Supporting Information

Kuhl et al. 10.1073/pnas.1016939108

SI Results

Recall Accuracy at Posttest. Posttest trials were coded as recalled if subjects recalled at least some detail of the target image beyond the category (face vs. scene) with which they were cued. The reason for this liberal coding scheme was because subjects often provided detailed descriptions of the images without actually producing the specific label that corresponded to the image. Notably, in cases where details were recalled without the specific label, it was often not clear that the visual memories were any weaker than when the specific verbal label was correctly recalled. For example, for the image of Robert DeNiro, subjects may have recalled "actor, mafia guy," or for the Acropolis of Athens, subjects may have recalled "ancient ruins." Likewise, in some cases subjects described details specific to the photographs that were used (e.g., "building with interesting shadows"). Overall, the most frequent response type at posttest was successful recall of the target associate (recalled: mean = 54.8%). Of these trials, 63.1% corresponded to recall of the specific verbal label, and the remaining 36.9% corresponded to recall of details beyond the category but without the specific label. The next most frequent response type were trials on which subjects indicated they did not remember the target (don't know: mean = 39.6%), and finally trials on which subjects retrieved an incorrect associate (error: mean = 5.6%). AB pairs yielded more don't know responses than DE pairs (mean = 41.9% vs. mean = 37.3%; $t_{17} = 3.18$, $P < 0.01$); AB and DE pairs yielded a similar rate of errors (mean = 5.7% vs. mean = 5.6%; $t_{17} = 0.12$, $P = 0.91$). Note, however, that comparison of don't know and error responses for AB and DE pairs were nonindependent of the primary comparison of recall rate.

Importantly, target associates reported as specific hits during the initial retrieval phase were much better remembered at posttest than were those reported as general hits. Indeed, the majority (mean = 77.1%) of associates reported as specific hits at initial retrieval were successfully remembered at posttest, whereas subsequent retrieval occurred much less frequently for general hits (mean = 34.4%; $t_{17} = 13.69$, $P < 0.001$). Thus, overt recall performance at posttest validated the covert recall data collected during the scanned retrieval phases, because specific and general hit trials from the retrieval phase robustly differed in terms of retrieval success at posttest.

Classification of Encoding Trials. To confirm that the pattern classifier was able to discriminate face- and scene-related neural activity during the encoding rounds, we used a cross-validation technique, training and testing the classifier on subsets of the encoding data. Classification accuracy for encoding trials was near ceiling (mean = 98.3%), documenting the separability of face- and scene-related cortical patterns during encoding. Though classification accuracy did not differ for competitive (AC; 98.0%) vs. noncompetitive (AB/DE; 98.5%) encoding trials ($t_{17} = 0.72$, $P = 0.48$), a more-sensitive continuous measure of classifier evidence (*SI Methods*) revealed that competitive encoding trials were characterized by modestly weaker classifier evidence for the target visual category, relative to noncompetitive encoding trials (mean = 0.874 vs. mean = 0.884; $t_{17} = 2.22$, $P = 0.04$). It should be emphasized that competitive encoding trials were equivalent to noncompetitive trials in terms of the visual input—they differed only in terms of their mnemonic history. Thus, though competitive encoding trials produced VOTC responses that could be classified with high accuracy, there was nonetheless subtle evidence that past associations impacted VOTC responses during present encoding.

Univariate Analysis of AC Retrieval Events as a Function of Prior AB Reactivation. Behavioral and neural data from the present study suggest that B terms interfere with AC retrieval events. One consequence of this interference should be increased demands on cognitive control mechanisms during AC retrieval. To test whether the strength of AB associations impacted frontoparietal responses during AC retrieval, we created an additional univariate model in which, for each AC retrieval event, a regressor was included representing the fidelity with which corresponding B terms had been previously reactivated. In other words, if a B term had been strongly reactivated before AC encoding, did AC retrieval events display any evidence of this proactive interference? As with analyses in the main text, we separated AC retrieval events into high, medium, and low bins, reflecting the fidelity of corresponding AB reactivation; we then contrasted the high and low bins to test for regions displaying greater activation during AC retrieval when corresponding AB pairs had been reactivated with high fidelity. At a standard threshold ($P < 0.001$, five-voxel extent threshold) we observed a single cluster in the left midventrolateral prefrontal cortex (*pars triangularis*, Brodmann's area 45; MNI coordinates: $x = -51$, $y = 27$, $z = 9$). A highly consistent region of the left midventrolateral prefrontal cortex has previously been implicated in resolving proactive interference in working memory (1, 2) and during task switching (3), consistent with the idea that this region was engaged during AC retrieval in support of resolving interference from prior associations (i.e., AB pairs). Left mid-VLPFC activation has also been observed in situations where weak episodic memories are retrieved amid competition from more dominant memories (4, 5), again consistent with a role for this region in resolving mnemonic interference.

SI Discussion

Integration Across Associations. In the AB/AC paradigm used here, recall of AC pairs is putatively impaired because B terms interfere with retrieval of C terms. We suggest that this interference is reflected in the coactivation of B and C terms during AC retrieval. However, simultaneous reactivation of B and C terms may reflect integration across AB and AC associations. That is, during the encoding of AC pairs, B terms may have been reactivated (6) and integrated into the new (AC) associations (7). During AC retrieval, both terms would then be reactivated as part of an integrated set. We next consider several lines of evidence that potentially address the role of integration in the present study.

First, VOTC responses during encoding were at least consistent with the possibility that B terms were reactivated while AC pairs were encoded, as evidence for the target (presented) category was modestly lower for competitive (AC) than noncompetitive (AB/DE) events (*SI Results*). However, though reactivation of B terms during AC encoding would putatively be necessary for integration to occur, it would not, on its own, require that integration take place. Thus, though not inconsistent with an integration account, these encoding data do not compel such an account.

Second, our behavioral results, at first pass, suggest that integration was unlikely to play a major role, in light of evidence that integration powerfully reduces interference-related forgetting (8), whereas the behavioral costs of interference observed here were quite robust. However, a more subtle possibility is that integration of B and C terms did take place, eliciting simultaneous reactivation of B and C terms, but memory for C terms was impaired because relevant contextual information concerning the temporal order of the associations—i.e., which item was more recent—was not maintained. Importantly, forgetting of this type, related to forgetting the relevant temporal source (9), is qualitatively different from a failure to reactivate the target memory or outright forgetting of C terms. Thus, the presence of interference-related forgetting is also inconclusive with respect to establishing the role of integration.

Finally, and most informatively, we considered whether recall of B and C terms was independent. If B and C terms are directly integrated, memory for corresponding B and C terms should be positively correlated. However, we did not observe a relationship between recall of AC pairs during the scanned retrieval rounds and recall of AB pairs at posttest. Specifically, memory for AB pairs at posttest was similar if corresponding AC pairs were specific hits (mean = 52.5%) vs. other (general hits, don't know, or error; mean = 53.8%). Similarly, the percentage of specific hits during AC retrieval was similar if corresponding AB pairs were recalled (mean = 43.7%) vs. other (don't know or error; mean = 45.0%). Analysis of conditional independence (Mantel–Haenszel test) revealed no violation of independence ($\chi^2 = 0.085$, df = 1, $P = 0.77$). Though this null relationship argues against a strong integration account of the present data, the present study was not specifically designed to test this possibility, and the relationship between reactivation and integration is worthy of future study.

Reactivation During Competitive Encoding. Kuhl et al. (6) recently found that during new learning, older, competing memories are reactivated, thereby protecting them against forgetting. The relationship between competitive encoding and reduced forgetting of past events was mediated by the hippocampus and reflected in cortical and subcortical reactivation of contextual features. Although the present study differs in several regards, including materials and analysis approach, the two studies share several parallel findings. First, here we also observed a link between the hippocampus and cortical reactivation, as recall of specific event details was associated with both robust VOTC reactivation and marked hippocampal engagement (Table S1). Second, both of these studies indicate that reactivation of previously encoded, competing events during ongoing mnemonic processing (either encoding or retrieval) is associated with reduced competitor forgetting. Third, although we primarily focus on reactivation at retrieval in the present study, there was also evidence suggesting that competing memories were reactivated during encoding as well. Namely, relative to noncompetitive encoding, competitive encoding trials were associated with a subtle decrease in classifier-based evidence for the target (perceived) category (*SI Results*), indicating that overlap in mnemonic associations impacted VOTC responses during present encoding—putatively because previously encoded memories were reactivated. Indeed, pooling data across both studies, we observed a robust relationship between individual differences in hippocampal activation during competitive encoding and protection against competitor forgetting (Fig. S5). Specifically, greater hippocampal activation during competitive encoding was associated with less forgetting of competing events. Together, these data provide converging evidence that (*i*) the hippocampus supports cortical reactivation of past events and (*ii*) that reactivation of past events—whether it occurs during present encoding or retrieval—can protect against forgetting of events past.

Cognitive Control and Competitive Remembering. The present results reveal a relationship between frontoparietal BOLD responses and the fidelity of memory reactivation within the VOTC. These findings relate to prior work that has demonstrated that frontally mediated cognitive control mechanisms are engaged during competitive remembering (10). For example, Kuhl et al. (10) have shown that during competitive retrieval, responses in anterior cingulate cortex and right inferior frontal gyrus decrease in direct relation to the weakening of competing memories; these regions also display heightened responses when weak memories are retrieved in the face of more dominant memories (4). In the present study, competition was putatively maximal during AC retrieval events that were associated with low-fidelity VOTC reactivation. Indeed, direct interrogation of anterior cingulate and right inferior frontal regions of interest revealed marked activation during low-fidelity AC retrieval events (Fig. S3 *B* and *C*). Thus, complementing prior evidence that these regions are engaged in relation to the competitive status of retrieval targets, here we observed a relationship between the engagement of these regions and neural expressions of mnemonic competition within posterior cortical sites. These results provide evidence of a relationship between prefrontal regions that putatively support cognitive control operations and posterior cortical expressions of competitive visual remembering.

SI Methods

Participants. Subjects consisted of 18 (10 female), right-handed, native English speakers between the ages of 18 and 27 y (mean = 22 y). Informed consent was obtained in accordance with procedures approved by the Stanford University Institutional Review Board. Subjects were paid $20/h for their participation.

Materials. Stimuli consisted of 96 nouns, 72 images of faces, and 72 images of scenes. Nouns were drawn from the Medical Research Council Psycholinguistic Database (http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm) and ranged in length from four to eight letters (mean = 5.4), with a Kucera–Francis written frequency of at least 5 (mean = 20.7), and a concreteness rating of at least 500 (mean = 600). Faces consisted of black-and-white photographs of well-known male and female actors and musicians (e.g., Robert DeNiro). Faces included hair and varied in emotional expression, but were cropped such that other body parts were not visible. Scenes consisted of black-and-white photographs of well-known locations, including natural landscapes (e.g., Niagara Falls) and manmade structures (e.g., Empire State Building). All face/scene images were 225 × 225 pixels, with a resolution of 150 pixels/inch. An additional eight nouns, four faces, and four scenes were used as filler items (all with similar properties to the nonfiller stimuli). All stimuli were randomly assigned to conditions for each subject.

Procedures. ***Encoding rounds.*** Encoding trials (4 s) consisted of a single noun (cue) presented directly above either a face or scene (associate), with the name of the associate (e.g., Robert De Niro) presented beneath the image (Fig. 1). Subjects were instructed to try to remember the association between each cue and associate such that they would later be able to retrieve the associate when presented with the cue. No overt response was required during encoding trials.

During the first encoding round, all cues and associates were novel. During encoding rounds 2–7, some cues were repeated but paired with novel associates. In total, half of all cues were paired with two associates (competitive condition) and half were paired with one associate (noncompetitive condition). Critically, for cues in the competitive condition, if the first associate was studied during encoding round n (AB trials; A = cue, B = associate), the second associate was studied during encoding round $n + 1$ (AC trials). Thus, AB trials were evenly distributed across encoding rounds 1–6, whereas AC trials were evenly distributed across encoding rounds 2–7. Importantly, for cues in the competitive condition, one associate was always a face and one associate was always a scene (i.e., the B and C terms were always from distinct categories). Cues in the noncompetitive condition were paired with novel associates and evenly distributed across encoding rounds 1–6 (DE trials). Thus, the distinction between AB and DE pairs was entirely based on whether an overlapping (competing) pair was subsequently learned. In summary, the first encoding round contained AB and DE trials (pseudorandomly intermixed); encoding rounds 2–6 contained AB, AC, and DE trials (pseudorandomly intermixed); and the last encoding round contained AC trials as well as an equal number of filler trials consisting of novel cues and novel associates (pseudorandomly intermixed). In total, there were 48 encoding trials in each condition (AB, AC, DE). Half of the trials in each condition contained face associates; half contained scene associates. The number of face/scene associates was also balanced within each encoding round.

Each encoding trial was followed by an 8-s baseline period (Fig. 1) beginning with the presentation of a fixation cross (800 ms), followed by six arrows (800 ms each). Each arrow was followed by a brief fixation cross (400 ms). Subjects were instructed to indicate the direction (left/right pointing) of each arrow via a button box held in their right hand (11). Arrow orientation was randomly determined.

Retrieval rounds. Retrieval rounds probed subjects' memory for each of the pairs—and only those pairs—that were encoded in the immediately preceding encoding round. Each trial (5 s) consisted of a single cue presented above a square (equal in size to associate images). The interior of the square was black (matching the background screen color), thereby giving the impression of an empty box. The outline of the square was white for the first 4 s of the trial, then changed to red for 1 s, indicating that the trial was about to end. Subjects were instructed to covertly recall the associate that was presented with each cue in the immediately preceding study round. Subjects were made aware that, in some cases, a cue would be paired with more than one associate (AB, AC pairs), but that a cue would never be paired with more than one associate within a single encoding round. In such cases, subjects were instructed to always retrieve the associate from the immediately preceding encoding round—that is, the most recent associate. Subjects were not explicitly told that when a cue was paired with two associates, the associates would always be from different categories (i.e., one face, one scene).

Because memory for AB pairs was assessed during the retrieval round before corresponding AC pairs were encoded, both AB and DE retrieval trials were noncompetitive (as during encoding). Thus, for AB and DE trials, the B and E terms represented the retrieval targets, with no relevant competitors for these trials. For AC trials, the C term represented the retrieval target, and the previously encoded B term represented the competitor.

Subjects indicated their retrieval success by making one of five responses via the button box: (*i*) "don't know" indicated they could not remember anything about the associate; (*ii*) "face–specific" indicated they successfully recalled the specific associate and it was a face; (*iii*) "face–general" indicated they recalled that the associate was a face but could not recall the specific image; and likewise for (*iv*) scene–specific and (*v*) scene–general. For all subjects, don't know responses corresponded to the subject's right thumb; the assignment of the remaining four responses was counterbalanced across subjects. Subjects could make their response at any point during the trial; no emphasis was placed on responding quickly.

Retrieval trials were followed by a 7-s baseline period during which a fixation cross was presented; no responses were required during this period.

Posttest. After exiting the scanner, memory for all AB and DE pairs was assessed again in a single posttest, allowing for measurement of the consequence that AC encoding/retrieval had on memory for AB pairs, relative to DE pairs. Notably, the posttest differed from retrieval rounds in that it involved overt retrieval of the associate images, thus allowing for validation of subjects' covert responses collected during the scanned retrieval rounds. On each trial (5 s), subjects were presented with a cue, above a square, as during retrieval. However, within each box was the word "face" or "scene," which oriented subjects to the category of the associate that they were to retrieve. Subjects were instructed that for each trial, they should retrieve the associate that they previously studied with the presented cue and that matched the indicated category cue (face/scene). Subjects were explicitly told that if a cue had been paired with more than one associate, here they would always be cued to retrieve the first associate. Subjects were instructed to respond aloud and to indicate the name of the associate image as it appeared during the encoding rounds, but, if they could not remember the name of the image, they should indicate any additional details that they did remember (e.g., male or female, manmade structure or natural scene, etc.). Because each posttest trial included a cue indicating the category of the target associate (face/scene), if subjects simply indicated the category of the associate (i.e., repeated aloud the category cue), this was not recorded as an actual response. Thus, whereas memory for the target category was an acceptable response during the retrieval rounds, this was not the case for the posttest.

AB and DE pairs were pseudorandomly intermixed during the posttest, equating for average testing position of each condition. Each trial was followed by a 1-s fixation cross.

fMRI Data Analysis. ***Preprocessing.*** Functional data were corrected for slice timing and head motion. Structural images were coregistered to functional images and segmented into gray matter, white matter, and cerebrospinal fluid. Gray matter images were stripped of remaining skull and normalized to a gray matter Montreal Neurological Institute template. Normalized gray matter images were used for normalization of the structural and functional images. Images were resampled to 3-mm cubic voxels and smoothed with a Gaussian kernel (8 mm FWHM).

Univariate analyses. Data were analyzed under the assumptions of the general linear model (GLM). Trials were modeled using a canonical hemodynamic response function and its first-order temporal derivative. Encoding and retrieval data were modeled separately with scan session (round) treated as a covariate (details of each GLM are included below). Linear contrasts were used to obtain subject-specific estimates for each effect of interest, which were then entered into a second-level, random-effects analysis using a one-sample t test against a contrast value of zero at each voxel. Unless otherwise noted, a threshold of $P < 0.001$, uncorrected, with a five-voxel extent threshold, was used for group-level contrasts. All contrast maps were overlaid on a mean anatomical image. Unless otherwise noted, region-of-interest (ROI) analyses were performed by extracting beta values from all significantly active voxels within a 6-mm radius of local maxima.

GLM for encoding data. A single GLM was generated for analysis of the encoding data. The model included regressors for six conditions, representing the two visual categories (faces vs. scenes) and three pair types (DE, AB, and AC pairs). A seventh regressor represented filler pairs that were included in the experiment.

GLM for retrieval data. Three GLMs were generated for analysis of the retrieval data. First, a model was generated that represented two levels of retrieval success (specific hits vs. general hits/don't knows) and the three pair types (DE, AB, AC), plus a regressor representing all excluded and filler trials. This model was used to identify regions of interest that were sensitive to detailed episodic retrieval.

Second, a model was generated that included regressors representing the visual category of the target image (face or scene), the pair type (DE, AB, AC), and the strength of classifier-based evidence for the target image (low, medium, high). Importantly, the classifier evidence bins were separately generated within each condition. For example, the 24 face-AB trials were sorted according to the strength of target evidence; the lowest eight trials constituted the low bin, the middle eight trials the medium bin, and the highest eight trials the high bin; likewise for the remaining five conditions (scene-AB, face-AC, scene-AC, face-DE, and scene-DE). A separate regressor was included to represent filler trials.

Finally, a third model was generated specifically for AC trials that included regressors representing the visual category of the target image (face or scene) and the strength of classifier-based evidence for target reactivation during the preceding AB retrieval trials. That is, AC retrieval events were modeled in terms of how strongly corresponding B terms had previously been reactivated. As above, responses were binned into three groups (low, medium, high). Separate regressors represented filler trials and AB/DE trials.

Multivoxel pattern analyses. All fMRI data used for classification analyses were high-pass filtered (0.01 Hz), detrended, and z-scored (mean response for each voxel across time = 0). Classifier analyses were based on penalized logistic regression using L2-norm regularization. All classification analyses involved training a classifier on a sample of fMRI data and testing the classifier on a distinct sample of data. To assess reactivation during retrieval, classification analyses were conducted for which the training set consisted of the encoding data and the testing set consisted of the retrieval data. For these cross-phase analyses, a penalty parameter of 100 was used. Additionally, classification analysis of the encoding data was performed, for which all but one of the encoding rounds constituted the training set, and the left-out set constituted the testing set; classification was then repeated such that every encoding round contributed to both the training and testing sets. For these intraphase analyses, a penalty parameter of 10 was used. Penalty parameters were based on preliminary analyses and not the result of optimization procedures.

For all classification analyses, voxel inclusion was restricted using an anatomically defined mask of the VOTC. The mask was generated using the Anatomical Automatic Labeling atlas (http://www.cyceron.fr/web/aal__anatomical_automatic_labeling.html) and consisted of the union of the masks labeled as left fusiform, right fusiform, left parahippocampal, and right parahippocampal. The mask consisted of 2,553 total voxels (Fig. S2). No additional feature selection was performed. Our use of a VOTC mask—as opposed to using a whole-brain classifier—was motivated by three factors: (*i*) prior work has shown the VOTC to be highly sensitive to differences between face and scene perception; (*ii*) we were interested in characterizing mnemonic reactivation specifically within higher-level visual areas, and (*iii*) reducing the dimensionality of the classification space in a principled way has been shown to improve classification performance by eliminating uninformative voxels (12).

All of the classifications considered in the present study represented classification between face vs. scene categories. For each trial in the testing set, the logistic regression classifier generated a scalar probability estimate that the trial corresponded to a face vs. scene (by construction, these probability estimates summed to unity). On each trial, the classifier's guess represented the category with the higher probability and was coded as correct or incorrect based on whether the guess corresponded to the target category for that trial. Critically, for retrieval trials, the target category represented the category of the image that subjects were intended to retrieve, as opposed to the category of what they reported retrieving.

Classification data were considered in three ways. First, we computed classification accuracy—the percentage of trials that the classifier correctly categorized. Second, we computed mean classifier evidence—the average probability that the classifier assigned to the target category of each trial. Capitalizing on the fact that the classifier's predictions were probabilistic rather than binary, this measure of classification performance potentially provided a more sensitive index of category discriminability than classification accuracy. Third, receiver operating characteristic (ROC) curves were generated for some of the core analyses. ROC curves were constructed by ranking the trial-by-trial classifier evidence scores according to how strongly each favored faces vs. scenes, and then charting the relationship between the true positive rate [arbitrarily defined here as the probability of classifying a face as a face, or P (face|face)] and the false positive rate [P (face|scene)] across a range of potential classifier decision thresholds. AUC was used as an index of classifier performance derived from the ROC curves, reflecting the mean accuracy with which a randomly chosen pair of trials from each class would be correctly classified.

Given that each trial in the encoding and retrieval phases corresponded to several volumes of fMRI data (six 2-s volumes per trial), trial-level classifier data were obtained by averaging several temporally contiguous volumes to generate a single brain activity pattern for each trial. For data from the encoding phase, TRs 3–4 (corresponding to 4–8 s poststimulus onset) were averaged; for the retrieval phase, TRs 3–6 (corresponding to 4–12 s poststimulus onset) were averaged. A wider window was used for retrieval trials because of variability in subjects' reaction times. Though all statistical analyses were based on trial-level classifier data, in some cases we display TR-by-TR classification performance for retrieval trails. In such cases, classification was separately applied to each of the six volumes corresponding to a single retrieval trial, producing a time course of classification performance. TR-by-TR classification is reported for illustrative purposes and to confirm that temporal evolution of classifier performance generally conformed to the shape of the canonical hemodynamic response function.

1. Jonides J, Smith EE, Marshuetz C, Koeppe RA, Reuter-Lorenz PA (1998) Inhibition in verbal working memory revealed by brain activation. *Proc Natl Acad Sci USA* 95:8410–8413.
2. Badre D, Wagner AD (2005) Frontal lobe mechanisms that resolve proactive interference. *Cereb Cortex* 15:2003–2012.
3. Badre D, Wagner AD (2006) Computational and neurobiological mechanisms underlying cognitive flexibility. *Proc Natl Acad Sci USA* 103:7186–7191.
4. Kuhl BA, Kahn I, Dudukovic NM, Wagner AD (2008) Overcoming suppression in order to remember: Contributions from anterior cingulate and ventrolateral prefrontal cortex. *Cogn Affect Behav Neurosci* 8:211–221.
5. Wimber M, et al. (2008) Neural markers of inhibition in human memory retrieval. *J Neurosci* 28:13419–13427.
6. Kuhl BA, Shah AT, DuBrow S, Wagner AD (2010) Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. *Nat Neurosci* 13:501–506.
7. Shohamy D, Wagner AD (2008) Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron* 60:378–389.
8. Anderson MC, McCulloch JC (1999) Integration as a general boundary condition on retrieval-induced forgetting. *J Exp Psychol Learn Mem Cogn* 25:608–629.
9. Johnson MK, Hashtroudi S, Lindsay DS (1993) Source monitoring. *Psychol Bull* 114:3–28.
10. Kuhl BA, Dudukovic NM, Kahn I, Wagner AD (2007) Decreased demands on cognitive control reveal the neural processing benefits of forgetting. *Nat Neurosci* 10:908–914.
11. Stark CEL, Squire LR (2000) Functional magnetic resonance imaging (fMRI) activity in the hippocampal region during recognition memory. *J Neurosci* 20:7776–7781.
12. Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends Cogn Sci* 10:424–430.

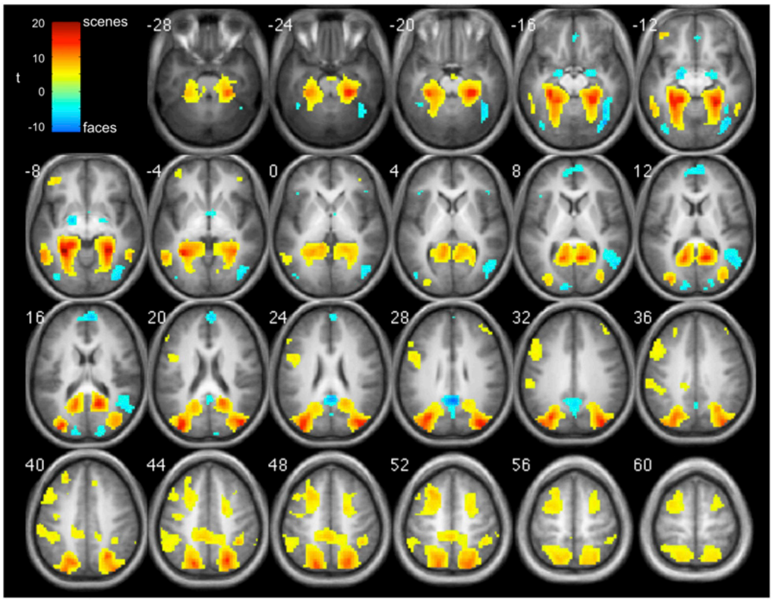

Fig. S1. Category-sensitive neural responses during encoding. Warm colors indicate regions that preferentially responded to scenes; cool colors indicate regions that preferentially responded to faces. $P < 0.001$, five-voxel extent threshold for each tail.

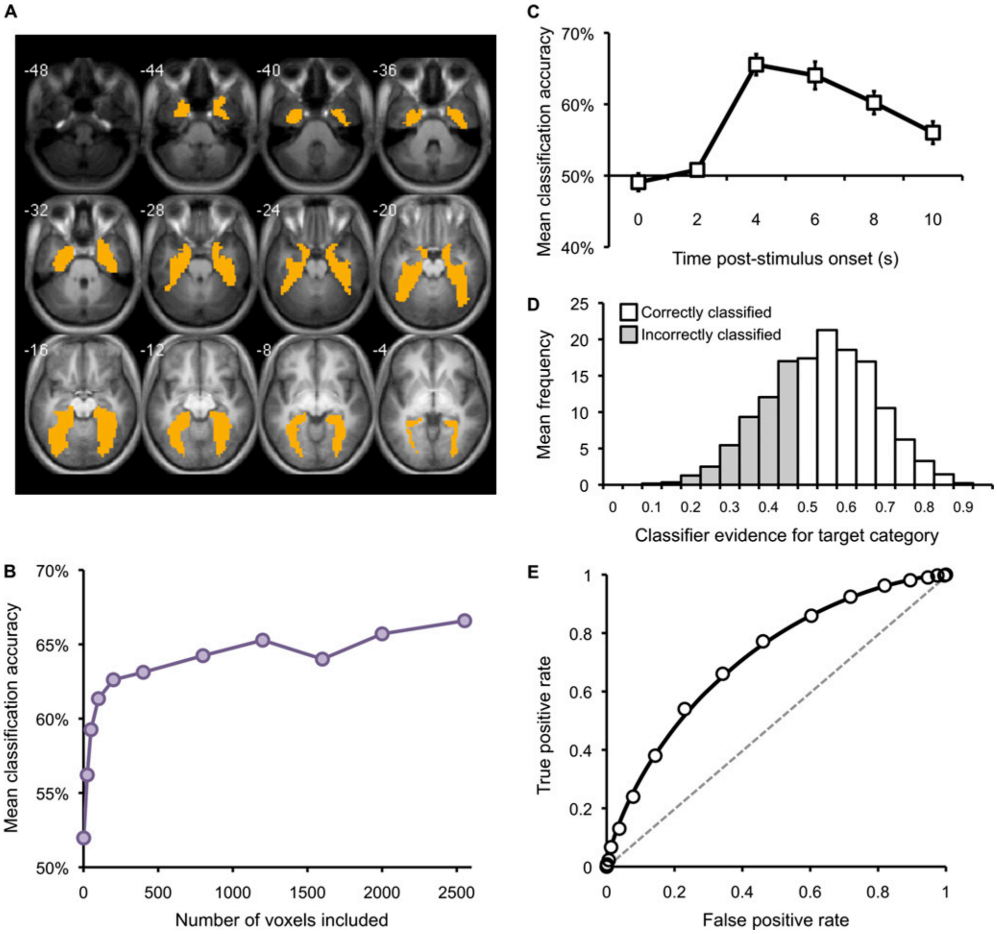

Fig. S2. Voxel selection for pattern classification. (*A*) Mask of all VOTC voxels used for pattern classification analyses. (*B*) Mean classification accuracy of all retrieval trials as a function of the number of VOTC voxels used for classification. (*C*) Classification accuracy for all retrieval trials as a function of acquisition volume (2 s per volume). Error bars represent SEM. (*D*) Distribution of continuous measure of classifier evidence for all retrieval trials. (*E*) ROC curve showing true positive rate [arbitrarily defined as *P* (face|face)] as a function of false positive rate [*P* (face|scene)]. AUC for fitted ROC curve = 0.713.

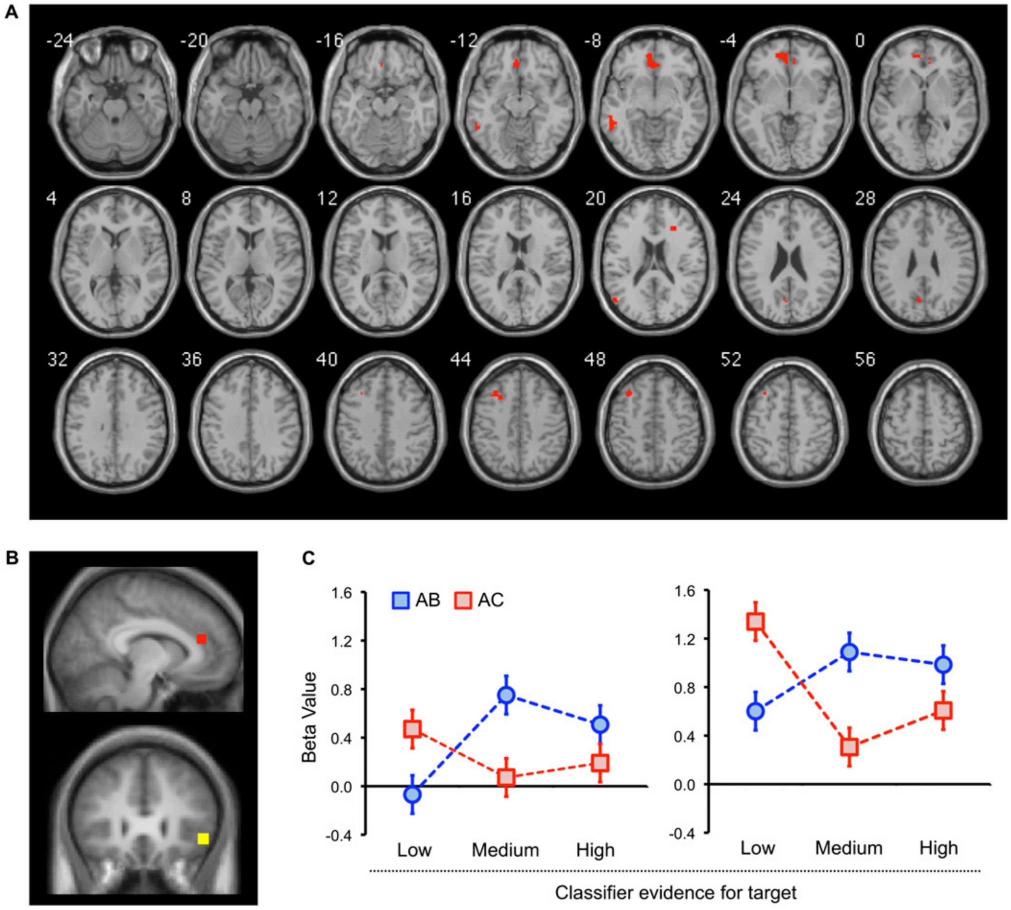

Fig. S3. (*A*) Results of conjunction analysis described in main test, revealing regions that displayed each of the following criteria: (*i*) a main effect of retrieval success for DE pairs (specific hit trials > general hit and don't know trials), $P < 0.005$, and (*ii*) an interaction between degree of classifier-based evidence for VOTC reactivation and level of competition [(high- > low-evidence AB trials) > (high- > low-evidence AC trials)], $P < 0.005$. (*B*) Anterior cingulate cortex and right inferior frontal gyrus ROIs generated from foci previously implicated in competitive retrieval (10). Anterior cingulate cortex ROI: 6-mm sphere, centered at MNI coordinates $x = -9$, $y = 36$, $z = 18$. Right inferior frontal gyrus ROI: 6-mm sphere centered at $x = 48$, $y = 27$, $z = -6$. (*C*) Beta values from each ROI revealed an interaction between classifier evidence and competition (anterior cingulate cortex: $F_{2,34} = 6.82$, $P < 0.005$; right inferior frontal gyrus: $F_{2,34} = 5.10$, $P < 0.05$). For each ROI, this interaction was characterized by greater responses for low-evidence AC trials than low-evidence AB trials (anterior cingulate cortex: $t_{17} = 2.45$, $P < 0.05$; right inferior frontal gyrus: $t_{17} = 2.81$, $P < 0.05$). Error bars indicate within-subject SEM.

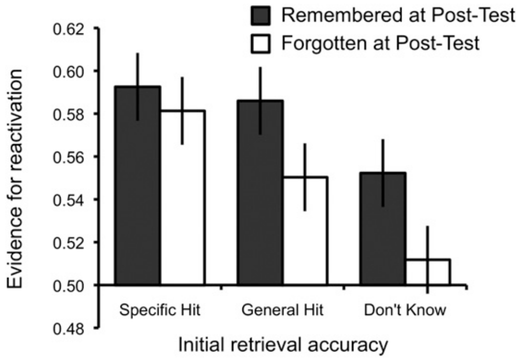

Fig. S4. Fidelity of reactivation for noncompetitive retrieval trials (AB and DE pairs) as a function of initial retrieval success and subsequent memory at posttest. Data are from 13 subjects (five subjects were excluded because there were zero trials in at least one of the relevant bins). A significant main effect of subsequent memory at posttest ($F_{1,12} = 12.73$, $P < 0.005$) indicated that higher-fidelity reactivation of AB/DE pairs during initial retrieval attempts was associated with better subsequent memory for these pairs at posttest. This effect did not interact with initial retrieval success ($F < 1$). Error bars indicate within-subject SEM.

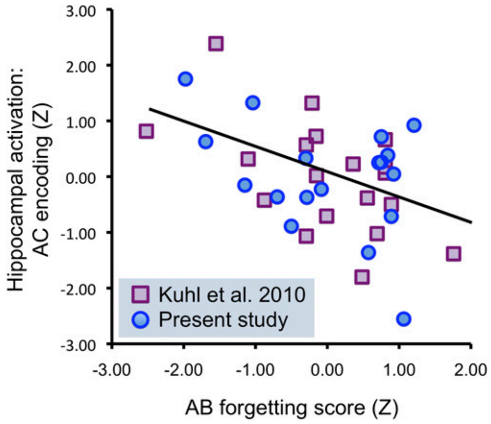

Fig. S5. Between-subject correlation between hippocampal activation during AC encoding and subsequent memory for AB pairs. Data are pooled from a prior study (6) ($n = 19$) and the present study ($n = 18$). Across studies, there was a significant negative correlation ($r = -0.47$, $P < 0.005$), indicating that greater hippocampal activation during AC encoding was associated with reduced forgetting of AB pairs. The strength of the correlation did not significantly differ across studies ($P > 0.5$). Data were normalized to Z scores within each study. Hippocampal activation reflected the beta value during AC encoding for two independently generated hippocampal regions of interest drawn from the prior study (6). Specifically, for the data extracted from the prior study, an ROI was generated from a within-subject contrast of AC encoding trials associated with subsequent remembering vs. forgetting of corresponding AB pairs, as described previously (6). For the data extracted from the present study, the hippocampal ROI was generated from the between-subject regression analysis reported previously (6) relating AC encoding activation to subsequent AB memory. AB forgetting represented the proportional forgetting of AB pairs, as expressed at posttest [(DE recall – AB recall)/(DE recall)].

Table S1. Regions displaying greater activation for specific hits vs. general hits/don't knows (DE pairs only)

			MNI coordinates		
Region	BA	Z	*x*	*y*	*z*
Hippocampus	–	4.86	−24	−12	−12
Hippocampus	–	4.73	27	−9	−15
Angular gyrus	39	4.58	−45	−66	33
Supramarginal gyrus	40	3.74	−48	−48	33
Hippocampus	–	4.23	36	−33	−9
Cerebellum	–	3.82	30	−69	−24
Cerebellum	–	3.28	42	−54	−27
Middle temporal gyrus	21	3.69	−57	−36	−9
Middle/superior temporal gyrus	21, 22	3.60	−57	−45	0
Middle temporal gyrus	21	3.63	−57	−12	−15
Angular gyrus	39	3.61	48	−72	36
Angular gyrus	39	3.58	39	−63	30
Cerebellum	–	3.60	−30	−36	−33
Postcentral gyrus	43	3.45	−60	−21	21
Postcentral gyrus	43	3.32	−51	−21	21
Cerebellum	–	3.40	0	−63	−12
Cerebellum	–	3.27	0	−72	−6
Middle temporal gyrus	21	3.30	60	−12	−12
Cerebellum	–	3.27	−21	−57	−21
Superior parietal lobule	7	3.27	33	−72	51

$P < 0.001$, five-voxel extent threshold. Local maxima within each cluster (>8 mm apart) are indicated by indentation. Foci within white matter are not reported. BA, Brodmann's area.

Table S2. Regions more active for low-evidence AC trials vs. low-evidence AB trials

			MNI coordinates		
Region	BA	Z	*x*	*y*	*z*
Precuneus	31	4.34	−18	−60	24
Precuneus	7	4.30	−12	−63	30
Precuneus	7	3.97	12	−63	30
Posterior cingulate gyrus	31	3.53	9	−48	24
Precuneus	7, 31	3.53	18	−51	30
Posterior cingulate gyrus	31	3.43	−3	−39	27
Middle frontal gyrus	10, 46	4.10	−36	51	9
Middle frontal gyrus	46	3.58	−45	42	18
Insula	13	4.01	39	6	6
Anterior cingulate/medial PFC	32	3.69	6	33	−15
Anterior cingulate/medial PFC	32	3.27	0	39	−9
Precuneus	7	3.56	−6	−60	63
Thalamus	–	3.51	3	−12	0
Fusiform	19	3.46	36	−72	−18
Cingulate gyrus	23	3.38	9	−27	33
Inferior parietal lobule	40	3.35	−45	−54	51
Medial superior frontal gyrus	6, 8	3.32	3	36	42
Cerebellum	–	3.32	21	−78	−18

$P < 0.001$, five-voxel extent threshold. Local maxima within each cluster (>8 mm apart) are indicated by indentation. BA, Brodmann's area.

Table S3. Relationship between DLPFC foci in the present study and a DLPFC region previously implicated in perceptual decision-making

			MNI coordinates		
	Z	P_{SVC}	*x*	*y*	*z*
Specific hits > (general hits/don't knows), DE pairs only	3.06	0.017*	−18	27	39
	3.00	0.020*	−30	21	39
(High- > low-evidence AB) > (high- > low- evidence AC)	2.81	0.036*	−30	27	39
	2.72	0.046*	−27	21	42
High evidence > low evidence, AB/DE pairs only	2.80	0.036*	−24	21	42
	2.62	0.054	−27	27	42

A DLPFC ROI was generated, centered at coordinates previously associated with tracking the strength of perceptual evidence (MNI coordinates: $x = -24$, $y = 24$, $z = 36$; 8-mm radius sphere) (1). This ROI was used for small volume correction (SVC) of three contrasts: (*i*) a contrast of retrieval trials associated with detailed episodic retrieval vs. less-detailed or failed retrieval, (*ii*) a test for an interaction between pair type (AB vs. AC) and strength of classifier evidence (high vs. low), and (*iii*) a contrast of noncompetitive retrieval trials associated with high- vs. low-classifier evidence. Each contrast revealed responses within the DLPFC ROI that were significant following SVC, confirming that the DLPFC foci reported here were anatomically consistent with the DLPFC region previously implicated in perceptual decision-making. For each contrast: $P < 0.005$, five-voxel extent threshold. Small volume correction was performed at the voxel level using a family-wise error rate correction. Asterisk (*) denotes foci that were significant following SVC.

1. Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859–862.