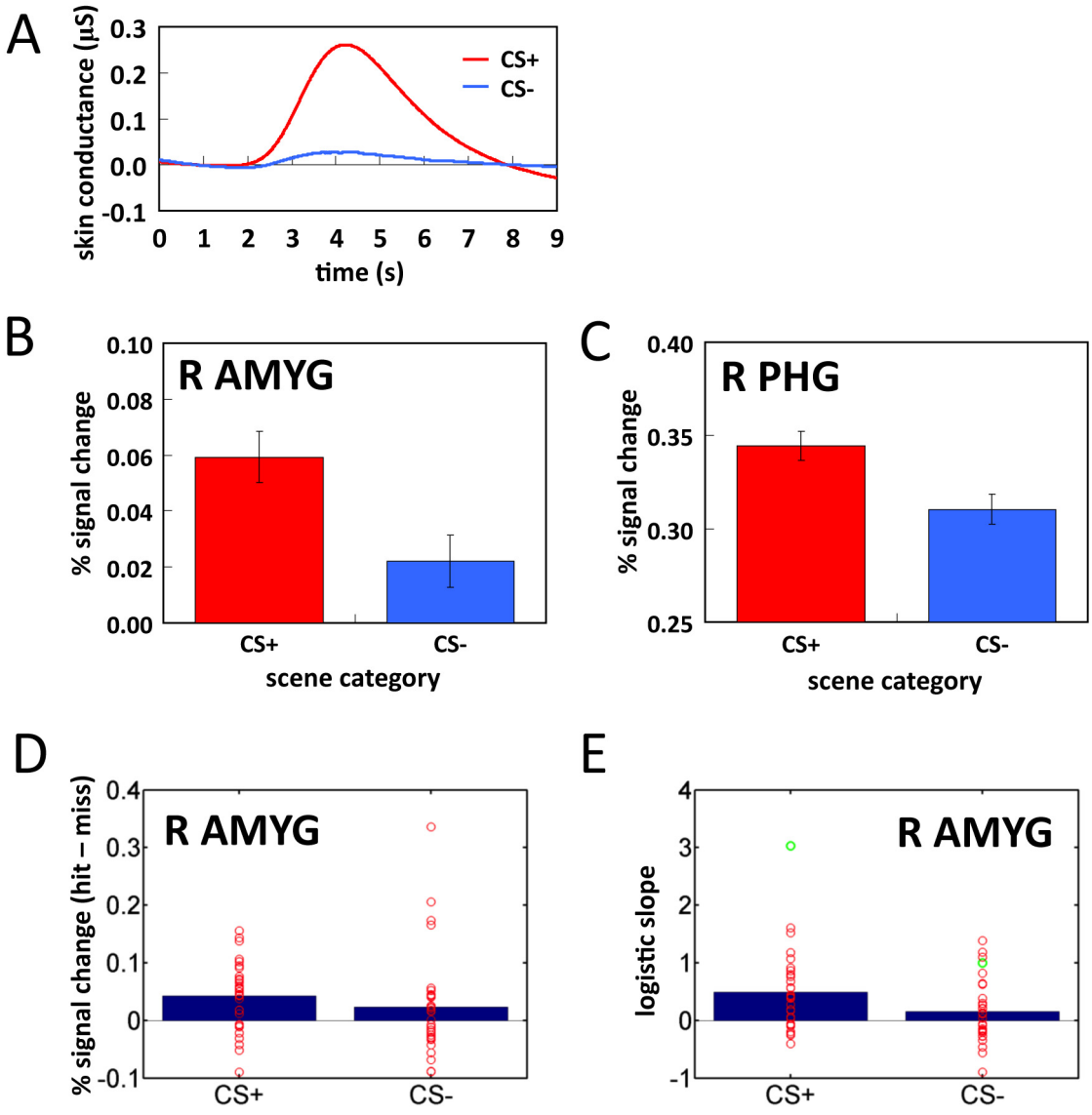

Fig. S1. SCRs and fMRI responses. (*A*–*C*) Results during the conditioning phase. (*A*) The SCR for each trial was indexed by the peak amplitude of 4–6 s post stimulus onset. (*B*) Right amygdala (AMYG) ROI data. (*C*) Right PHG ROI data. (*D* and *E*) Results during the dual-task phase. (*D*) Differential responses (hits - misses) plotted for the CS^+ and CS^- conditions in the right amygdala. (*E*) Logistic regression slopes plotted for the CS^+ and CS^- conditions in the right amygdala. Error bars in panels *B* and *C* denote the standard within-subject error term (26). The circles in panels *D* and *E* indicate individual data points.

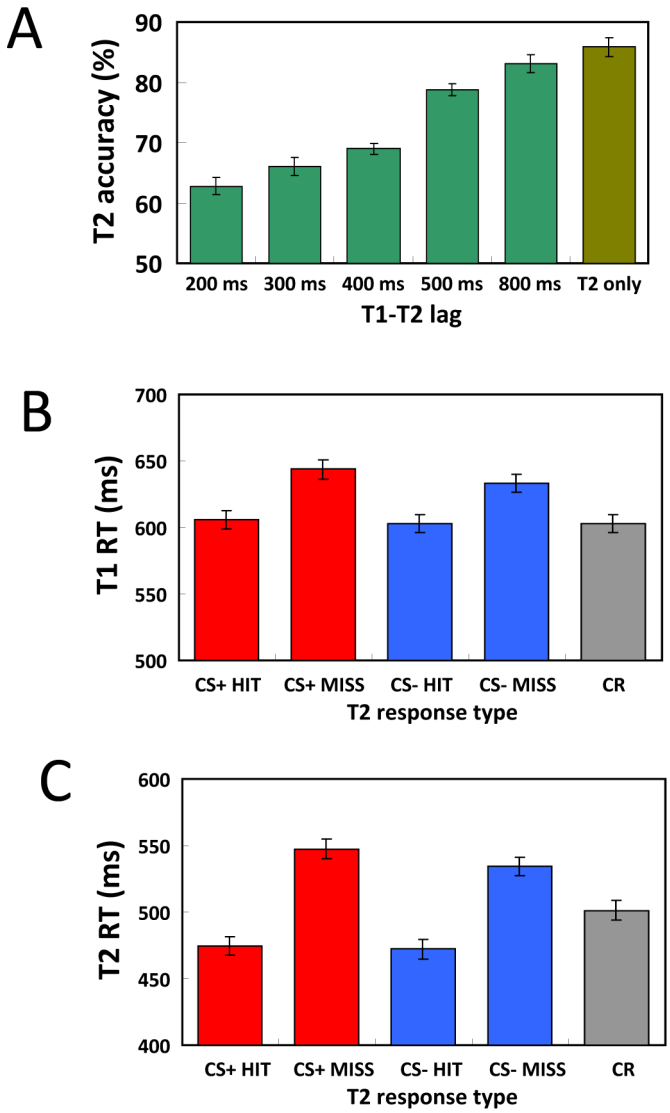

Fig. S2. (*A*) Behavioral performance during session outside the scanner. Mean T2 detection accuracy is shown as a function of T1-T2 lag and for T2-only trials. The typical patterns of T1-T2 lag dependence can be seen, such that the AB is fully eliminated with an 800-ms lag. During single-task trials, only the T2 scene target was presented. (*B*) T1-related RT data during fMRI sessions as a function of T2 response. (*C*) T2-related RT data during fMRI sessions as a function of T2 responses. Error bars denote the standard within-subject error term (26).

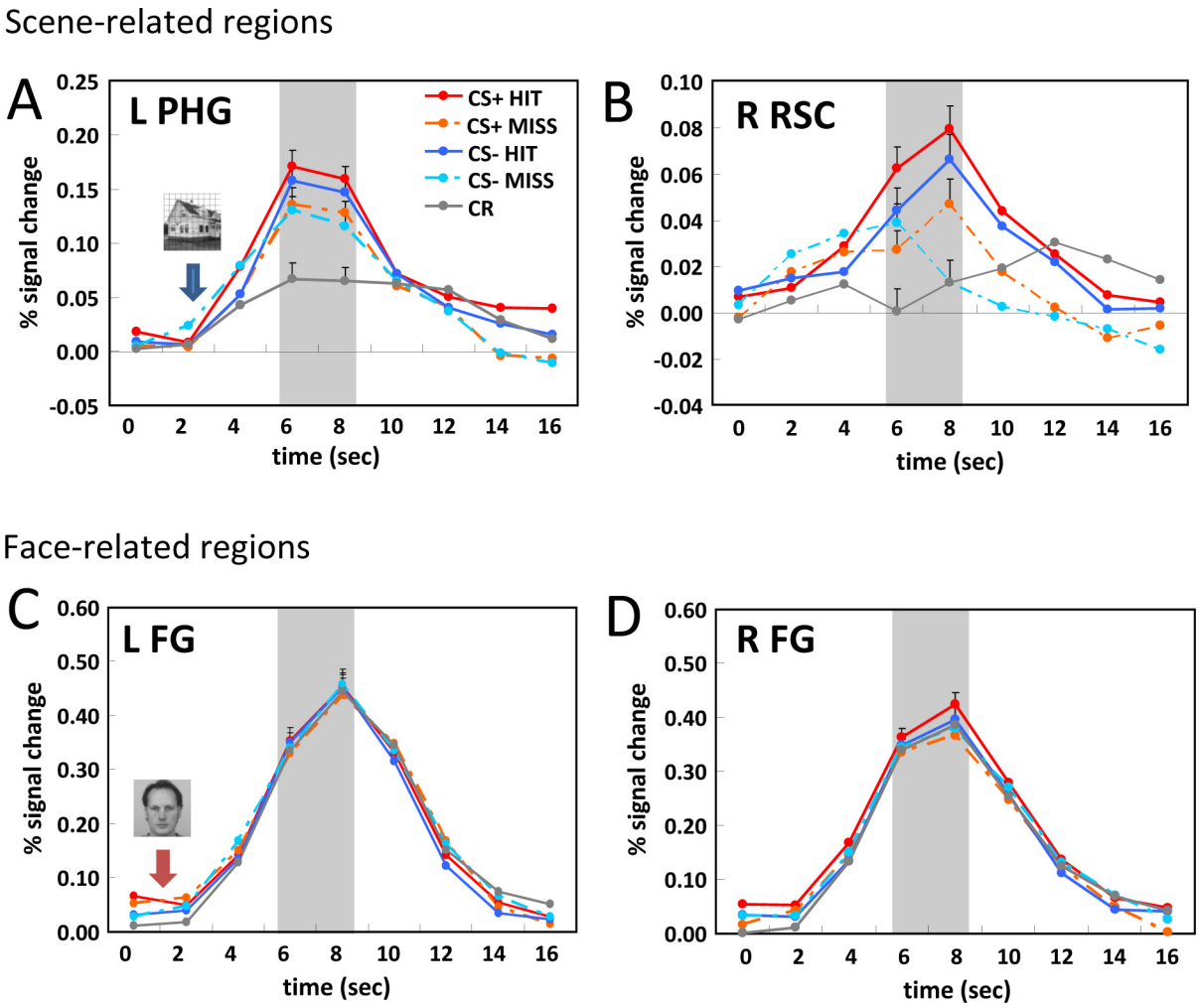

Fig. S3. Mean trial-evoked responses in visually responsive ROIs. (*A*) Left PHG. (*B*) Right retrosplenial cortex (RSC). (*C*) Left FG. (*D*) Right FG. Error bars denote the standard within-subject error term (26).

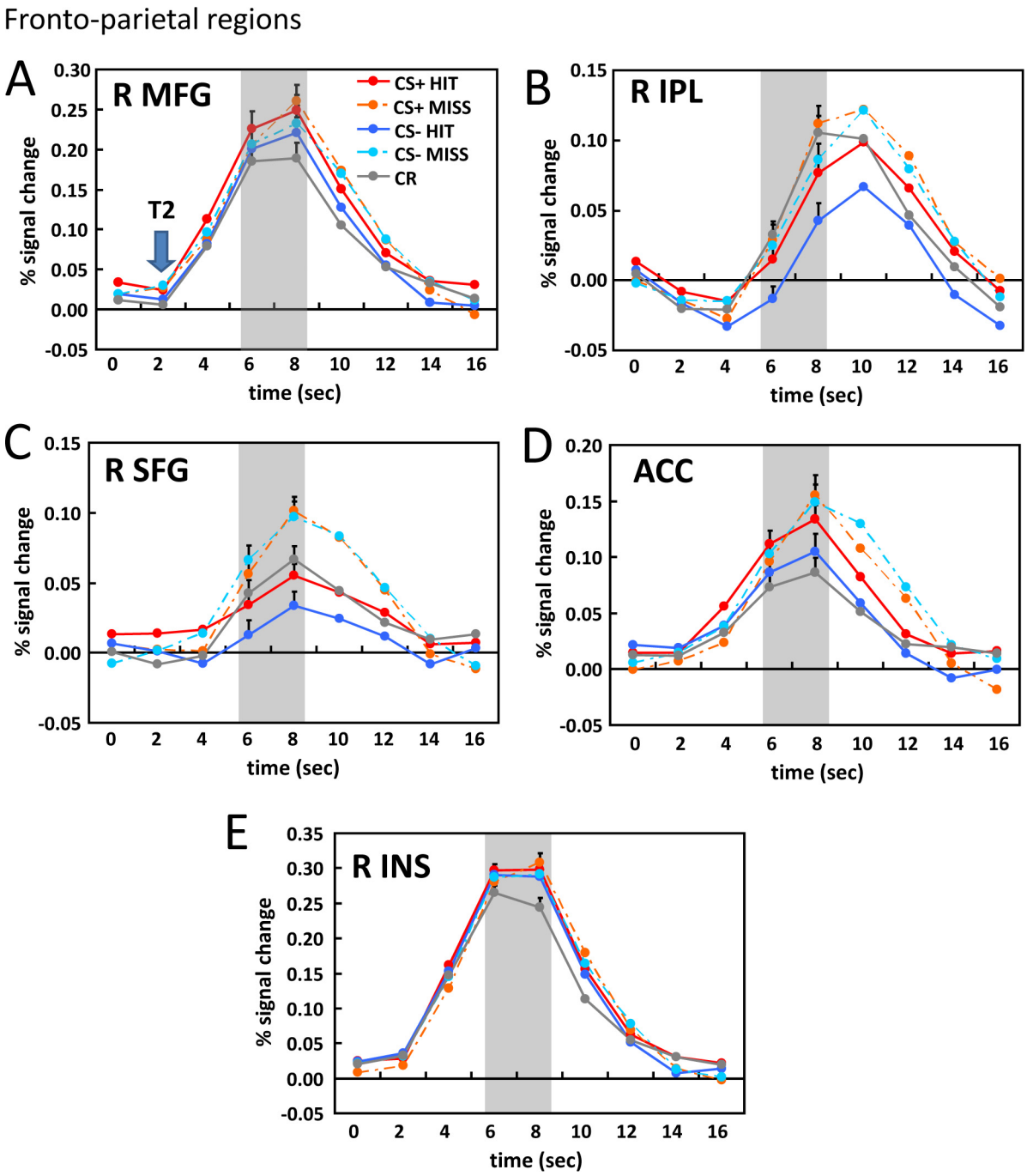

Fig. S4. Mean trial-evoked responses in frontoparietal ROIs. (*A*) Right MFG. (*B*) Right IPL. (*C*) Right superior frontal gyrus (SFG). (*D*) Anterior cingulate cortex (ACC). (*E*) Right anterior insula (INS). Error bars denote the standard within-subject error term (26).

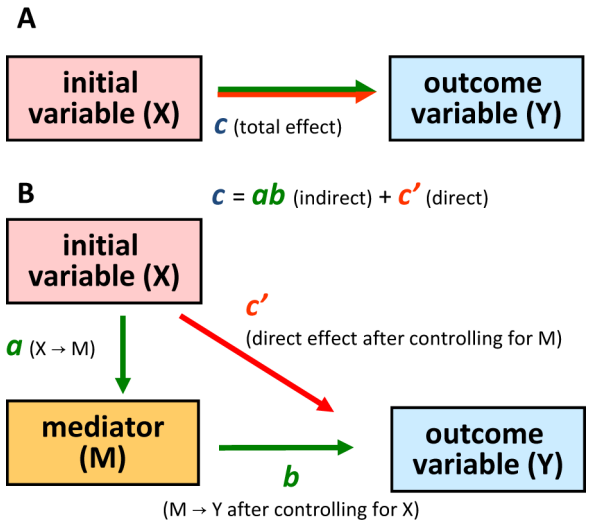

Fig. S5. Mediation analysis. (*A*) The simplest path model with two variables. (*B*) The total effect *c* of panel *A* can be expressed in terms of both an indirect effect *ab* involving a mediator variable, *M*, and a direct effect *c*′.

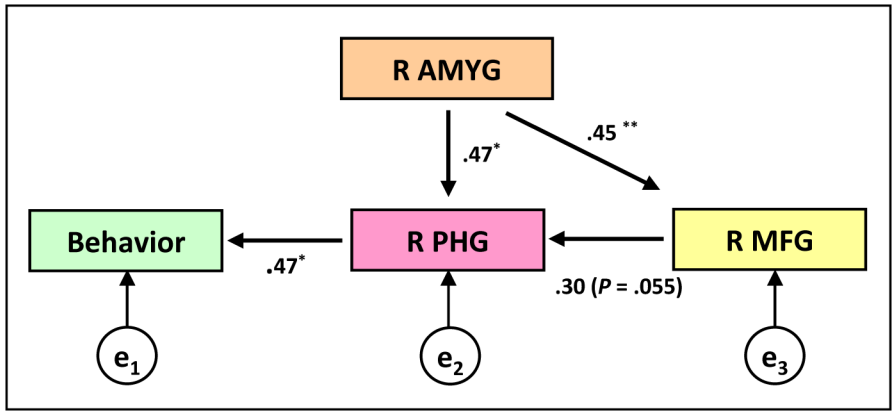

Fig. S6. Structural equation modeling (SEM) analysis. Goodness of fit index (GFI) = 0.981; adjusted goodness of fit index (AGFI) = 0.906. The path coefficients between the amygdala and behavior and between the MFG and behavior were not statistically significant ($Ps > 0.47$). The terms e_1, e_2, and e_3 indicate measurement errors. *, $P < 0.05$; **, $P < 0.01$; R, right; AMYG, amygdala; PHG, parahippocampal gyrus; MFG, middle frontal gyrus.

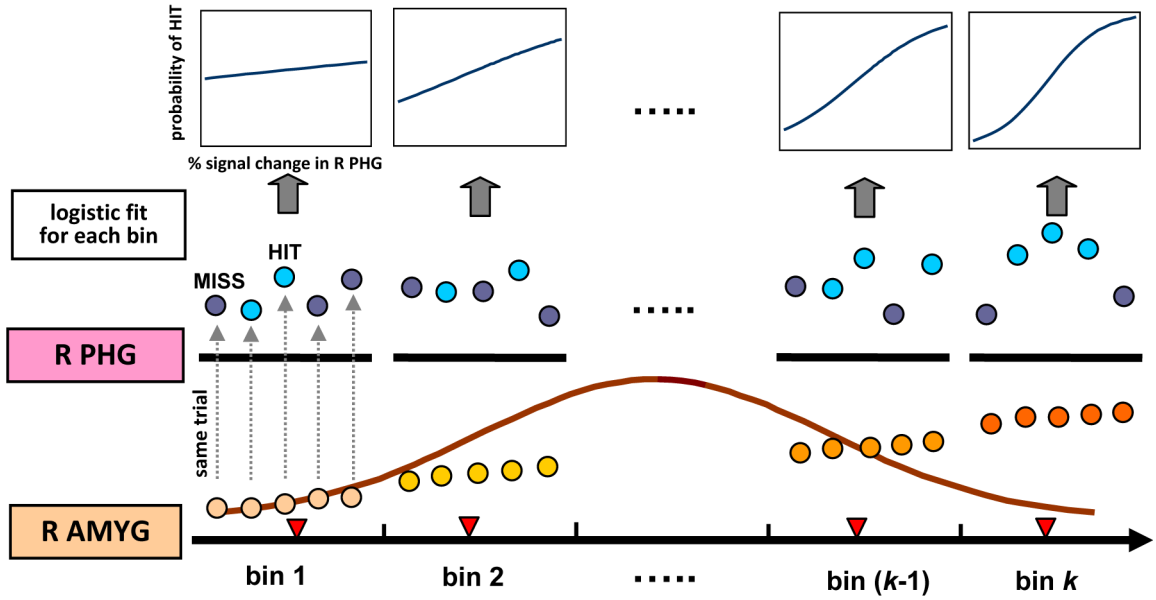

Fig. S7. Trial-by-trial fluctuations in the amygdala and the magnitude of the visual-cortex-to-behavior relationship. The schematic diagram highlights the main steps in assessing this relationship. The filled circles represent the magnitude of amygdala (AMYG, bottom row) and visual cortex (PHG, middle row) responses, as indicated via the height from the horizontal bars. The top row illustrates idealized logistic regression fits showing an increase in the slope value as a function of the magnitude of evoked responses in the amygdala. The red triangles indicate the positions of the median response for each amygdala bin, which were correlated with the slope of the logistic fit for that same bin. The results with actual data are shown in Fig. 6*B* of the main text.

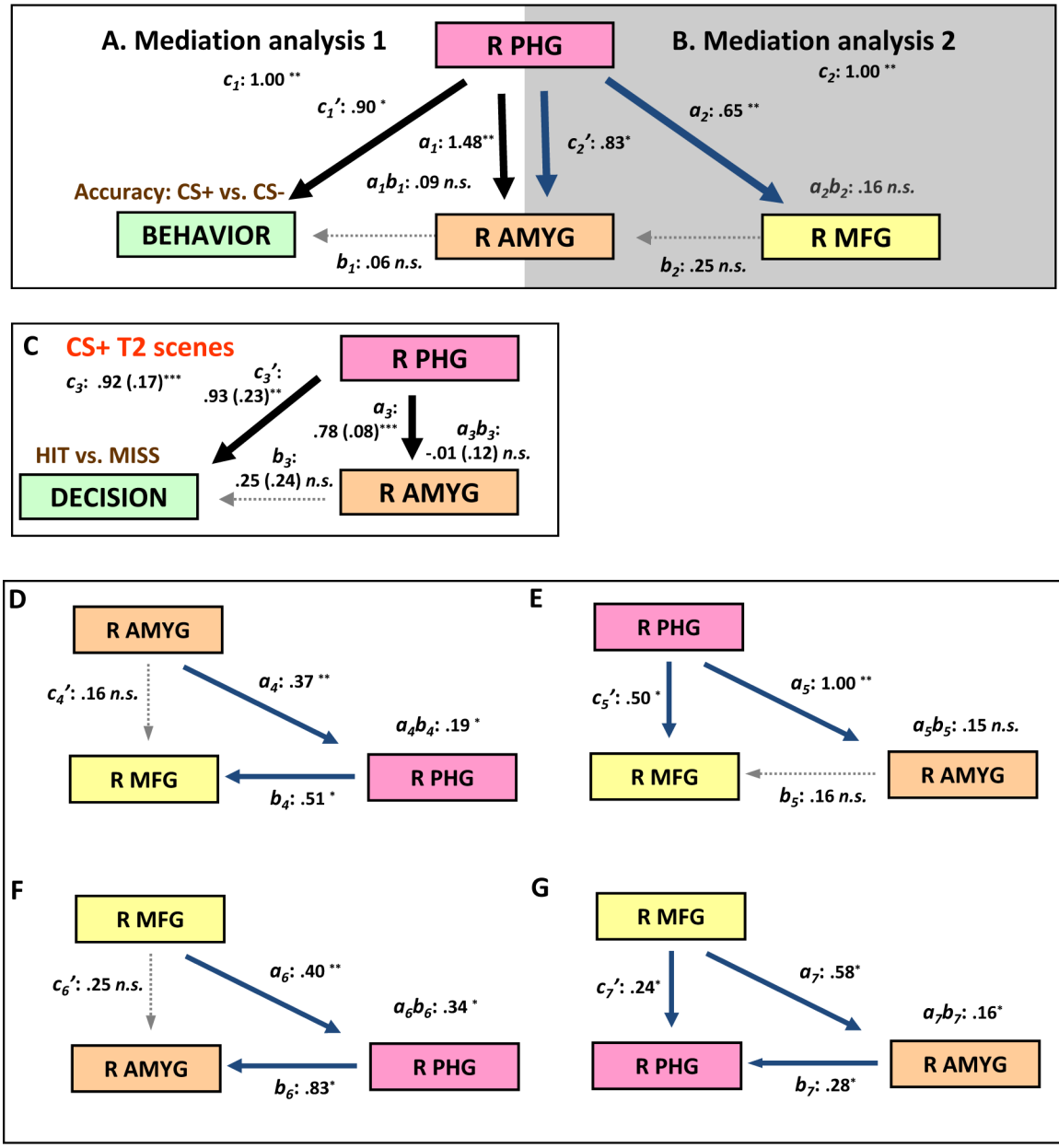

Fig. S8. Additional network analyses. (*A* and *B*) Mediation analyses based on mean responses when the positions of the amygdala and the PHG were exchanged relative to those of Fig. 4 of the main text. (*C*) Mediation analysis based on trial-by-trial responses when the positions of the amygdala and the PHG were exchanged relative to those of Fig. 6*A* of the main text. (*D*–*G*) Mediation analyses based on mean responses for the remaining configurations.

Table S1. ROI analysis for Hits (CS^+ vs. CS^-) and Misses (CS^+ vs. CS^-)

Region	L/R	Talairach *x*	Talairach *y*	Talairach *z*	ROI voxels	HIT: CS^+ vs. CS^-, *t*	MISS: CS^+ vs. CS^-, *t*
T2-related sensory regions							
Parahippocampal gyrus	L	Individual ROI ($N = 17$)			M = 27.6	1.30	0.72
	R	Individual ROI ($N = 30$)			M = 21.7	3.93***	0.12
Retrosplenial cortex (RSC)	R	11	−48	10	14	2.79**	0.95
T1-related sensory regions							
Fusiform gyrus	L	Individual ROI ($N = 23$)			M = 43.3	0.18	−0.51
	R	Individual ROI ($N = 30$)			M = 27.6	1.45	−0.57
Amygdala and thalamus							
Amygdala	R	28	−8	−10	11	2.66*	0.51
Thalamus	L	−15	−14	13	26	1.90	−1.18
	R	14	−22	16	14	0.82	−0.41
Fronto-parietal regions							
Middle frontal gyrus (MFG)	L	−40	21	27	33	3.32***	1.06
	R	46	19	24	18	3.21***	0.94
Anterior cingulate cortex (ACC)	L/R	−3	25	25	19	1.80	0.40
Anterior insula (INS)	L	−28	17	4	40	3.65***	0.91
	R	28	17	9	8	1.07	0.35
Superior frontal gyrus (SFG)	R	15	4	56	23	2.38**	−0.23
Inferior parietal lobule (IPL)	L	−28	−53	41	24	1.11	−0.70
	R	49	−47	33	26	2.99**	0.99
Anterior/posterior orbital gyrus	R	29	29	−2	14	2.58*	−0.33

$N = 30$; L, left; R, right; *, $P < 0.05$; **, $P < 0.01$, ***, $P < 0.005$. Individual ROI: Regions obtained based on separate "localizer" run. The number of subjects for whom the ROIs were extracted is indicated in parenthesis.