

Fidelity of neural reactivation reveals competition between memories

Brice A. Kuhl^{a,b,1}, Jesse Rissman^a, Marvin M. Chun^b, and Anthony D. Wagner^{a,c}

^aDepartment of Psychology and ^cNeurosciences Program, Stanford University, Stanford, CA 94305; and ^bDepartment of Psychology, Yale University, New Haven, CT 06511

Edited by Edward E. Smith, Columbia University, New York, NY, and approved February 14, 2011 (received for review November 10, 2010)

Remembering an event from the past is often complicated by the fact that our memories are cluttered with similar events. Though competition is a fundamental part of remembering, there is little evidence of how mnemonic competition is neurally represented. Here, we assessed whether competition between visual memories is captured in the relative degree to which target vs. competing memories are reactivated within the ventral occipitotemporal cortex (VOTC). To assess reactivation, we used multivoxel pattern analysis of fMRI data, quantifying the degree to which retrieval events elicited patterns of neural activity that matched those elicited during encoding. Consistent with recent evidence, we found that retrieval of visual memories was associated with robust VOTC reactivation and that the degree of reactivation scaled with behavioral expressions of target memory retrieval. Critically, competitive remembering was associated with more ambiguous patterns of VOTC reactivation, putatively reflecting simultaneous reactivation of target and competing memories. Indeed, the more weakly that target memories were reactivated, the more likely that competing memories were later remembered. Moreover, when VOTC reactivation indicated that conflict between target and competing memories was high, frontoparietal mechanisms were markedly engaged, revealing specific neural mechanisms that tracked competing mnemonic evidence. Together, these findings provide unique evidence that neural reactivation captures competition between individual memories, providing insight into how well target memories are retrieved in the present and how likely competing memories will be remembered in the future.

forgetting | pattern classification

Our ability to remember an event from the past is powerfully influenced by competition arising from memories of similar or overlapping events (1–3). For example, in searching for today's parking space, we may find ourselves standing where we parked yesterday. Though competition between memories is almost ubiquitous, and a primary reason why we forget, there is surprisingly little evidence of how competition between memories is neurally represented. In part, the lack of evidence reflects a methodological challenge of how to measure neural competition between memories. Here, we consider whether competition between memories can be measured by, and understood in terms of, the relative degree to which memories are neurally reactivated—that is, the degree to which patterns of neural activity present during event encoding are reinstated at retrieval. By this view, competitive remembering may strongly parallel competitive perception (4)—a domain that has been more extensively studied.

When competition exists between visual stimuli, responses within the ventral occipitotemporal cortex (VOTC) are strongly modulated by how attention is allocated. For example, when faces and scenes are concurrently or sequentially presented, increased activity is observed in fusiform or parahippocampal gyri according to whether faces or scenes are attended, respectively (5–9). Similarly, VOTC responses are tightly correlated with both spontaneously fluctuating perceptions (10, 11) and misperceptions (12) of visual stimuli. VOTC responses also scale with gradations in the strength of visual stimuli (13–15), thus reflecting the quality of perceptual events. Importantly, perceptual evidence emerging from the VOTC is also correlated with the engagement of fron-

toparietal structures (15–19), putatively reflecting the translation of perceptual evidence to goal-relevant behavior.

Building on the literature relating competitive perception to VOTC responses, we used a memory task in which subjects learned cue-associate pairs for which the cues were words and the associates were images of faces or scenes. Competition between individual memories was created through AB/AC learning: subjects first encoded and retrieved novel cue-associate pairs (noncompetitive, AB pairs) and subsequently encoded and retrieved overlapping pairs that contained a repeated cue paired with a novel associate (competitive, AC pairs; Fig. 1). A separate set of novel cue-associate pairs (noncompetitive, DE pairs) were not followed by overlapping pairs, thus functioning as a control condition. To allow for separation of VOTC reactivation related to target vs. competitor memories, the B and C images for a given AB/AC set were always from distinct visual categories (i.e., a face and a scene). During the critical retrieval phase, subjects were presented with word cues and attempted to covertly recall target associates, indicating by button press whether they were able to recall the specific image vs. a more general memory for the category of the image (face/scene). The scanned encoding and retrieval rounds were followed by a behavioral posttest that reassessed memory for AB and DE pairs, allowing for measurement of the impact that AC learning had on memory for previously learned AB pairs.

Reactivation was assessed via multivoxel pattern analysis (MVPA)—a technique well-suited to measuring reactivation (20–22). Accordingly, we first trained a pattern classifier to discriminate patterns of activity within the VOTC that were associated with the encoding of words paired with faces vs. scenes; we then applied the classifier to retrieval trials to estimate the degree to which VOTC responses reflected face vs. scene reactivation—that is, the fidelity with which target memories were retrieved. Importantly, because the classifier measured the relative evidence for face vs. scene reactivation, here, fidelity of reactivation refers to the relative strength of target reactivation. We predicted that competitive remembering would be associated with reactivation of both target and competitor memories, thereby reducing the fidelity of VOTC reactivation. Moreover, we predicted that measures of competitive reactivation would be related to both how well target memories were retrieved in the present and how likely competing memories would be remembered in the future. Finally, we predicted that neural evidence of competitive reactivation would correspond to the engagement of frontoparietal structures that evaluate mnemonic—and perhaps perceptual—evidence.

Results

Behavioral Results. Retrieval rounds. Responses during the scanned retrieval rounds were separated according to whether subjects

Author contributions: B.A.K., M.M.C., and A.D.W. designed research; B.A.K. performed research; B.A.K. and J.R. analyzed data; and B.A.K., J.R., M.M.C., and A.D.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: brice.kuhl@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1016939108/-DCSupplemental.

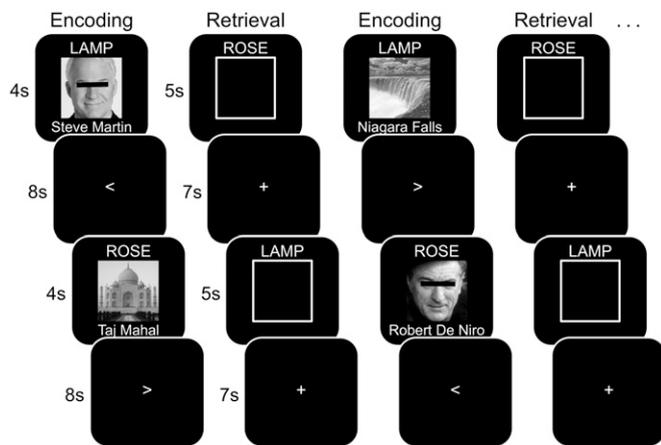


Fig. 1. Encoding and retrieval rounds. During encoding, subjects viewed nouns (cues) paired with images of well-known faces or scenes (associates). The name of each face/scene was presented below the image. During retrieval, subjects were presented with cues and instructed to recall the corresponding associate, indicating the category (face/scene) via button press. Encoding and retrieval rounds alternated for a total of seven rounds each. Half of the cues were paired with one associate (DE pairs), and half of the cues were paired with two associates (one face, one scene). If a cue was paired with two associates, the first associate (AB pair) was encoded and retrieved before encoding and retrieval of the second associate (AC pair). During retrieval, subjects were instructed to recall the most recent associate for each cue. Images of faces were not obscured in the actual experiment.

correctly indicated the category (face/scene) of the target image (hit); indicated that they did not remember the target image (don't know); or incorrectly classified the category of the target image (error). Hits were further subdivided according to whether subjects reported remembering specific or general details of the image. Because errors were relatively infrequent, they were not subdivided into specific and general errors.

Overall, subjects were fairly successful at retrieving target images: specific hit = 53.2% (mean); general hit = 23.5%; don't know = 14.7%; error = 6.4%; no response = 2.1% (Fig. 2). Critically, competition significantly impacted retrieval success, as evidenced by a robust reduction in the rate of specific hits ($t_{17} = 5.10$, $P < 0.001$; Fig. 2). Though the reduction in the rate of specific hits necessarily corresponded to a higher rate of responses in the remaining bins—and thus nonindependence of additional tests—competition increased general hits ($t_{17} = 2.22$, $P < 0.05$), don't knows ($t_{17} = 2.23$, $P < 0.05$), and errors ($t_{17} = 2.38$, $P < 0.05$). The rate of specific hits, general hits, don't knows, and errors did not differ for faces vs. scenes (all P 's > 0.4).

Posttest. Recall performance at posttest probed memory for AB and DE pairs from the scanned encoding/retrieval rounds. Posttest trials were coded as successful (recalled) if subjects retrieved at least some detail of the target image beyond the category. As with

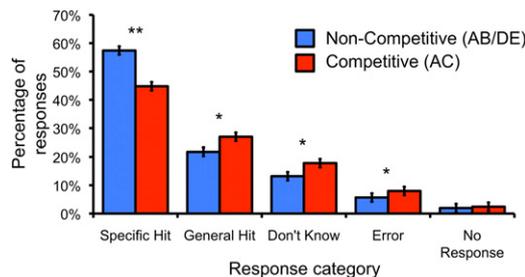


Fig. 2. Behavioral accuracy during retrieval rounds as a function of retrieval competition. Error bars indicate within-subject SEM.

the retrieval rounds, competition negatively impacted performance at posttest, evidenced by a lower rate of recalled AB pairs than DE pairs (mean = 52.4% vs. mean = 57.2%; $t_{17} = 2.32$, $P < 0.05$; *SI Results*).

fMRI Results. Category-selective encoding. Univariate analysis of the encoding data revealed robust category-sensitive neural responses within VOTC (Fig. S1), including greater responses to face trials within fusiform gyrus and greater responses to scene trials in parahippocampal cortex, confirming category-sensitivity within the VOTC. To next permit quantification of the fidelity of cortical reactivation at retrieval, we trained a pattern classifier to discriminate face vs. scene encoding trials using the encoding data from all voxels within an anatomically defined VOTC mask (*SI Methods* and Fig. S2A). MVPA of encoding trials confirmed highly robust sensitivity to face- vs. scene-related encoding responses in the VOTC (*SI Results*).

Cortical reactivation at retrieval. To assess reactivation at retrieval, pattern classification of retrieval trials was performed using the encoding data as the training set and the retrieval data as the testing set. Because the training set was restricted to the encoding data, classification of retrieval data could only succeed to the extent that neural responses that differentiated faces vs. scenes at encoding were reactivated at retrieval.

Considering all retrieval trials, irrespective of the subject's response or condition, trial-by-trial classification accuracy of the target associate's perceptual category (i.e., whether a retrieved item was a face vs. scene) was well above chance (mean = 66.6%, $t_{17} = 7.70$, $P < 0.001$), providing strong evidence that encoding-related activity within VOTC was reactivated at retrieval. This cued-recall result builds on prior evidence of neural reactivation of episodic memories revealed by MVPA, including during free recall (22), source memory (21), and item recognition (20). Notably, classification accuracy scaled with the number of voxels included in the VOTC mask (Fig. S2B), consistent with the idea that the representation of visual stimuli is distributed across the VOTC (23).

We next considered classifier performance in several additional ways. First, by computing classifier performance at each volume [i.e., repetition time (TR)-by-TR classification], we confirmed that classification accuracy generally conformed to the pattern of a hemodynamic response function, with peak accuracy achieved 4–8 s poststimulus onset (Fig. S2C). Second, the distribution of classifier evidence revealed that trials varied considerably in the strength of evidence for target reactivation (Fig. S2D). Finally, this distribution of evidence allowed for generation of receiver operating characteristic (ROC) curves (Fig. S2E); the area under these curves (AUC) provides another useful index of classifier performance (24, 25).

Cortical reactivation and retrieval strength. Having confirmed the approach, our first fundamental objective was to determine whether cortical reactivation scales with retrieval performance. Supporting this possibility, classification accuracy for specific hits, general hits, don't know, and error trials revealed a significant main effect of retrieval performance ($F_{3,51} = 21.70$, $P < 0.001$; Fig. 3A–D). Though classification accuracy for don't know trials did not differ from chance ($t_{17} = 0.27$, $P = 0.79$), classification accuracy for general hit trials was significantly above chance ($t_{17} = 6.48$, $P < 0.001$) and significantly greater than accuracy for don't know trials ($t_{17} = 3.45$, $P < 0.005$). Moreover, as the specificity of retrieval increased, so did the fidelity of cortical reactivation: classification accuracy for specific hit trials was well above chance ($t_{17} = 10.19$, $P < 0.001$) and significantly greater than accuracy for general hit trials ($t_{17} = 2.80$, $P < 0.05$). Finally, classification accuracy for error trials was numerically, but not significantly, below chance ($t_{17} = -1.53$, $P < 0.14$). Measures of AUC mirrored these results (Fig. 3E).

Mnemonic competition and cortical reactivation. Our second fundamental objective was to determine whether and how competition impacts the reactivation of target memories during retrieval. To this end, we first computed mean classification accuracy separately for noncompetitive DE/AB trials and for competition-

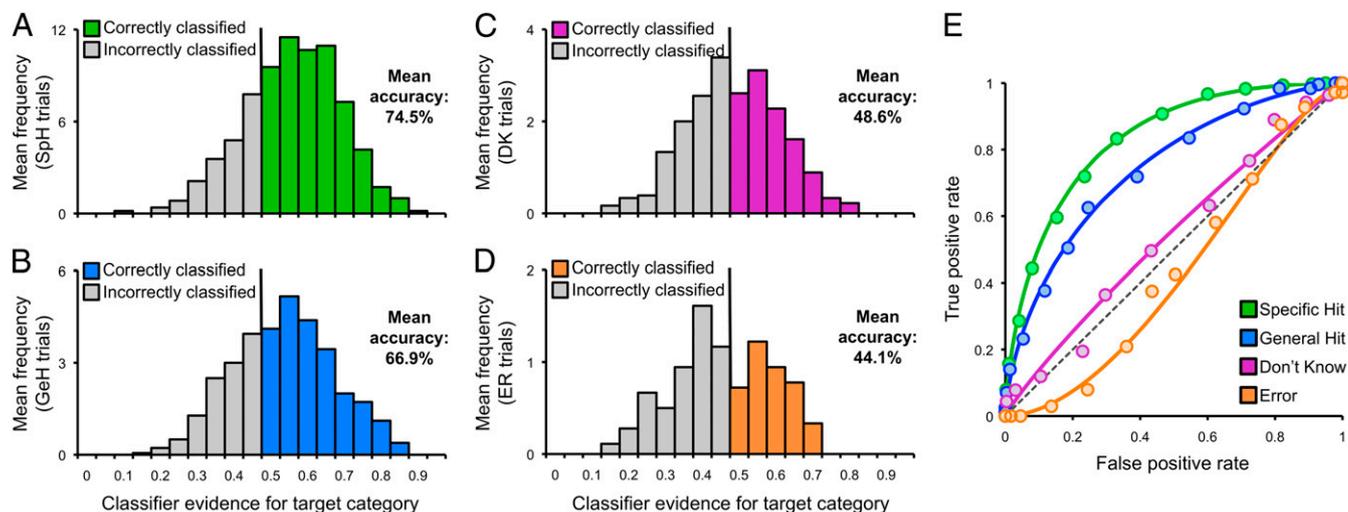


Fig. 3. Classification performance as a function of behavioral response. (A–D) Distribution of classifier evidence for target category and mean classification accuracy for (A) specific hits, (B) general hits, (C) don't knows, and (D) error trials. Individual trials were correctly classified if target evidence exceeded 0.5. (E) ROC curves for each response type; data are based on 14 of the 18 subjects (four subjects excluded due to an insufficient number of trials in at least one of the response bins). AUCs for fitted ROC curves: specific hit = 0.832; general hit = 0.745; don't know = 0.544; error = 0.427.

laden AC trials, irrespective of behavioral performance. Critically, classification accuracy—and thereby evidence for neural reactivation—was lower for competitive retrieval trials than noncompetitive retrieval trials ($t_{17} = 2.39, P < 0.05$), consistent with worse behavioral retrieval success for competitive retrieval trials. Notably, classification accuracy for AB and DE trials did not differ ($t_{17} = 0.49, P = 0.63$), whereas AB vs. AC classification accuracy differed significantly ($t_{17} = 2.61, P < 0.05$; Fig. 4). For subsequent comparisons of competitive vs. noncompetitive classification, we focus exclusively on AB vs. AC trials, as these conditions shared a common retrieval cue (A term).

Though the above data indicate that competition reduced both retrieval success and the fidelity of neural reactivation, it is possible that neural evidence of competition was present even when retrieval performance was matched across AB and AC trials. To this end, we conducted an ANOVA with factors of trial type (AB vs. AC trials) and retrieval specificity (specific vs.

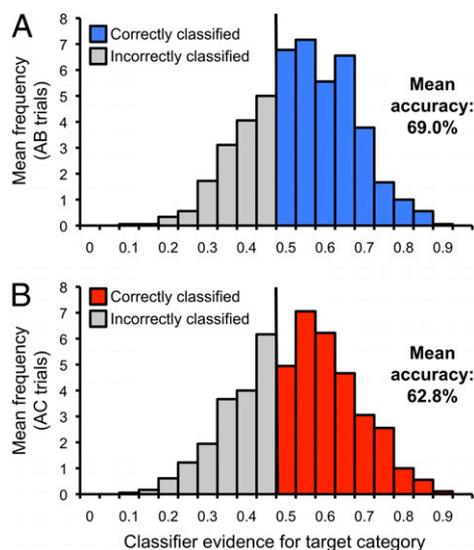


Fig. 4. Classification performance as a function of retrieval competition. Distribution of classifier evidence for target category and mean classification accuracy for (A) AB retrieval trials and (B) AC trials.

general hit trials), allowing for a comparison of the degree of target reactivation as a function of competition while controlling for subjects' responses. Indeed, there was a trend toward poorer classification success for AC than AB retrieval trials ($F_{1,17} = 3.99, P = 0.06$), and no evidence that this effect interacted with retrieval success level (specific vs. general hits; $F < 1$). Together, these results indicate that competitive retrieval events were associated with lower-fidelity reactivation of retrieval targets.

Critically, we next sought to determine whether reductions in the fidelity of target reactivation provided insight into the long-term fate of competing memories. Specifically, we considered whether trial-by-trial measures of classifier evidence for AC reactivation were predictive of posttest memory for AB pairs (i.e., the associations that competed during AC retrieval). To address this, we separated AC retrieval trials according to whether they corresponded to AB pairs that were subsequently remembered vs. forgotten at posttest and according to behavioral measures of AC memory during the scanned retrieval rounds (specific vs. general hits), thus controlling for behavioral evidence of AC learning. This analysis indicated that lower-fidelity reactivation of C terms—or, conversely, greater evidence for reactivation of competing B terms—was associated with a greater likelihood of subsequently remembering the corresponding AB pair at posttest ($F_{1,17} = 8.00, P < 0.05$; Fig. 5). Though the magnitude of this effect was numerically greater for general hit trials, the effect did not interact with AC retrieval performance (specific vs. general hits, $F_{1,17} = 1.10, P = 0.31$). It is important to emphasize that this analysis was based only on AC trials for which subjects correctly identified the category of the C term (i.e., specific and general hit trials), indicating that although there was behavioral evidence that the target category was successfully retrieved, there was nonetheless variance in the fidelity of reactivation that was predictive of subsequent AB memory. Indeed, this relationship was still evident when the analysis was further restricted to only those AC trials that both the subject and the classifier categorized correctly ($F_{1,15} = 5.80, P < 0.05$; two subjects were excluded due to empty cells). Thus, the diminished AC reactivation on trials where subjects went on to later remember the corresponding AB pair cannot be wholly attributed to subjects occasionally reactivating the B associate more strongly than the C associate. Additionally, the relationship did not simply reflect how well AB pairs were initially learned, as the fidelity of initial AB reactivation was not related to the fidelity of AC reactivation (mean r across subjects = $-0.04, P > 0.20$). Together, these data indicate that reduced-fidelity re-

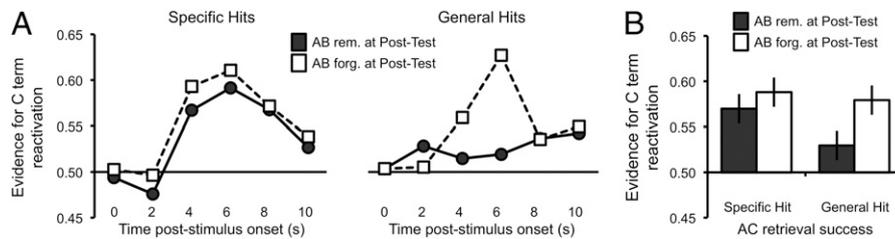


Fig. 5. Relationship between fidelity of reactivation during AC retrieval and subsequent memory for AB pairs. (A) Mean classifier evidence for C term reactivation during AC retrieval as a function of acquisition volume (2 s per volume), behavioral measures of AC retrieval success (specific hit vs. general hit), and subsequent memory for corresponding AB pairs at posttest (remembered vs. forgotten). (B) Mean classifier evidence for C term reactivation during AC retrieval as a function of AC retrieval success (specific hits vs. general hits) and subsequent memory for corresponding AB pairs at posttest [recalled vs. forgotten (don't know or error)]. Error bars indicate within-subject SEM.

membership of new associations corresponded to better subsequent memory for previously encoded, competing associations. **Frontoparietal mechanisms and low-fidelity reactivation.** The preceding analyses indicate that low-fidelity remembering of AC pairs is positively associated with later memory for AB pairs. One interpretation of this relationship is that past associations (B terms) are reactivated during the retrieval of newer associations (AC pairs), and this coactivation of B and C terms yields low-fidelity reactivation (i.e., ambiguous classifier evidence) but ultimately benefits AB retention. By this account, low-fidelity AC trials can be characterized in terms of high, but conflicting, mnemonic evidence. Alternatively, it is possible that when AC pairs are weakly reactivated, memory for AB pairs is less likely to be disrupted, thus accounting for the inverse relationship between AC reactivation and subsequent AB memory. By this account, low-fidelity AC trials can be characterized in terms of low, but not necessarily conflicting, evidence.

To differentiate between these competing accounts, we first identified frontoparietal regions that positively tracked retrieval evidence, and then assessed how responses in these regions related to the fidelity of AB vs. AC reactivation. Specifically, we performed a conjunction analysis of two independent contrasts: (i) a contrast to identify regions that tracked behavioral expressions of mnemonic evidence, and (ii) a contrast to identify regions that were differentially related to the fidelity of reactivation across AB vs. AC trials. To first identify regions that tracked mnemonic evidence, we contrasted specific hit trials against general hit and don't know trials. Importantly, we restricted this contrast to only DE pairs, so as to obtain a contrast that was independent of the critical AB/AC trials (Table S1). To next identify regions that were differentially engaged in relation to the fidelity of reactivation across AB vs. AC retrieval events, we first separated AB and AC trials according to the fidelity of classifier-based evidence for VOTC reactivation. AB trials were sorted into three bins of equal size: low-, medium-, and high-fidelity reactivation, and likewise for AC trials. It should be emphasized that these bins only represented the relative strength of target evidence; strong reactivation of both face and scene representations would correspond to low-fidelity reactivation because of the lack of evidence selectively favoring the target category. Using these binned data, we then tested for a voxel-level interaction between the fidelity of reactivation (high vs. low) and pair type (AB vs. AC), with the prediction being that AC trials associated with low-fidelity reactivation would elicit relatively high engagement of regions that tracked mnemonic evidence.

A conjunction analysis of the two contrasts (each thresholded at $P < 0.005$) revealed overlapping activation in several regions, including the dorsolateral prefrontal cortex (DLPFC), medial prefrontal cortex, and lateral and medial parietal cortex (Fig. 6A and Fig. S3A). Critically, these regions were characterized by relatively high activation during low-fidelity AC retrieval events (Fig. 6B). Indeed, a direct contrast of low-fidelity AC retrieval events vs. low-fidelity AB retrieval events revealed highly similar frontoparietal regions (Table S2). These data are particularly striking considering that, following our trial binning procedure,

low-fidelity AC retrieval events were associated with even lower-fidelity reactivation (mean = 0.41) than low-fidelity AB retrieval events (mean = 0.43; $P < 0.01$). High-fidelity AB (mean = 0.70) vs. AC (mean = 0.70) trials did not differ ($P = 0.73$). Thus, whereas low-fidelity AC retrieval events were associated with weaker classifier-based evidence that target memories were reactivated, they were associated with elevated frontoparietal responses, suggesting that a relatively high amount of episodic information was nonetheless retrieved. Together, these data strongly favor the interpretation that low-fidelity reactivation during AC retrieval reflected robust but nonselective retrieval—that is, reactivation of both target and competitor memories.

Discussion

Emerging evidence indicates that neural reactivation is a fundamental component of event remembering (26). The present results provide unique evidence that reactivation reflects not only how successfully memories are retrieved but also how competition impacts remembering—namely, that competitive remembering of visual memories is associated with lower-fidelity

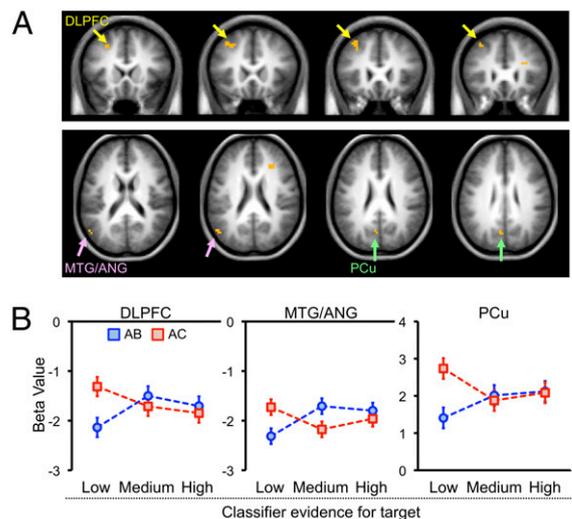


Fig. 6. Frontoparietal regions that tracked mnemonic evidence and fidelity of reactivation. (A) Conjunction analysis: regions that tracked retrieval of specific event details (specific hit trials > general hit/don't know trials; DE pairs only; $P < 0.005$) and displayed an interaction between fidelity of reactivation (low vs. high) and trial type (AB vs. AC; $P < 0.005$). DLPFC, dorsolateral prefrontal cortex; MTG, middle temporal gyrus; ANG, angular gyrus; PCu, precuneus. For complete set of implicated regions, see Fig. S3. (B) Responses in several regions of interest drawn from the conjunction analysis revealed that the interaction between fidelity of reactivation and trial type was characterized by marked activation for AC trials associated with low-fidelity reactivation. Error bars indicate within-subject SEM.

reactivation within VOTC. Critically, these decreases in the fidelity of target reactivation were predictive of subsequent competitor memory. These data provide a striking link between neural expressions of memory competition and the corresponding consequences for future remembering.

The present findings point to a strong parallel between visual perception and visual remembering (4). First, although competition impacted the fidelity of reactivation, VOTC reactivation was clearly modulated by retrieval goals. That is, the majority of competitive retrieval trials were associated with VOTC responses that favored the target category (Fig. 4B). This finding parallels evidence that, when competing visual stimuli are presented, VOTC responses track attended stimuli (5–9). Thus, both during competitive perception and competitive remembering, patterns of responses within VOTC provide insight into the representations that are favored via perceptual or mnemonic selection. Second, we observed a strong relationship between behavioral measures of retrieval specificity—that is, the strength of retrieved information—and the fidelity of VOTC reactivation (Fig. 3). Thus, the degree to which relevant VOTC structures are engaged corresponds to the strength of visual evidence, whether that evidence comes from external inputs that drive current perception (15, 19) or episodic remembering. Together, these findings reveal that, like visual perception, visual remembering is intimately related to the evoked patterns of activation across VOTC.

A primary goal of the present study was to characterize how competition impacts neural reactivation. Does competition reduce the fidelity of VOTC reactivation? If so, are such reductions due to reinstatement of competing memories, and are these reductions predictive of later memory outcomes? Our findings suggest that retrieval cues associated with competing images elicited lower-fidelity reactivation of target categories, relative to retrieval cues associated with a single image. This finding is consistent with the observed behavioral costs associated with competition—namely, a reduction in the rate with which specific event details were retrieved. Interestingly, we also observed a trend for lower-fidelity reactivation during competitive retrieval even when behavioral accuracy was matched, suggesting that VOTC responses may reflect costs that are not otherwise apparent in behavior.

Given that reductions in the fidelity of AC reactivation putatively reflect interference from corresponding AB pairs, it is notable that the fidelity of AC reactivation did not significantly correlate with the fidelity of corresponding AB reactivation. To the extent that AB pairs compete with AC pairs, a negative correlation would have been predicted. Though the data were numerically in this direction (the correlation coefficient was negative in 12 of 18 subjects, and the group mean was negative), we may have lacked sufficient power for this subtle analysis. In particular, AC reactivation should be a product of both the degree to which AC pairs are successfully encoded, which is unaccounted for in this analysis, and the degree to which AB pairs interfere. As this finding is ultimately inconclusive, this issue is worth future consideration. It is of note, however, that we did observe a positive correlation between the fidelity of AB reactivation and prefrontal engagement during corresponding AC retrieval (*SI Results*), consistent with the idea that the strength of AB associations does influence AC retrieval.

Perhaps the most compelling aspect of VOTC responses in the present study is that variance in the fidelity of AC reactivation was predictive of future remembering of competing memories. The lower the evidence for target memory reactivation—and, therefore, the stronger the evidence for competitor reactivation—the more likely that competing memories were later remembered. Critically, this relationship was observed even when only considering trials that both the subject and the classifier categorized correctly. In other words, even when targets were successfully retrieved and VOTC reactivation was biased toward target representations, there was nonetheless meaningful variance in the fidelity of reactivation that reflected the influence of competing memories. These data indicate that retrieval success was graded, and these gradations were diagnostic of future remembering.

Though we primarily focus on this variance in relation to competitive dynamics, there was also variance in the fidelity of reactivation for noncompetitive retrieval trials (Fig. 4A). Indeed, variance in the fidelity of AB/DE reactivation was predictive of subsequent memory for these pairs (Fig. S4). Thus, AB memory at posttest was a function of both how strongly AB pairs were initially reactivated and how weakly AC pairs were subsequently reactivated. Collectively, the present findings provide evidence for a relationship between distributed patterns of neural reactivation and subsequent mnemonic outcomes. As such, these results are highly relevant to a growing literature that considers the powerful ways in which current acts of retrieval can influence future remembering (27–29).

Consideration of frontoparietal responses during retrieval provided strong evidence that competing items were reactivated during target retrieval, thereby reducing the fidelity of classifier-based evidence for reactivation. Specifically, low-fidelity AC retrieval events disproportionately engaged frontoparietal regions that tracked mnemonic evidence (Fig. 6). This relationship raises two important questions. First, why was reactivation of B terms during AC retrieval associated with better subsequent AB memory? On the one hand, our findings are surprising in light of reconsolidation theory, which posits that reactivation renders memories susceptible to disruption (30, 31). On the other hand, the benefits of AB reactivation during AC retrieval are consistent with evidence documenting the powerful benefits of event retrieval for subsequent memory retention (29), as well as with event integration theories, which posit that the reinstatement of older associations during processing of newer associations can lead to the direct binding of past and present in memory (32, 33). Similarly, behavioral evidence indicates that integration can reduce forgetting of otherwise competing associations (34). Indeed, the relationship observed here strongly parallels recent evidence that reactivation of competing memories during encoding can protect competing memories against forgetting (35) (*SI Discussion*). However, seemingly inconsistent with a strong role of integration, behavioral evidence indicated that recall of AB and AC pairs was conditionally independent ($\chi^2 = 0.085$, $df = 1$, $P = 0.77$; *SI Discussion*). A more subtle but theoretically well-articulated possibility is that reactivation may either increase or decrease the likelihood of forgetting depending on how strongly competing memories are reactivated. In other words, memory disruption may be a nonmonotonic function of reactivation strength (36–38). Understanding the situations in which reactivation is beneficial vs. disruptive represents an important avenue for future investigation.

A second question is whether the frontoparietal regions that tracked mnemonic evidence are convergent with those engaged during visual perception. In particular, we observed a relationship between classifier evidence and responses within a subregion of the left DLPFC, extending from the superior frontal sulcus to the middle frontal gyrus. Notably, a highly similar DLPFC subregion has been argued to play a fundamental role in perceptual decision-making (15, 17; cf. ref. 39). For example, when viewing noise-degraded face vs. scene images, DLPFC activation increases as a function of the strength of visual evidence, regardless of category, whereas category-sensitive VOTC responses increase according to the strength of category-preferred visual stimuli (15). Moreover, DLPFC responses directly scale with both the strength of VOTC responses and subjects' behavioral performance (15).

The present data are strikingly consistent with the data implicating DLPFC in perceptual decision-making, as our data similarly reveal that DLPFC activation scaled with both behavioral performance (i.e., DLPFC was modulated by retrieval success) and VOTC responses (i.e., DLPFC activation was related to classifier evidence; *Table S3*). Importantly, follow-up analyses confirmed the anatomical consistency of the present DLPFC foci with the DLPFC region previously implicated in perceptual decision-making (*Table S3*). The strong convergence across these distinct domains suggests a commonality in the operations that DLPFC performs during visual perception and visual remembering—namely, these collective findings indicate that DLPFC is engaged in relation to the fidelity or strength of

responses within VOTC, regardless of whether these responses are driven by current visual input or memory-based reactivation of visual events. These functional interactions between the prefrontal cortex and the VOTC putatively enable perceptual or mnemonic evidence to be translated to goal-relevant behavioral responses (15, 17, 18). Importantly, these interactions are thought to be fundamental to the implementation of cognitive control in service of competitive remembering (*SI Discussion*).

Collectively, the present results constitute unique evidence relating patterns of neural reactivation elicited during competitive remembering to the quality of information retrieved in the present and to memory outcomes experienced in the future. These results parallel findings from studies of visual perception, indicating that VOTC structures are modulated by current mnemonic goals while reflecting the costs associated with mnemonic competition. Importantly, our results also point to overlap in frontoparietal mechanisms that operate upon perceived vs. remembered representations. More broadly, our results reveal that reactivation provides unique insight into understanding competitive dynamics between memories, and is central to both the experience and consequences of episodic remembering.

Methods

Procedure. The experiment was comprised of four phases: encoding, retrieval, a face/scene localizer task (not considered here), and a posttest. All phases except the posttest were conducted during fMRI scanning. Seven encoding rounds and seven retrieval rounds occurred in alternation. During encoding rounds, subjects studied words (cues) paired with either faces or scenes (associates) for 4 s each. Subjects encoded a total of 48 pairs in each of three conditions (AB, AC, DE). AB and DE pairs appeared in encoding rounds 1–6 (eight pairs per condition per round); AC pairs appeared in rounds 2–7 (eight pairs per round). AB pairs and corresponding AC pairs always appeared in encoding rounds n and $n + 1$, respectively. Encoding trials were separated by

an 8-s active baseline period (Fig. 1). During retrieval rounds, subjects' memory was tested for each pair from the immediately preceding encoding round. Subjects were presented with cues and instructed to covertly recall the corresponding associate. Each trial lasted 5 s, and subjects indicated their retrieval success by making one of five responses using a five-key button box: (i) don't know, (ii) face-specific, indicating that they remembered the specific image and that it was a face, (iii) face-general, indicating that they had a nonspecific memory of a face, (iv) scene-specific, and (v) scene-general. Retrieval trials were separated by a 7-s baseline during which a fixation cross was presented. After exiting the scanner, subjects completed the posttest. For each posttest trial (5 s), subjects were presented with a cue along with the instruction to retrieve either the face or the scene that was previously studied with that cue. For cues that had been associated with more than one image, subjects were always prompted to retrieve the first associate (B term). Subjects responded aloud during posttest. Each trial was followed by a 1-s fixation cross. For additional details of each phase, see *SI Methods*.

fMRI Data Analysis. fMRI scanning was conducted at the Lucas Center at Stanford University on a 3.0T GE Signa MRI system (GE Medical Systems). Functional images were obtained using a T2*-weighted 2D gradient echo spiral-in/out pulse sequence; TR = 2 s; echo time (TE) = 30 ms; flip angle = 75°; 30 slices, $3.4 \times 3.4 \times 4$ mm; axial oblique sequential acquisition. Encoding rounds corresponded to seven functional scans (940 volumes total), and likewise for retrieval rounds. Image preprocessing and data analysis were performed using SPM5 (Wellcome Department of Cognitive Neurology, London). Pattern classification analyses were conducted using the Princeton Multi-Voxel Pattern Analysis Toolbox (<http://code.google.com/p/princeton-mvpa-toolbox/>) and custom code implemented in MATLAB (MathWorks).

ACKNOWLEDGMENTS. This work was supported by National Institute of Mental Health Grants 5R01-MH080309 and 5R01-MH076932 (to A.D.W.); R01-EY014193 and P30-EY000785 (to M.M.C.); and EY019624-02 (to B.A.K.).

- Anderson JR (1974) Retrieval of propositional information from long-term memory. *Cogn Psychol* 6:451–474.
- Anderson MC (2003) Rethinking interference theory: Executive control and the mechanisms of forgetting. *J Mem Lang* 49:415–445.
- Mensink G, Raaijmakers JG (1988) A model for interference and forgetting. *Psychol Rev* 95:434–455.
- Anderson MC, Spellman BA (1995) On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychol Rev* 102:68–100.
- Dudukovic NM, Preston AR, Archie JJ, Glover GH, Wagner AD (2011) High-resolution fMRI reveals match enhancement and attentional modulation in the human medial temporal lobe. *J Cogn Neurosci* 23:670–682.
- Gazzaley A, Cooney JW, McEvoy K, Knight RT, D'Esposito M (2005) Top-down enhancement and suppression of the magnitude and speed of neural activity. *J Cogn Neurosci* 17:507–517.
- O'Craven KM, Downing PE, Kanwisher N (1999) fMRI evidence for objects as the units of attentional selection. *Nature* 401:584–587.
- Yi D-J, Kelley TA, Marois R, Chun MM (2006) Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain Res* 1080:53–62.
- Yi D-J, Chun MM (2005) Attentional modulation of learning-related repetition attenuation effects in human parahippocampal cortex. *J Neurosci* 25:3593–3600.
- Lumer ED, Rees G (1999) Covariation of activity in visual and prefrontal cortex associated with subjective visual perception. *Proc Natl Acad Sci USA* 96:1669–1673.
- Tong F, Nakayama K, Vaughan JT, Kanwisher N (1998) Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21:753–759.
- Summerfield C, Egnor T, Mangels J, Hirsch J (2006) Mistaking a house for a face: Neural correlates of misperception in healthy humans. *Cereb Cortex* 16:500–508.
- Bar M, et al. (2001) Cortical mechanisms specific to explicit visual object recognition. *Neuron* 29:529–535.
- Grill-Spector K, Kushnir T, Hendler T, Malach R (2000) The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci* 3:837–843.
- Heekeren HR, Marrett S, Bandettini PA, Ungerleider LG (2004) A general mechanism for perceptual decision-making in the human brain. *Nature* 431:859–862.
- Gold JI, Shadlen MN (2001) Neural computations that underlie decisions about sensory stimuli. *Trends Cogn Sci* 5:10–16.
- Heekeren HR, Marrett S, Ruff DA, Bandettini PA, Ungerleider LG (2006) Involvement of human left dorsolateral prefrontal cortex in perceptual decision making is independent of response modality. *Proc Natl Acad Sci USA* 103:10023–10028.
- Kim JN, Shadlen MN (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci* 2:176–185.
- Ploran EJ, et al. (2007) Evidence accumulation and the moment of recognition: Dissociating perceptual recognition processes using fMRI. *J Neurosci* 27:11912–11924.
- Johnson JD, McDuff SGR, Rugg MD, Norman KA (2009) Recollection, familiarity, and cortical reinstatement: A multivoxel pattern analysis. *Neuron* 63:697–708.
- McDuff SGR, Frankel HC, Norman KA (2009) Multivoxel pattern analysis reveals increased memory targeting and reduced use of retrieved details during single-agenda source monitoring. *J Neurosci* 29:508–516.
- Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310:1963–1966.
- Haxby JV, et al. (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293:2425–2430.
- Bradley AP (1997) The use of area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit* 30:1145–1159.
- Rissman J, Greely HT, Wagner AD (2010) Detecting individual memories through the neural decoding of memory states and past experience. *Proc Natl Acad Sci USA* 107:9849–9854.
- Danker JF, Anderson JR (2010) The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychol Bull* 136:87–102.
- Anderson MC, Bjork RA, Bjork EL (1994) Remembering can cause forgetting: Retrieval dynamics in long-term memory. *J Exp Psychol Learn Mem Cogn* 20:1063–1087.
- Dudukovic NM, Dubrow S, Wagner AD (2009) Attention during memory retrieval enhances future remembering. *Mem Cognit* 37:953–961.
- Roediger HL, Karpicke JD (2006) Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol Sci* 17:249–255.
- Dudai Y (2004) The neurobiology of consolidations, or, how stable is the engram? *Annu Rev Psychol* 55:51–86.
- Nader K (2003) Memory traces unbound. *Trends Neurosci* 26:65–72.
- Hupbach A, Gomez R, Hardt O, Nadel L (2007) Reconsolidation of episodic memories: A subtle reminder triggers integration of new information. *Learn Mem* 14:47–53.
- Shohamy D, Wagner AD (2008) Integrating memories in the human brain: Hippocampal-midbrain encoding of overlapping events. *Neuron* 60:378–389.
- Anderson MC, McCulloch JC (1999) Integration as a general boundary condition on retrieval-induced forgetting. *J Exp Psychol Learn Mem Cogn* 25:608–629.
- 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.
- Newman EL, Norman KA (2010) Moderate excitation leads to weakening of perceptual representations. *Cereb Cortex* 20:2760–2770.
- Norman KA, Newman EL, Detre G (2007) A neural network model of retrieval-induced forgetting. *Psychol Rev* 114:887–953.
- Anderson MC, Levy BJ (2010) On the relationship between interference and inhibition in cognition. *Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork*, ed Benjamin AS (Psychology Press, London), pp 107–132.
- Ho TC, Brown S, Serences JT (2009) Domain general mechanisms of perceptual decision making in human cortex. *J Neurosci* 29:8675–8687.