

Memory Reactivation Predicts Resistance to Retroactive Interference: Evidence from Multivariate Classification and Pattern Similarity Analyses

 Joshua D. Koen and Michael D. Rugg

Center for Vital Longevity and School of Behavioral and Brain Sciences, University of Texas at Dallas, Dallas, Texas 75235

Memory reactivation—the reinstatement of processes and representations engaged when an event is initially experienced—is believed to play an important role in strengthening and updating episodic memory. The present study examines how memory reactivation during a potentially interfering event influences memory for a previously experienced event. Participants underwent fMRI during the encoding phase of an AB/AC interference task in which some words were presented twice in association with two different encoding tasks (AB and AC trials) and other words were presented once (DE trials). The later memory test required retrieval of the encoding tasks associated with each of the study words. Retroactive interference was evident for the AB encoding task and was particularly strong when the AC encoding task was remembered rather than forgotten. We used multivariate classification and pattern similarity analysis (PSA) to measure reactivation of the AB encoding task during AC trials. The results demonstrated that reactivation of generic task information measured with multivariate classification predicted subsequent memory for the AB encoding task regardless of whether interference was strong and weak (trials for which the AC encoding task was remembered or forgotten, respectively). In contrast, reactivation of neural patterns idiosyncratic to a given AB trial measured with PSA only predicted memory when the strength of interference was low. These results suggest that reactivation of features of an initial experience shared across numerous events in the same category, but not features idiosyncratic to a particular event, are important in resisting retroactive interference caused by new learning.

Key words: encoding; episodic memory; forgetting; multivoxel pattern analysis

Significance Statement

Reactivating a previously encoded memory is believed to provide an opportunity to strengthen the memory, but also to return the memory to a labile state, making it susceptible to interference. However, there is debate as to how memory reactivation elicited by a potentially interfering event influences subsequent retrieval of the memory. The findings of the current study indicate that reactivating features idiosyncratic to a particular experience during interference only influences subsequent memory when interference is relatively weak. Critically, reactivation of generic contextual information predicts subsequent source memory when retroactive interference is either strong and weak. The results indicate that reactivation of generic information about a prior episode mitigates forgetting due to retroactive interference.

Introduction

The durability of an episodic memory depends both on how an event is initially encoded (Craik and Lockhart, 1972; Rugg et al., 2008) and on factors that operate after encoding (McGaugh,

2000; Dudai, 2004). Notably, engaging in new learning can cause retroactive interference, reducing the accessibility of existing memories (Postman and Underwood, 1973; Anderson and Neely, 1996; Wixted, 2004). The standard paradigm for investigating retroactive interference is the AB/AC task. The typical finding is that memory is worse for associations (AB trials) when one member of the association is shared in a subsequent association (AC trials) than for associations where a member is not subsequently represented (DE trials). However, not every memory subjected to interference becomes inaccessible, and the factors that contribute to whether a memory will remain accessible after an interfering event are unclear.

Received Nov. 13, 2015; revised Feb. 25, 2016; accepted Feb. 27, 2016.

Author contributions: J.D.K. and M.D.R. designed research; J.D.K. and M.D.R. performed research; J.D.K. and M.D.R. analyzed data; J.D.K. and M.D.R. wrote the paper.

This research was supported by National Institute of Mental Health Grant 1R01MH074528 (M.D.R.) and a Ruth L. Kirschstein National Research Service Award from the National Institute on Aging (Grant 1F32AG049583; J.D.K.).

The authors declare no competing financial interests.

Correspondence should be addressed to Joshua D. Koen, University of Texas at Dallas, Center for Vital Longevity, 1600 Viceroy Drive, Suite 800, Dallas, TX 75235. E-mail: joshua.koen@utdallas.edu.

DOI:10.1523/JNEUROSCI.4099-15.2016

Copyright © 2016 the authors 0270-6474/16/364389-11\$15.00/0

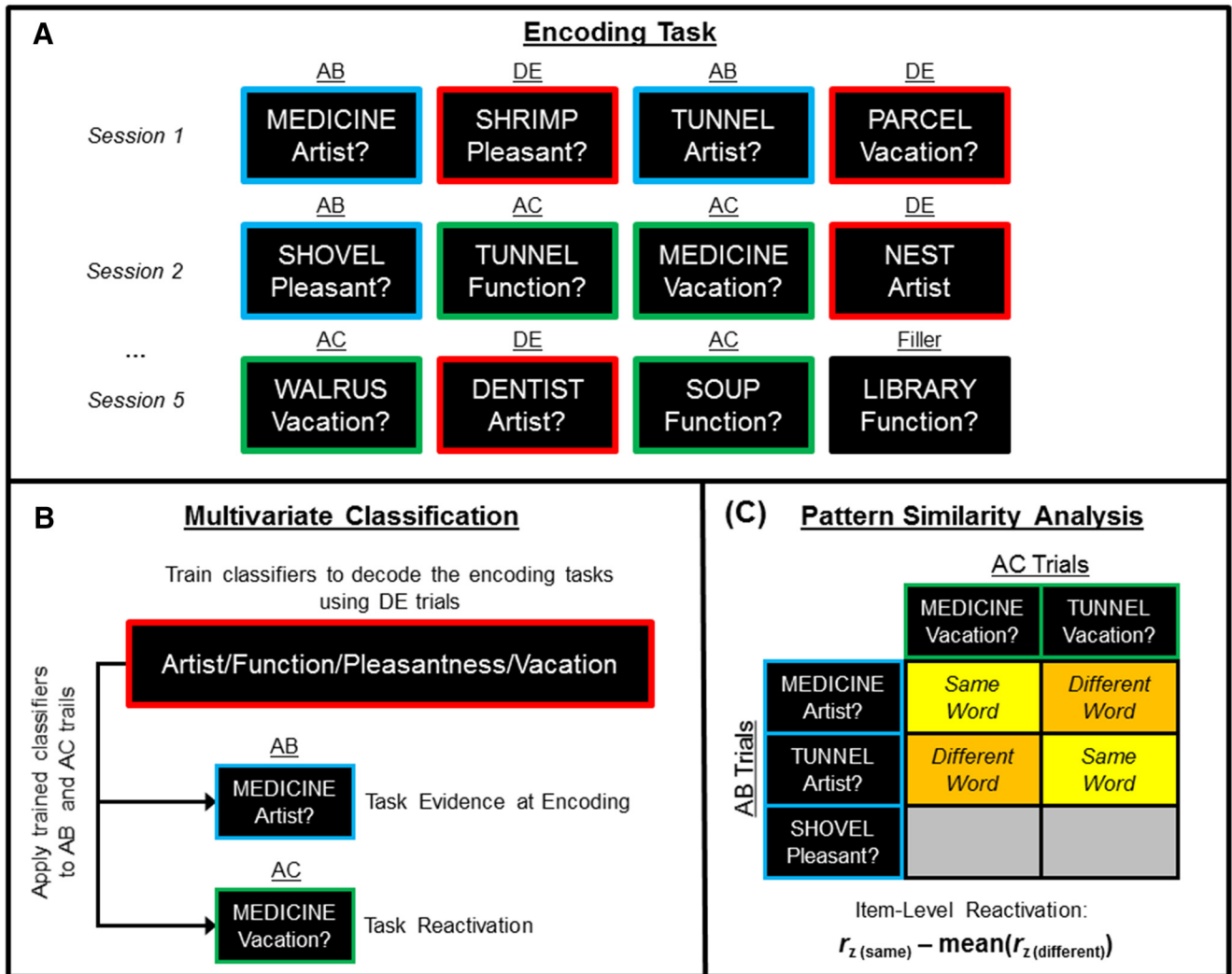


Figure 1. An overview of the source memory AB/AC paradigm and approach to using MVPA to measure memory reactivation in the current experiment. **A**, Participants studied words while performing one of four encoding tasks. Words in the AB (blue boxes) and AC trials (green boxes) were presented twice in two different encoding tasks, whereas words for DE trials (red boxes) and filler trials were presented in a single encoding task. During retrieval, participants first made an old/new decision about the word, and for words receiving an “old” response, they were asked to retrieve all of the encoding tasks performed on the word. **B**, A multivariate classifier was trained to discriminate the four encoding tasks using the DE trials and tested on the AB and AC trials to measure task evidence at encoding (i.e., classifier evidence for the AB encoding task during the AB trial) and task reactivation (i.e., classifier evidence for the AB encoding task during the AC trial sharing the same word), respectively. **C**, Pattern similarity analysis was used to measure item reactivation by correlating the neural patterns elicited by AB and AC trials sharing the same word and subtracting the average correlation between AC and AB trials with different words that shared the same AB encoding task.

Here, we investigate whether memory reactivation—the reinstatement of processes and representations engaged when an event is initially experienced—is predictive of whether a memory persists after a potentially interfering event is encountered. Existing memories can be reactivated when there is sufficient amount of overlap between the original and a new learning episode (McClelland et al., 1995; Norman and O’Reilly, 2003). Memory reactivation during new learning is believed to facilitate generalization across different episodes (Shohamy and Wagner, 2008; Zeithamova and Preston, 2010; Zeithamova et al., 2012a,b; Schlichting et al., 2014; Richter et al., 2015; Schlichting and Preston, 2015) and, more generally, is thought to return the memory to a labile state in which it can be strengthened, weakened, or updated (Sara, 2000; Nader, 2003; McKenzie and Eichenbaum, 2011). However, it is unclear how reactivation during new learning influences subsequent memory for the previously encoded memory. Whereas some evidence indicates that reactivating old memories makes them susceptible to retroactive interference

(Forcato et al., 2007, 2009; Hupbach et al., 2007), other evidence suggests that memory reactivation mitigates the effects of retroactive interference (Kuhl et al., 2010; but see, Zeithamova and Preston, 2010; Richter et al., 2015). We address this issue here, going beyond prior studies by investigating the effects of weak versus strong interference and employing two complementary metrics of reactivation.

The primary aim of the present study was to determine how memory reactivation during new learning predicts subsequent memory for the previously experienced event. Using a source memory version of the AB/AC interference paradigm (Fig. 1A), we measured evidence for the reactivation of existing memories during potentially interfering events using multivoxel pattern analysis (MVPA), a technique well suited to measure memory reactivation (Rissman and Wagner, 2012; Rugg et al., 2015). A second aim of the study was to examine whether different types of memory reactivation, namely, “task-level” and “item-level” reactivation, differentially predict subsequent mem-

ory (Fig. 1). We used multivariate classification (Norman et al., 2006), which generalizes over specific events to measure commonalities in neural patterns between exemplars of the same category, to measure task reactivation (evidence for the AB encoding task during the potentially interfering AC trial). We measured item-level reactivation with pattern similarity analysis (PSA; Kriegeskorte et al., 2008), which can identify neural patterns idiosyncratic to a given experimental item–context association (Ritchey et al., 2013; Wing et al., 2015). It is well established that the strength of the neural pattern elicited by an event predicts subsequent memory for the event (Kuhl et al., 2012; Gordon et al., 2014). Therefore, a third aim of the present study was to determine whether memory reactivation elicited by an interfering event promotes subsequent memory independently of the strength of the neural patterns elicited during encoding. To the extent that reactivating older memories renders them susceptible to interference, lower levels of memory reactivation should be associated with AB trials for which subsequent memory is successful relative to when it is unsuccessful. However, if memory reactivation mitigates the effects of interference, then we expect that stronger reactivation will be predictive of subsequent AB memory.

Materials and Methods

Participants

Twenty-eight adults (16 females) aged between 18 and 30 years (mean, 22) consented to participate in the experiment and were financially compensated \$30 per hour for their time. All were native English speakers, right-handed, and free from neurological and psychiatric disorder according to self-report. The experiment was approved by the Institutional Review Boards of the University of Texas at Dallas and the University of Texas Southwestern Medical Center.

Five participants were excluded, for falling asleep during the scanning session ($n = 2$), technical issues with the MRI scanner ($n = 2$), or withdrawing from the study before completion ($n = 1$). Four additional participants were excluded from the analysis, either because of a high proportion of missed responses in the encoding phase ($n = 2$) that indicated that they were not paying attention to the task or because they had no trials in a condition of interest ($n = 2$). The remaining 19 participants (13 females; mean age, 23) contributed data to the results reported below.

Materials

The critical experimental trials comprised four different encoding tasks (Artist, Function, Pleasantness, Vacation) paired with 300 words taken from the MRC Psycholinguistic Database (Coltheart, 2007). Figure 1 shows example stimuli. The words were all concrete nouns and ranged from four to nine letters in length (mean, 5.36; SD, 1.33), from 1 to 75 in Kucera and Francis (1967) written frequency (mean, 23; SD, 19), and from 500 and 662 in concreteness ratings (mean, 584; SD, 31). For each participant, the 300 critical words were randomly divided into four groups: 96 words served as words for the AB/AC trials, 80 served as words for the DE (control) trials, 100 served as new items during the memory test, and 24 words served as filler items in the fifth and final encoding session. Words assigned to the AB/AC trials were each presented twice during encoding and paired with two different encoding tasks. One task served as the encoding task for the first (AB) presentation (e.g., DOG-Artist?) and the other task served as the encoding task for the second (AC) presentation (e.g., DOG-Vacation?). The AB encoding task was selected from all four tasks. The AC encoding task, however, was selected from the three remaining encoding tasks after excluding the task selected for the AB encoding task. There were an equal number of words in the AB/AC condition assigned to the 12 possible task pairings. The words for the DE and filler trials were paired with a single encoding task selected from the set of all four possible tasks with the constraint that each encoding task occurred equally often for the two classes of trials.

The items in the encoding phase comprised the AB, AC, DE, and filler trials. The words for AB/AC trials were separated into four groups of 24. One group of words was presented in each of the first four encoding

sessions as the AB trials. AC trials only occurred in the final four encoding sessions and comprised the words from the AB trials from the prior encoding session. The presentation of the corresponding AB and AC words always occurred in consecutive sessions, with a minimum lag of 16 trials and a maximum lag of 126 trials between the two presentations. The DE trials were divided into five groups of 16, and one group was presented in each of the five encoding sessions. The 24 filler trials were presented only in the final encoding session. In total, there were 24 AB trials and 16 DE trials in the first encoding session; 24 AB trials, 24 AC trials, and 16 DE trials in the second, third, and fourth encoding sessions, respectively; and 24 AC trials, 16 DE trials, and 24 filler trials in the fifth encoding session. The test trials comprised the words from AB/AC trials, the words from DE trials, and 100 new words.

An additional 26 words with similar characteristics to the critical words were used for practice. Twenty were presented in the practice encoding and test phases, and six served as lures on the practice test phase. All of the trials in the practice encoding phase were identical to the DE trials in the critical phase (i.e., none of the cues were repeated) because we did not want to alert participants to cue repetitions before the critical encoding sessions.

All stimuli were presented using Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent_2000.php) written in Matlab (Mathworks). In the scanner, stimuli were projected to a screen at the rear of the magnet bore and viewed through a mirror mounted on the head coil. Responses during the scanned encoding session were made with a four-button MRI compatible response box. The test phase was completed on a laptop computer outside the scanner. Words were presented in the center of the screen in 32 point Helvetica font.

Procedure

Following informed consent, participants completed the practice encoding phase outside the MRI scanner and then entered the scanner to complete the five critical encoding sessions. Figure 1 depicts the general procedure for the encoding phase. During encoding, words were presented in the center of the screen for 4 s each and followed by a 4–8 s intertrial interval (ITI; 8–12 s stimulus-onset asynchrony; 9 s average). For each trial type, 62.5, 25, and 12.5% of the trials in each phase were followed by a 4, 6, and 8 s ITI, respectively. The presentation order in each encoding session was pseudorandomized such that the different item types (i.e., AB, AC, DE, or filler) did not occur for more than three consecutive trials.

During encoding, participants were cued to make the Artist, Function, Pleasantness, or Vacation judgment when the word appeared. Each judgment was made using a one-to-four scale. Participants were instructed to respond with their right hand while the word was on the screen. For the Artist task (Johnson et al., 2009; McDuff et al., 2009), participants were instructed to rate how difficult it would be for an artist to draw the object denoted by the word (1, easy; 4, hard). When cued for the Function task (Johnson et al., 2009; McDuff et al., 2009), participants were instructed to report the number of functions for the object denoted by the word that came to mind. If four or more functions came to mind, participants were instructed to report four functions. For the Pleasantness task, participants reported how pleasant they found the object denoted by the word (1, unpleasant; 4, pleasant). For the Vacation task (Nairne et al., 2008), participants were asked before the practice encoding phase to think about going on an extended, foreign vacation and were told to use this imagined vacation throughout the experiment. When cued to perform the Vacation task, participants were instructed to report how relevant the object denoted by the word would be to the vacation (1, irrelevant; 4, relevant). The task cues appeared with the onset of the word and were located below the word centered horizontally along with the corresponding response scale. Following the final encoding session, participants remained in the scanner for the structural scan.

Approximately 20 min after the conclusion of the encoding phase, participants completed a surprise memory test in a separate testing room outside the scanner. This phase comprised a practice test phase followed by a critical test phase. Before the practice test, participants were instructed that some words from the encoding phase had been presented twice during the phase, whereas others were presented once, and were

given the instructions for making their memory decisions. Participants were not informed about which words had been presented once or twice. They were informed that they would have to make between one and three memory judgments—item, source, and order—for each word in the test list. For the item judgment, which was made for every test probe, participants were asked to identify words from the encoding list (i.e., words from AB/AC and DE trials) as “old” and words that did not appear during encoding as “new” using the “d” and “f” keys, respectively. If a “new” response was given, the participant moved on to the next trial. However, an “old” response prompted participants to perform the source judgment.

For the source judgment, participants were instructed to attempt to recall all of the tasks that were performed on the word during encoding (similar to the modified modified free recall procedure; cf. Postman and Underwood, 1973). The response options for this memory decision included the labels of the four encoding tasks (i.e., “Artist,” “Function,” “Pleasant,” and “Vacation”) along with a “Don’t Know” response, and participants entered these judgments using the “g,” “h,” “j,” “k,” and “l” keys, respectively. Participants were instructed to give a “Don’t Know” response only if they were unable to remember any of the tasks that were performed on the word during encoding, and were further instructed to select a task only if they were confident in their decision. The instructions for the source memory decision differed slightly if participants believed the word was presented once (which was true for DE trials) or twice (which was true for AB/AC trials). If a participant believed the word was presented twice, the instructions were to try and recall both tasks that were performed on the word during encoding. They were further instructed that if both tasks were confidently remembered, they were to select each task by pressing the corresponding keys on the keyboard. If only one task could be confidently recalled, then participants were instructed to select only the recalled task, even if they were sure the word was presented twice. We assume that participants either forgot or were not confident of the task that was not reported, or that they believed the word was only presented once during encoding. If a participant believed that a word was only presented once during the encoding session, instructions were to report the task that was performed on the word. It was emphasized that selecting two tasks for a word presented once was incorrect, even if the correct task was one of the two tasks selected. DE trials that went on to attract two task responses were excluded from all behavioral and fMRI analyses; the proportion of such responses was low (mean, 0.04; SE, 0.01). For each trial, participants were allowed to select one or two tasks. The selected tasks were highlighted in a blue box and could be toggled on or off by pressing the corresponding key on the keyboard. Given that the “Don’t Know” response was to be selected only when participants were unable to remember any of the encoding tasks, the program did not allow the “Don’t Know” response to be selected in conjunction with a task response. Once the responses were entered, participants were instructed to press the “Enter” key to lock in their response.

If only one task was selected on the source judgment described above, the program moved on to the next trial. However, if two tasks were selected, participants were prompted to perform the order judgment. For this judgment, the two tasks selected in the Source judgment remained on the screen, along with the “Don’t Know” response. Participants were instructed to press the key corresponding to the task that was performed first. Similar to the source judgment, participants were instructed to select a task only if they were confident in their decision. However, if they were not confident or could not remember which task was performed first, they were instructed to give a “Don’t Know” response. All reported analyses were collapsed across the order judgment, and it will not be discussed further. Following completion of the test phase, participants were debriefed and thanked for their time.

Behavioral data measures

The dependent variables of interest from the behavioral data were item recognition performance, source memory accuracy, and reaction time (RT) for the encoding judgments. Item recognition was calculated as the proportion of “old” responses given to AB/AC words, DE words, and new words during the test phase. Source accuracy was calculated as the proportion of accurate source judgments for items accorded an accurate recognition judgment. Source accuracy was calculated in two ways for the

AB and AC encoding tasks. First, source accuracies for the AB and AC encoding tasks were each calculated independently of performance on the other task; that is, AB source accuracy was calculated as the proportion of times the correct AB task was selected when an “old” item response was made, regardless of whether or not the AC task was remembered or forgotten and whether one or two tasks were selected. AC accuracy was calculated in a similar fashion by ignoring AB source accuracy.

Second, we calculated two different conditional source accuracy measures: AB source accuracy conditional on remembering the AC encoding task and AC source accuracy conditional on remembering the AB encoding task. This was achieved by splitting the trials into four cells that crossed AB source memory (hit vs miss) with AC source memory. The AB Hit–AC Hit cell comprised trials in which both the AB and AC encoding tasks were selected. The AB Hit–AC Miss and AB Miss–AC Hit cells comprised trials in which the AB or AC encoding task, respectively, was remembered and the other encoding task was forgotten. The AB Miss–AC Miss cell comprised trials in which both encoding tasks were forgotten. This last bin included trials with two encoding tasks selected that were both incorrect, trials with a single encoding task selected that was incorrect, and trials in which a “Don’t Know” response was selected on the source judgment. The conditional source accuracy measures were computed by holding memory for one task constant [e.g., AB source accuracy when the AC task was remembered was calculated as $AB\ Hit-AC\ Hit / (AB\ Hit-AC\ Hit + AB\ Miss-AC\ Hit)$]. It is important to point out that the same binning procedure described above was used in the subsequent memory analysis of the MVPA measures.

fMRI data acquisition

MRI data were acquired with a 3T Philips Achieva MRI scanner (Philips Medical Systems) equipped with a 32 channel receiver head coil. Functional images were acquired with a blood oxygenation level-dependent (BOLD), T2*-weighted echoplanar imaging (EPI) sequence (SENSE factor, 1.5; flip angle, 70°; 80 × 80 matrix; FOV, 240 × 240 mm; TR, 2000 ms; TE, 30 ms; 34 ascending slices; slice thickness, 3 mm; slice gap, 1 mm) and were oriented parallel to line connecting the anterior and posterior commissures. Three “dummy” scans were acquired at the start of each EPI session and discarded to allow for equilibration of tissue magnetization. A total of 186 functional volumes were acquired during the first encoding session, and 294 functional volumes were acquired in each of the last four encoding sessions. T1-weighted images (MPRAGE sequence; 240 × 240 matrix; 1 mm isotropic voxels) were acquired for anatomical reference following the last EPI session.

fMRI data preprocessing

The functional data were preprocessed with Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology, London, UK) implemented in Matlab 2012b (The Mathworks). First, the functional data were subjected to a two-pass realignment procedure, whereby images were first realigned to the first image of a session and then realigned to a mean EPI image. Next, slice timing differences were corrected using sinc interpolation with reference to the middle slice. The images were then reoriented and spatially normalized to a standard EPI template in MNI space (Cocosco et al., 1997). Finally, the functional data from the five encoding sessions were concatenated.

The structural scan for each participant was spatially normalized to MNI space using an affine transformation. The normalized scans were used to create an across-participant average anatomical image.

Multivoxel pattern analysis

Overview. We used two MVPA approaches, namely, multivariate classification and PSA, to assess how reactivation of the AB association during the presentation of the AC trial sharing the same word was related to subsequent source memory for the AB encoding task. Figure 1 outlines the general approach used to measure task reactivation with multivariate classification, and item reactivation with PSA. In addition, we used multivariate classification analysis to assess the neural patterns elicited during the initial encoding of the AB and AC encoding tasks. These analyses were conducted using a combination of SPM8, the Princeton MVPA Toolbox (<https://github.com/princetonuniversity/princeton-mvpa-toolbox>), and custom Matlab scripts.

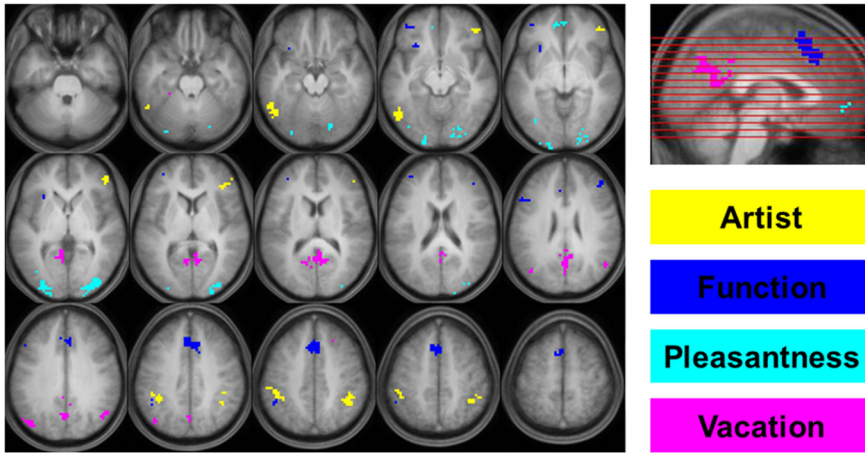


Figure 2. The 1000 voxel feature set used for the multivariate classification and pattern-similarity MPVA. The 250 task-selective voxels for each encoding task are shown in different colors and overlaid on the across-participant average T1-weighted structural scan in MNI space. The axial slices depicted are spaced every 6 mm with the most inferior (top left) and superior (bottom right) corresponding to $z = -27$ and $z = 57$, respectively.

Feature selection. The features used for MVPA were the voxels that showed the largest mean signal differences across the four encoding tasks, as estimated from a univariate GLM analysis. To ensure the independence of the feature selection stage and MVPA of the AB and AC trials, only the DE and filler trials were used to select the features.

The univariate analyses were conducted in two stages. In the first stage, each participant's data were analyzed separately. Each trial was modeled as a 2 s boxcar convolved with a canonical hemodynamic response function, along with its temporal and dispersion derivatives (Friston et al., 1998). The DE and filler trials were sorted according to the encoding task (i.e., Artist, Function, Pleasantness, or Vacation). Trials in which participants did not enter a response to the encoding judgment as well as AB and AC trials were modeled as a single covariate of no interest. Each first-level GLM also included 10 nuisance covariates (six motion parameters and four session-specific means for encoding sessions 2–5). The data were high-pass filtered at 1/128 Hz, and temporal autocorrelation in the error covariance was corrected with an AR(1) model (Friston et al., 2002).

In the second stage, the parameter estimates for the four encoding tasks were carried forward to a one-way ANOVA as implemented in SPM8. Four contrasts were defined, each of which identified voxels that were more active for one of the tasks than for the average of the other three tasks [e.g., Artist > (Function + Pleasantness + Vacation)]. To identify task-selective voxels, each of the contrasts described above was exclusively masked at $p < 0.05$ (one-tailed) with the analogous contrast for the remaining three encoding tasks. For example, to identify task-selective voxels for the Artist task, the contrast Artist > (Function + Pleasantness + Vacation) (contrast weights of [3, -1, -1, -1]) was exclusively masked with (1) Function > (Artist + Pleasantness + Vacation), (2) Pleasantness > (Artist + Function + Vacation), and (3) Vacation > (Artist + Pleasantness + Function). The features used in the classification MVPA and PSA comprised the 250 voxels showing the largest t values in each of the four task-selective contrasts. To ensure that the voxels selected for the MVPA analysis were restricted to gray matter, each of the task selective contrasts described above were inclusively masked with a gray matter mask, which was defined by voxels with a gray matter probability greater than 0.2 in the default SPM gray matter tissue probability map. This resulted in a 1000 voxel feature set that was used for every participant. The voxels that comprised the feature set are depicted in Figure 2, where it can be seen that the four tasks were associated with elevated activity in distinct cortical clusters.

Importantly, the trials used in the feature selection and classifier training steps did not overlap with the trials used in the testing set. The primary analyses focused on the measures obtained from the multivariate classification analyses and PSA from trial types that were separate from those used for feature selection and training. Thus, using the same

trials both to select features and to train the classifier is unlikely to have led to inflation or bias in the measures of interest.

Data preprocessing. The unsmoothed functional data from the 1000 voxels in the feature set were subjected to several preprocessing steps before MVPA. First, linear and quadratic trends were removed from the time series of each voxel and the data z scored across volumes within each of the five encoding sessions. Second, estimates of the BOLD signal for each trial were obtained by averaging across the third and fourth TRs after stimulus onset (i.e., the period encompassing the peak of the evoked hemodynamic response). Trials that did not receive a response during the encoding task or that received a “new” item recognition response on the later memory test were excluded. There were differences in RT between the four encoding tasks (results not reported, but are available from the first author). To remove possible contributions of these RT differences to neural pattern differences, we regressed the BOLD signal in each voxel of the feature set on trialwise RT (cf. Todd et al., 2013). The

z -transformed residuals from these regression analyses were retained and used in the multivariate classification and PSA described below.

Multivariate classification analyses

General overview. The purpose of the multivariate classification analysis was to obtain three experimental measures of interest for the AB and AC trials: AB task evidence at encoding (i.e., evidence for the AB encoding task during an AB trial), AC task evidence at encoding (i.e., evidence for the AC encoding task during an AC trial), and task reactivation (i.e., evidence for the AB encoding task during the AC trial sharing the same word). In the following sections, we outline our approach to training and testing the classifiers and then describe the specific classifiers that were used to obtain the three experimental measures just described.

Approach to classifier training and testing. All classifiers were L2-penalized logistic regression models ($\lambda = 1$). We created classifiers that ensured independence between the trial types used to train the classifiers and the trials the classifiers were applied to. There were two aspects to this. First, classifiers were trained using a different trial type (e.g., DE trials) than the trial types used to obtain the experimental measures of interest (e.g., AB and AC trials). Thus, our measures of task encoding and reactivation avoided possible bias caused by the inclusion of trials in the training set (AB trials) that shared words with trials in the testing set (AC trials).

Second, to avoid introducing classifier bias caused by within-session autocorrelation of the BOLD signal (Mumford et al., 2014), the subset of trials comprising the training set were drawn from different scanning sessions than the trials comprising the test set. This was achieved using a fivefold leave-one-session-out approach to classifier training and testing. In each fold, the DE trials from four encoding sessions (e.g., sessions 1–4) were used as the training set. The trained classifier was subsequently tested on all of the trials from the left out session (e.g., session 5). This process was repeated until all five encoding sessions served as the testing set. The procedure described above allowed for complete independence between the training and testing sets. The approach meant that for each participant, five different classifier sets were required, one for each fold. As will be described below, this resulted in a total of 20 DE classifiers per participant.

Multivariate classification. The primary goal of our analytic approach was to measure task reactivation that was independent of the AC encoding task. To do this, we used four different classifiers that each discriminated between three of the four encoding tasks. The task combinations for the four classifiers were (1) Artist versus Function versus Vacation, (2) Artist versus Function versus Pleasantness, (3) Artist versus Vacation versus Pleasantness, and (4) Function versus Vacation versus Pleasant. As noted above,

this resulted in a total of 20 different trained classifiers for each participant (five folds each with the four classifiers just described).

The trained classifiers were used to obtain three measures of interest: AB task evidence (from AB trials), AB task reactivation (from AC trials), and AC task evidence (from AC trials). For each fold, the trained classifiers were applied to the AB and AC trials of the testing trial set. Only one of the four different classifiers (in each fold) was used to obtain the AB task evidence and reactivation measures, and a different classifier was used to obtain the AC task evidence measure. For both the AB task evidence and reactivation measures, we extracted the evidence values from the classifier that discriminated between the task of interest (the AB encoding task) and the two encoding tasks that were not performed on that word. For example, if, for a given word, the AB and AC tasks were the Artist and Vacation tasks, respectively, then the evidence for the Artist task would be extracted from the classifier with the Artist versus Function versus Pleasantness task combination because this classifier excluded the AC encoding task (i.e., the Vacation task). AC task evidence was measured in a similar manner using the classifier that discriminated the AC encoding task from the two encoding tasks not previously associated with the word (i.e., the classifier that excluded the AB encoding task). Thus, the specific classifiers used to measure AB task evidence and reactivation, and AC task evidence depended on the particular combination and order of the AB and AC encoding tasks for each word. Evidence values were normalized to a range between 0 and 1.

The classifiers were validated by examining classification accuracy of the DE trials in the encoding session labeled as the test set in each fold of the analysis. Classification accuracy was collapsed across the five folds but calculated separately for each encoding task and each of the four classifiers outlined above. This resulted in 12 classifier accuracy measures per participant (three tasks in each of four different classifiers).

Pattern similarity analysis

We used PSA to measure reactivation of task specific item–context associations, and thus to assess how strongly the idiosyncratic neural pattern elicited by the combination of a given word and AB encoding task were present during the AC trial that shared the same word. Figure 1C provides a visual depiction of the procedure. We calculated the Fisher z -transformed Pearson's correlation coefficient between the BOLD patterns for each AC trial with each of the AB trials in the previous encoding session. These correlations were divided into three groups based on the AC trial's word and corresponding AB encoding task: same word and same AB encoding task, different word and same AB encoding task, and different word and different AB encoding task. Note that for the purpose of this analysis, the latter category of correlations is not considered. The correlation between AB and AC trials that shared the same word, hereafter referred to as same word correlations, carries information about the degree to which idiosyncratic patterns associated with an AB trial were reactivated during the AC trial. However, the same word correlations do not provide a pure measure of reactivation of the initial AB item–context associations, and can also be driven by generic task reactivation (cf. Wing et al., 2015). To address this issue, we obtained a measure of generic task reactivation (Ritchey et al., 2013; Wing et al., 2015) by averaging across the correlations between AC and AB trials sharing the same AB encoding task but with different words, hereafter referred to as different word correlations. The item reactivation measure was obtained by subtracting the average of the different word correlations from the same word correlation. This approach removes the fraction of the same word correlation that is driven by generic task reactivation.

There are two important details about this analysis that bear mention. First, as discussed above, the AC trials were only correlated with the AB trials from the previous encoding session, which contained the AB trial sharing the same word as a given AC trial. The goal of this constraint was to mitigate differences in interitem lag between the same and different word correlations. Second, none of the correlations were conducted between two trials from the same scanning session, obviating bias caused by within-session autocorrelation in the BOLD signal (Mumford et al., 2014).

Across-participant analyses

Across-participant analyses of the behavioral and MRI data were conducted with SPSS 21. The Greenhouse–Geisser (Greenhouse and Geisser, 1959) procedure was used to correct the degrees of freedom for non-sphericity in ANOVA and is reflected in the reported degrees of freedom. Results were considered significant at $p < 0.05$.

Results

Behavioral results

Item recognition accuracy

A one-way repeated measures ANOVA revealed a significant difference between the hit rates for studied words from DE trials (mean, 0.76; SE, 0.03) and AB/AC trials (mean, 0.93; SE, 0.02) and the false alarm rate to new words (mean, 0.08; SE, 0.02; $F_{(1.27,22.88)} = 566.38, p < 0.001$). *Post hoc* contrasts indicated that participants were able to discriminate both types of studied words from new words ($t_{(18)}$ values $\geq 19.79, p$ values < 0.001). Moreover, participants were better able to discriminate words presented twice, as indicated by a significantly higher hit rate for words in the AB/AC condition relative to words in the DE condition ($t_{(18)} = 11.03, p < 0.001$).

Source memory accuracy

Retroactive inference was investigated by analyzing source accuracy for the AB, AC, and DE tasks (Fig. 3A). As discussed above in Materials and Methods (Methods), source accuracy for the AB encoding task was calculated independently of source accuracy for the AC encoding task and vice versa. If retroactive interference is present in the current experiment, then source accuracy will be lower for the AB than the DE encoding task. A one-way repeated-measures ANOVA showed a significant main effect of item type ($F_{(1.78,32.07)} = 87.26, p < 0.001$). *Post hoc* paired t tests revealed that source accuracy was highest for the DE task (mean, 0.76; SE, 0.03) relative to both the AB task (mean, 0.52; SE, 0.03; $t_{(18)} = 11.14, p < 0.001$), and the AC task (mean, 0.59; SE, 0.03; $t_{(18)} = 9.42, p < 0.001$). Source accuracy for the AB task was significantly lower than accuracy for the AC task, ($t_{(18)} = 4.18, p < 0.001$). A 3 (item type, DE, AB, AC) \times 4 (encoding task, Artist, Function, Pleasantness, Vacation) ANOVA did not reveal a significant main effect of encoding task ($F_{(2.80,50.40)} = 1.45, p = 0.24$), nor a significant interaction ($F_{(4.20,75.57)} = 0.11, p = 0.98$), which suggests that the encoding task did not moderate the above pattern of results.

It is possible that the length of the ITI that preceded or followed a trial might act to confound subsequent memory performance with our MVPA measures. Specifically, trials preceded or followed by a relatively long ITI might be associated both with better source memory and less noisy (and hence, stronger) MVPA measures. There was, however, no evidence for a main effect of length of preceding ITI on source accuracy ($F_{(1.53,27.57)} = 1.14, p = 0.32$), nor was there an interaction between preceding ITI and item type ($F_{(2.59,46.69)} = 0.53, p = 0.64$, Table 1). The same pattern was also evident for the ITI following a trial (main effect of ITI, $F_{(1.74,31.31)} = 0.70, p = 0.50$; interaction, $F_{(2.56,46.10)} = 1.57, p = 0.19$). Thus, we could find no evidence that source memory covaried with ITI, obviating any potential concerns regarding the possibility of a confound with ITI length.

Last, for the AB/AC trials, we examined source accuracy for each study task conditional on source accuracy for the alternate task (see Materials and Methods, Behavioral data measures). First, we calculated AB source accuracy conditional on whether or not the AC encoding task was remembered or forgotten. AB source accuracy was substantially lower when the AC source was

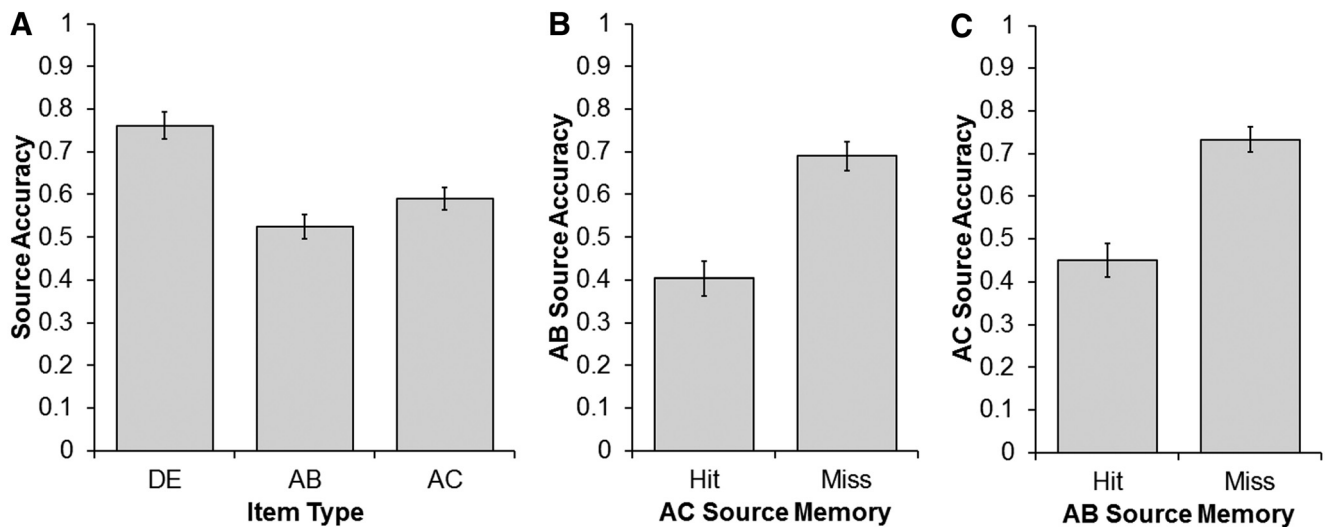


Figure 3. Source memory accuracy calculated from the behavioral data. **A**, Source accuracy for the DE, AB, and AC encoding tasks. Note that in this panel, AB source accuracy was calculated ignoring AC source memory and vice versa. **B**, Source accuracy for the AB encoding task conditional on whether the AC encoding task was remembered (hit) or forgotten (miss). **C**, Source accuracy for the AC encoding task conditional on whether the AB encoding task was remembered or forgotten. Error bars represent ± 1 SEM.

Table 1. Source accuracy as a function of item type and the ITIs of the preceding trial (top) and following a trial (bottom)

	Item type		
	DE	AB	AC
Preceding ITI			
4 s	0.75 (0.03)	0.53 (0.03)	0.60 (0.02)
6 s	0.79 (0.03)	0.51 (0.04)	0.56 (0.04)
8 s	0.76 (0.06)	0.50 (0.05)	0.56 (0.05)
Following ITI			
4 s	0.75 (0.03)	0.53 (0.03)	0.60 (0.03)
6 s	0.82 (0.03)	0.51 (0.04)	0.58 (0.04)
8 s	0.73 (0.04)	0.54 (0.05)	0.58 (0.04)

SEs are provided in parentheses.

remembered (mean, 0.40; SE, 0.04) than when the AC source was forgotten (mean, 0.69; SE, 0.03; $t_{(18)} = 9.23$, $p < 0.001$; Fig. 3B). Second, we calculated AC source accuracy conditional on source memory accuracy for the AB encoding task. Similar to the prior result, AC source accuracy was significantly lower when the AB encoding task was remembered (mean, 0.45; SE, 0.04) than when the AB task was forgotten (mean, 0.73; SE, 0.03; $t_{(18)} = 9.00$, $p < 0.001$; Fig. 3C). Together, these findings suggest that memory for the AB task can be used as a categorical indicator of the strength of interference caused by the AC task and vice versa. For example, AC trials on which the AC encoding task was later remembered can be thought of as exerting higher levels of retroactive interference than trials on which the AC task was forgotten.

MVPA results

First, we present the results from the classifier validation analysis to show that the classifiers were able to discriminate the four encoding tasks above chance levels. Second, we present the analysis of the task evidence at encoding, task reactivation, and item reactivation measures.

Classifier validation

The classifier accuracy measures, which were derived from cross-validation of the DE trials, are provided in Table 2. The across-participant average classification accuracy ranged from 49 to 61% (mean, 55%; SE, 2%) depending on the encoding task and clas-

Table 2. Classification accuracy of the four different classifiers

	Classifier 1	Classifier 2	Classifier 3	Classifier 4
Encoding task				
Artist	0.53 (0.04)	0.52 (0.04)	0.54 (0.04)	
Function	0.59 (0.04)	0.57 (0.04)		0.59 (0.04)
Pleasantness		0.53 (0.03)	0.49 (0.03)	0.54 (0.03)
Vacation	0.61 (0.04)		0.55 (0.04)	0.57 (0.03)

Values in parentheses reflect ± 1 SEM. Chance performance for was 33.33%. Classifier accuracy was collapsed across each fold of the fivefold approach. The tasks discriminated by the classifiers were as follows: Classifier 1, Artist versus Function versus Vacation; Classifier 2, Artist versus Function versus Pleasantness; Classifier 3, Artist versus Pleasantness versus Vacation; Classifier 4, Function versus Pleasantness versus Vacation.

sifier. The four different classifiers, collapsed across the five folds, were able to classify each task reliably above the chance level of 33.3% ($t_{(18)}$ values ≥ 4.43 , p values < 0.001). As described in Materials and Methods, the specific classifier used to measure task evidence at encoding and task reactivation depended on the pairing and order of the AB and AC encoding tasks. Therefore, it was important to determine whether classification accuracy for a particular encoding task varied as a function of the classifier. Only the Vacation task classification varied in this way ($F_{(1,48,26,68)} = 3.85$, $p = 0.05$). *Post hoc t* tests demonstrated that classification accuracy of the Vacation task was greater for the Artist versus Function versus Vacation classifier compared to the Artist versus Pleasantness versus Vacation classifier ($t_{(18)} = 4.21$, $p < 0.001$). However, there were no significant differences in accurately classifying the Vacation task when comparing the Artist versus Function versus Vacation and Function versus Pleasantness versus Vacation classifiers ($t_{(18)} = 1.67$, $p = 0.11$) or the Artist versus Pleasantness versus Vacation and Function versus Pleasantness versus Vacation classifiers ($t_{(18)} = 0.69$, $p = 0.50$). Last, we examined whether classification accuracy varied as a function of the encoding task included in the classifier. A series of four one-way ANOVAs (one for each classifier) did not find significant differences in classification accuracy among the encoding tasks (p values ≥ 0.20). Overall, these results suggest that the four classifiers reliably classified the different encoding tasks and that, with the one exception noted above, classification accuracy did not vary between the classifiers or the encoding tasks.

Subsequent memory analyses

Here, we report the results from the subsequent memory analysis of task reactivation (i.e., classifier evidence for the AB encoding task during AC trials sharing the same word), item reactivation (i.e., pattern similarity between corresponding AB and AC trials), AB task evidence at encoding (i.e., classifier evidence for the AB encoding task during AB trials), and AC task evidence at encoding (i.e., classifier evidence for the AC encoding task during AC trials). Each of the above measures were submitted to separate 2 (AB source memory) \times 2 (AC source memory) repeated measures ANOVAs (for a description of the trials in the four cells, see Materials and Methods, Behavioral data measures). For each participant, we separately averaged each MVPA measure across trials for the four cells of the factorial combination of subsequent AB and AC source memory.

We first report the findings that address the primary question of whether task reactivation, item reactivation, or both were associated with subsequent AB source memory. If memory reactivation during new learning renders the reactivated memory susceptible to retroactive interference, then subsequent AB source misses should be associated with higher levels of task or item reactivation than should subsequent AB source hits. However, if reactivation promotes retention of previously encoded memories, then we expect subsequent AB source hits to be associated with higher levels of reactivation relative to subsequent AB source misses. The ANOVA of the task reactivation measures revealed a significant main effect of AB source memory ($F_{(1,18)} = 5.49$, $p = 0.03$) in the absence of a main effect of AC source memory or an interaction between AB and AC source memory (p values ≥ 0.70). As can be seen in Figure 4A, AB task reactivation during AC trials was higher when the AB encoding task was subsequently remembered (mean, 0.35; SE, 0.01) relative to forgotten (mean, 0.32; SE, 0.01).

The ANOVA on item reactivation measures also revealed a significant main effect of subsequent AB source memory ($F_{(1,18)} = 5.41$, $p = 0.03$), again in the absence of an AC source memory effect ($F_{(1,18)} = 0.05$, $p = 0.83$). There was however also a significant interaction between subsequent AB and AC source memory ($F_{(1,18)} = 5.41$, $p = 0.03$; Fig. 4B). *Post hoc t* tests indicated that when the AC encoding task was forgotten item reactivation was greater for subsequent AB source hits (mean, 0.03; SE, 0.01) than for misses (mean, -0.004 ; SE, 0.01; $t_{(18)} = 2.95$, $p = 0.01$). When the AC encoding task was remembered, however, the difference in item reactivation between AB source hits (mean, 0.01; SE, 0.01) and misses (mean, 0.01; SE, 0.01) was not reliable ($t_{(18)} = 0.39$, $p = 0.70$).

Finally, the 2×2 ANOVA for AB task evidence during AB encoding (Fig. 4C) also revealed a significant main effect of subsequent AB source memory ($F_{(1,18)} = 7.71$, $p = 0.01$), demonstrating higher AB task evidence values when the AB encoding task was subsequently remembered (mean, 0.48; SE, 0.02) compared to when it was forgotten (mean, 0.46; SE, 0.01). There was no evidence for an effect of AC source memory, nor for an interaction between subsequent AB and AC source memory (p values ≥ 0.26). Inspection of Figure 4C strongly suggests however that, despite the absence of a significant interaction effect, task evidence at encoding predicted subsequent memory only when the strength of interference was relatively weak (AC source miss trials). This impression is supported by the finding that the difference in evidence values for the low interference condition was significant ($t_{(18)} = 2.13$, $p = 0.05$), whereas the difference in the high interference condition was far from significant ($t_{(18)} = 0.17$, $p = 0.87$). Although caution is warranted, we tentatively interpret

this pattern of results as evidence that the subsequent memory effect associated with the AB task evidence at encoding measure is modulated by strength of subsequent interference.

The above analyses do not establish that task evidence at encoding, task reactivation, and item reactivation were independently associated with subsequent AB source memory. We addressed this issue in two ways. First, we estimated the relationships across trials between these three MVPA measures for each participant by calculating the across-trial Pearson's correlation coefficients between the measures. The across-participant average of each of these correlations was close to zero (AB task evidence and task reactivation, $r_{\text{avg}} = 0.01$; range, -0.24 to $+0.21$; AB task evidence and item reactivation, $r_{\text{avg}} = 0.003$; range, 0.21 to $+0.29$; task and item reactivation, $r_{\text{avg}} = 0.007$; range, -0.23 to $+0.21$), indicating that there was little shared variance among the three variables at the trial level.

Second, to more directly address the question of independence, the ANOVAs reported previously were repeated using residualized MVPA measures. Thus, for the task reactivation measure, we performed a within-participant, across-trial regression to remove any variance that was shared with measures of AB task evidence at encoding or with item reactivation. Analogous procedures were conducted for the measures of item reactivation (controlling for task evidence and reactivation) and task evidence at encoding (controlling for task and item reactivation). These within-participant regressions were computed across all trials regardless of subsequent memory. The ANOVA results from these residualized MVPA measures were nearly identical to the results of the original analyses reported above. Thus, the ANOVA on the residualized measure of task reactivation revealed only a main effect of subsequent AB source memory ($F_{(1,18)} = 5.20$, $p = 0.04$). The ANOVA on item reactivation again showed a main effect of AB source memory ($F_{(1,18)} = 5.17$, $p = 0.04$) that was moderated by a significant interaction between AB and AC source memory ($F_{(1,18)} = 4.44$, $p = 0.05$). Last, the ANOVA of the residualized measures of task evidence was significantly higher for subsequent AB source hits (mean, 0.47; SE, 0.02) relative to misses (mean, 0.45; SE, 0.02; $F_{(1,18)} = 4.40$, $p = 0.05$).

Finally, we analyzed AC task evidence at encoding with a 2 (AB source memory) \times 2 (AC source memory) repeated measures ANOVA (Fig. 4D). The only reliable effect was a significant interaction between the two factors ($F_{(1,18)} = 7.26$, $p = 0.02$). This interaction was decomposed by contrasting subsequent AC source hits and misses within each level of AB source memory. When the AB encoding task was remembered (strong proactive interference) AC task evidence failed to predict subsequent memory (remembered AC trials, mean, 0.45; SE, 0.02; forgotten AC trials, mean, 0.48; SE, 0.01; $t_{(18)} = 1.84$, $p = 0.08$). However, when interference was weak (the AB encoding task was forgotten) the AC task evidence was significantly higher for trials in which the AC task was subsequently remembered (mean, 0.48; SE, 0.02) rather than forgotten (mean, 0.43; SE, 0.03; $t_{(18)} = 2.54$, $p = 0.02$).

One potential explanation for the interaction just described is that processes involved in reactivating the AB encoding trial (either at the task or item level) compete with processes that are important for encoding the AC task. This scenario predicts that task reactivation, item reactivation, or both should show a negative correlation with the AC task evidence measure. However, the average of the across-trial correlations for AC task evidence with AB task reactivation ($r_{\text{avg}} = 0.04$; range, -0.13 to $+0.22$) and AC task evidence with AB item reactivation ($r_{\text{avg}} = 0.03$; range, -0.17 to $+0.24$) were small and inconsistent with the above prediction.

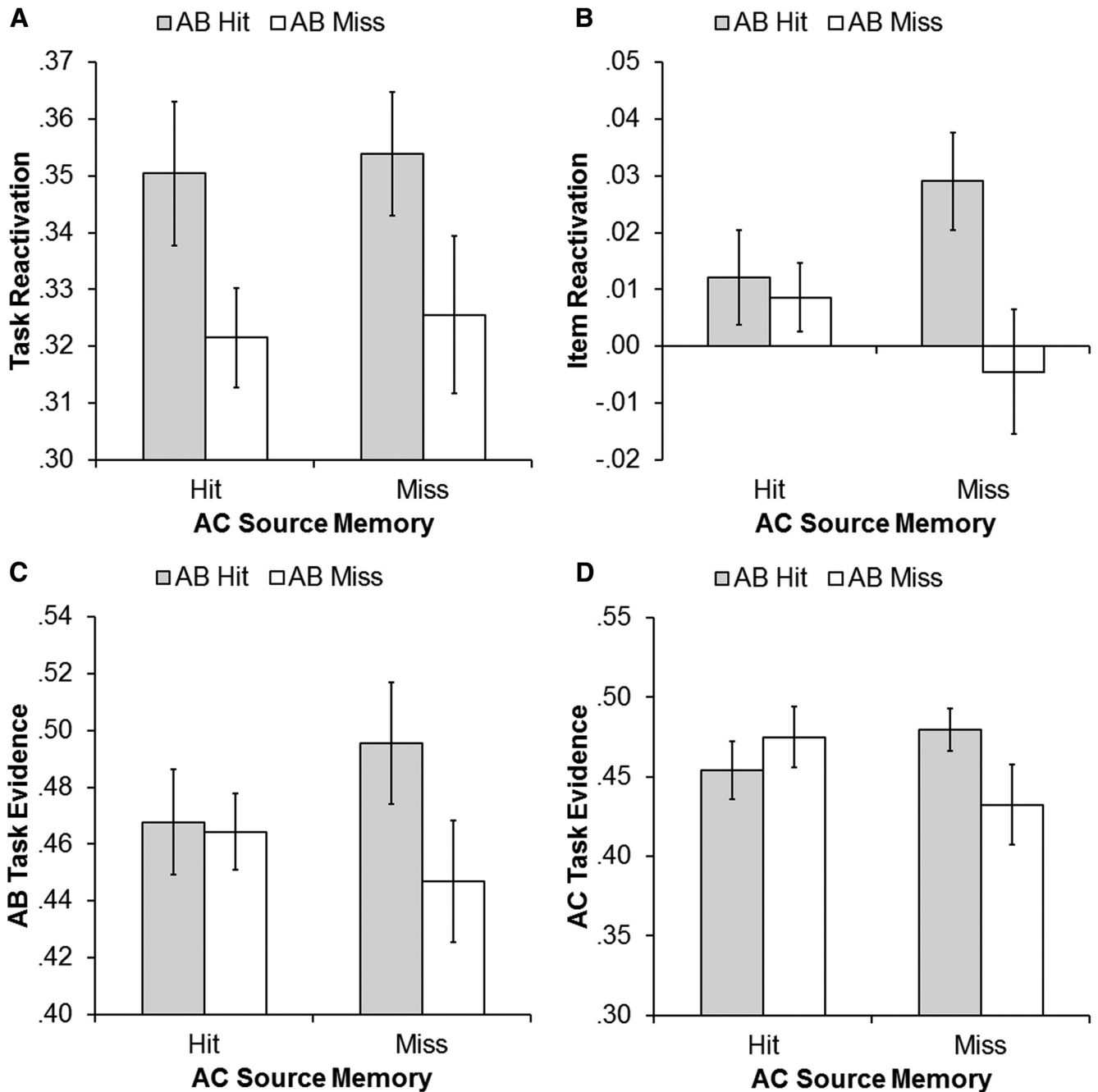


Figure 4. *A–D*, Results from the subsequent source memory analysis of the four MVPA measures: task reactivation (*A*), item reactivation (*B*), AB task evidence at encoding (*C*), and AC task evidence at encoding (*D*). Each panel depicts the across-participant average of the relevant MVPA metric for the four cells formed by treating subsequent source memory for the AB and AC encoding tasks as separate factors. Error bars represent ± 1 SEM.

Discussion

We used MVPA and an AB/AC source memory interference paradigm to investigate how subsequent memory for an initial event is affected by the degree to which it is reactivated when a potentially interfering event is subsequently encountered. In addition, we examined whether task and item-level reactivation have dissociable influences on subsequent memory for AB trials. The results demonstrated that estimates of task-level reactivation were higher when the AB encoding task was subsequently remembered compared to when it was forgotten. The results also showed an association between item-level reactivation, measured with PSA, and subsequent AB source memory. Critically, this latter effect interacted

with interference strength (operationalized by accuracy of AC source memory), such that item-level reactivation only showed a subsequent memory effect when strength of interference was low (when the AC encoding task was forgotten).

Our results are consistent with previous research in suggesting that reactivation of the memory of an event when a highly similar event is later encountered acts to reduce the likelihood it will be forgotten (Kuhl et al., 2010; but see, Zeithamova and Preston, 2010; Richter et al., 2015). Critically, the present results extend these prior findings by showing that two different types of memory reactivation—here termed task and item reactivation—can occur concurrently when an overlapping event is encountered

and are independently predictive of the accuracy of subsequent memory for the original event.

Multivariate classifiers generalize across individual trials and hence do not capture the idiosyncrasies of a particular item–context association (here, the study word and corresponding encoding task). Thus, the task reactivation metric measures commonalities between neural patterns belonging to the same task category. The metric potentially reflects the level of reactivation of a generic cognitive or task set elicited when an event was initially experienced. By contrast, as used here, PSA measured the reactivation of neural patterns that were specific to a given AB study trial (Wing et al., 2015). Our measure of item reactivation was therefore able to capture the degree to which neural patterns elicited by a specific AB trial were reactivated during the presentation of an AC trial sharing the same word. We assume that this metric indexed the reactivation of trial-specific features of an AB study event.

Our results are consistent with the idea that task reactivation and item reactivation measure different aspects of memory reactivation. First, there was little or no shared variance across-trials between these two measures. Second, and equally important, the association between task reactivation and subsequent source memory for the AB encoding task differed from the subsequent memory effects observed for item reactivation. Whereas task reactivation predicted subsequent AB source memory regardless of interference strength, item reactivation only showed a subsequent memory effect on trials associated with low interference strength. Furthermore, this pattern of results was unchanged after controlling for shared variance among the reactivation measures (as well as AB task evidence at encoding).

The results for the task reactivation measure extend the findings reported by Kuhl et al. (2010) and Richter et al. (2015), and suggest that reactivating the context in which an item was initially encountered helps memory for that context to persist in the face of later interference. Importantly, the data suggest that task reactivation predicted subsequent source memory for the AB encoding task even when the strength of retroactive interference—operationalized as AC source memory—was high. Another interesting aspect of this data is that task reactivation promoted subsequent source memory for the item–context association even though this reactivation metric is insensitive to the idiosyncratic features of the particular AB trial. This finding is reminiscent of prior results demonstrating that the reactivation of a study task context facilitates recall of items that were associated with the context (Polyn et al., 2005). It is currently unclear, however, why reactivation of the initial AB encoding task during a potentially interfering event improved memory for the item–context association. One possibility is that reactivating the AB encoding task provided an opportunity to reencode and thus to strengthen the association between the study word and its original encoding task. This hypothesis is consistent with the finding that task reactivation was higher for trials in which the AB encoding task was subsequently recollected, regardless of the strength of the interference caused by the AC trials.

The results from the conditional source memory analysis (Fig. 3) suggest there was competition between memory for the AB and AC encoding tasks. Specifically, source accuracy for one task (e.g., the AB task) was much worse when the other task (e.g., the AC task) was remembered relative to when it was forgotten. The MVPA results suggest that during periods of strong interference (whether retroactive or proactive), competition between memory for the two encoding tasks observed in the behavioral data might be caused by competition between processes supporting the reactivation of item-specific features of a prior event and processes important for the encoding a new, potentially interfer-

ing event. This hypothesis is motivated by the finding that both the item reactivation and AC task evidence measures showed a similar sensitivity to interference strength. Specifically, both measures predicted subsequent memory only when interference was relatively weak. However, as noted in Results, if the processes supporting AB reactivation and the instantiation of the AC task context were in competition, one might have expected the respective reactivation metrics to show a negative correlation. There was, however, no evidence for this. Clearly, further research is needed to elucidate the relationship between the processes supporting memory reactivation and encoding during new learning.

Task evidence during the initial encoding of both AB and AC trials was higher for subsequent source hits than for source misses, as has been reported previously (Kuhl et al., 2012; Gordon et al., 2014). However, task evidence at encoding appears not to be an invariable predictor of subsequent source memory. Although the interaction between AB and AC source memory for the AB task evidence at encoding measure was not reliable at our selected statistical threshold, the general pattern of results indicate that our measures of task evidence at encoding predicted subsequent source memory only when interference was relatively weak. An interesting implication of these results, elaborated on below, is that task evidence at encoding is, at best, only a partial index of the “strength” of encoding. Otherwise, task evidence might have been expected to predict subsequent memory regardless of interference strength.

Relatedly, it was perhaps surprising that there was little or no across-trial relationship between AB task evidence at encoding and either the task or the item reactivation measures. Indeed, these three classes of subsequent memory effect were independent of each other, even when retroactive interference was low. This result suggests that our measure of task evidence at encoding does not reflect processes that determine whether a memory is sufficiently well encoded to be capable of later reactivation. In the limit, of course, any measure of memory reactivation must be related to some metric of encoding success, since a complete failure to encode a memory would mean that there is nothing to later reactivate. This raises the question as to the nature of the encoding processes that were reflected by our task evidence metric. A speculative possibility is that the metric indexed the amount of attention given to the features of the word that were relevant to the specific encoding task (Uncapher and Rugg, 2009). In other words, multivariate classifiers might be sensitive not to the overall quality or fidelity of an encoded memory representation, but to factors, such as attention to specific details of the event, that are only partially correlated with how well it was encoded. An important caveat to this hypothesis arises, however, from the extensive literature demonstrating that the nature of the processing supporting effective encoding depends on how memory is subsequently tested (Tulving and Thomson, 1973; Morris et al., 1977; Otten, 2007; Park and Rugg, 2008). Therefore, the null correlation between task evidence during AB encoding and AB reactivation may reflect a mismatch between the encoding processes reflected by the task evidence metric and the processes supporting AB reactivation during the subsequent AC trial. It is possible that our encoding and reactivation measures might correlate if reactivation is measured in a different retrieval context, such as a direct memory test. Future work will be needed to fully understand the conditions in which multivariate measures of encoding and reactivation correlate with one another, as well as which of the numerous factors operating at encoding are captured in the evidence values of multivariate classifiers.

In conclusion, our results indicate that experiencing new, overlapping events elicits variable amounts of reactivation of pre-

viously encoded events. The fidelities of the reactivation of the memory of a study event at both generic (task) and item-specific levels independently predict subsequent memory. Critically, our findings suggest that the reactivation of the generic context in which an item was initially encountered, but not the reactivation of the idiosyncratic features of the item–context association, can prevent forgetting in the face of strong retroactive interference.

References

- Anderson MC, Neely JH (1996) Interference and inhibition in memory retrieval. In: *Memory* (Bjork EL, Bjork RA, eds), pp 237–313. San Diego: Academic.
- Cocosco CA, Kollokian V, Kwan RK, Evans AC (1997) BrainWeb: online interface to a 3D MRI simulated brain database. *Neuroimage* 5:S425.
- Coltheart M (2007) The MRC psycholinguistic database. *Q J Exp Psychol A* 33:497–505. [CrossRef](#)
- Craik FIM, Lockhart RS (1972) Levels of processing: a framework for memory research. *J Verbal Learn Verbal Behav* 11:671–684. [CrossRef](#)
- Dudai Y (2004) The neurobiology of consolidations, or, how stable is the engram? *Annu Rev Psychol* 55:51–86. [CrossRef](#) [Medline](#)
- Forcato C, Burgos VL, Argibay PF, Molina VA, Pedreira ME, Maldonado H (2007) Reconsolidation of declarative memory in humans. *Learn Mem* 14:295–303. [CrossRef](#) [Medline](#)
- Forcato C, Argibay PF, Pedreira ME, Maldonado H (2009) Human reconsolidation does not always occur when a memory is retrieved: The relevance of the reminder structure. *Neurobiol Learn Mem* 91:50–57. [CrossRef](#) [Medline](#)
- Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998) Event-related fMRI: characterizing differential responses. *Neuroimage* 7:30–40. [CrossRef](#) [Medline](#)
- Friston KJ, Penny W, Phillips C, Kiebel S, Hinton G, Ashburner J (2002) Classical and Bayesian inference in neuroimaging: theory. *Neuroimage* 16:465–483. [CrossRef](#) [Medline](#)
- Gordon AM, Rissman J, Kiani R, Wagner AD (2014) Cortical reinstatement mediates the relationship between content-specific encoding activity and subsequent recollection decisions. *Cereb Cortex* 24:3350–3364. [CrossRef](#) [Medline](#)
- Greenhouse SW, Geisser S (1959) On methods in the analysis of profile data. *Psychometrika* 24:95–112. [CrossRef](#)
- 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. [CrossRef](#) [Medline](#)
- Johnson JD, McDuff SG, Rugg MD, Norman KA (2009) Recollection, familiarity, and cortical reinstatement: a multivoxel pattern analysis. *Neuron* 63:697–708. [CrossRef](#) [Medline](#)
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis—connecting the branches of systems neuroscience. *Front Syst Neurosci* 2:4. [CrossRef](#) [Medline](#)
- Kucera H, Francis WN (1967) *Computational analysis of present-day American English*. Providence, RI: Brown UP.
- 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. [CrossRef](#) [Medline](#)
- Kuhl BA, Rissman J, Wagner AD (2012) Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia* 50:458–469. [CrossRef](#) [Medline](#)
- McClelland JL, McNaughton BL, O'Reilly RC (1995) Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol Rev* 102:419–457. [CrossRef](#) [Medline](#)
- McDuff SG, 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. [CrossRef](#) [Medline](#)
- McGaugh JL (2000) Memory—a century of consolidation. *Science* 287:248–251. [CrossRef](#) [Medline](#)
- McKenzie S, Eichenbaum H (2011) Consolidation and reconsolidation: two lives of memories? *Neuron* 71:224–233. [CrossRef](#) [Medline](#)
- Morris CD, Bransford JD, Franks JJ (1977) Levels of processing versus transfer appropriate processing. *J Verbal Learn Verbal Behav* 16:519–533. [CrossRef](#)
- Mumford JA, Davis T, Poldrack RA (2014) The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *Neuroimage* 103:130–138. [CrossRef](#) [Medline](#)
- Nader K (2003) Memory traces unbound. *Trends Neurosci* 26:65–72. [CrossRef](#) [Medline](#)
- Nairne JS, Pandeirada JN, Thompson SR (2008) Adaptive memory: the comparative value of survival processing. *Psychol Sci* 19:176–180. [CrossRef](#) [Medline](#)
- Norman KA, O'Reilly RC (2003) Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychol Rev* 110:611–646. [CrossRef](#) [Medline](#)
- 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. [CrossRef](#) [Medline](#)
- Otten LJ (2007) Fragments of a larger whole: retrieval cues constrain observed neural correlates of memory encoding. *Cereb Cortex* 17:2030–2038. [CrossRef](#) [Medline](#)
- Park H, Rugg MD (2008) The relationship between study processing and the effects of cue congruency at retrieval: fMRI support for transfer appropriate processing. *Cereb Cortex* 18:868–875. [CrossRef](#) [Medline](#)
- Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes retrieval during memory search. *Science* 310:1963–1966. [CrossRef](#) [Medline](#)
- Postman L, Underwood BJ (1973) Critical issues in interference theory. *Mem Cognit* 1:19–40. [CrossRef](#) [Medline](#)
- Richter FR, Chanales AJH, Kuhl BA (2015) Predicting the integration of overlapping memories by decoding mnemonic processing states during learning. *Neuroimage* 124:323–325. [CrossRef](#)
- Rissman J, Wagner AD (2012) Distributed representations in memory: insights from functional brain imaging. *Annu Rev Psychol* 63:101–128. [CrossRef](#) [Medline](#)
- Ritchey M, Wing EA, LaBar KS, Cabeza R (2013) Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cereb Cortex* 23:2818–2828. [Medline](#)
- Rugg MD, Johnson JD, Park H, Uncapher MR (2008) Encoding–retrieval overlap in human episodic memory: a functional neuroimaging perspective. *Prog Brain Res* 169:339–352. [CrossRef](#) [Medline](#)
- Rugg MD, Johnson JD, Uncapher MR (2015) Encoding and retrieval in episodic memory: insights from fMRI. In: *Handbook on the cognitive neuroscience of memory* (Duarte A, Barense MD, Addis DR, eds), pp 84–107. Oxford, UK: Wiley-Blackwell.
- Sara SJ (2000) Retrieval and reconsolidation: toward a neurobiology of remembering. *Learn Mem* 7:73–84. [CrossRef](#) [Medline](#)
- Schlichting ML, Preston AR (2015) Memory integration: neural mechanisms and implications for behavior. *Current Opin Behav Sci* 1:1–8. [CrossRef](#)
- Schlichting ML, Zeithamova D, Preston AR (2014) CA1 subfield contributions to memory integration and inference. *Hippocampus* 24:1248–1260. [CrossRef](#) [Medline](#)
- Shohamy D, Wagner AD (2008) Integrating memories in the human brain: hippocampal-midbrain encoding of overlapping events. *Neuron* 60:378–389. [CrossRef](#) [Medline](#)
- Todd MT, Nystrom LE, Cohen JD (2013) Confounds in multivariate pattern analysis: theory and rule representation case study. *Neuroimage* 77:157–165. [CrossRef](#) [Medline](#)
- Tulving E, Thomson DM (1973) Encoding specificity and retrieval processes in episodic memory. *Psychol Rev* 80:352–373. [CrossRef](#)
- Uncapher MR, Rugg MD (2009) Selecting for memory? The influence of selective attention on the mnemonic binding of contextual information. *J Neurosci* 29:8270–8279. [CrossRef](#) [Medline](#)
- Wing EA, Ritchey M, Cabeza R (2015) Reinstatement of individual past events revealed by the similarity of distributed activation patterns during encoding and retrieval. *J Cogn Neurosci* 27:679–691. [CrossRef](#) [Medline](#)
- Wixted JT (2004) The psychology and neuroscience of forgetting. *Annu Rev Psychol* 55:235–269. [CrossRef](#) [Medline](#)
- Zeithamova D, Preston AR (2010) Flexible memories: differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *J Neurosci* 30:14676–14684. [CrossRef](#) [Medline](#)
- Zeithamova D, Dominick AL, Preston AR (2012a) Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron* 75:168–179. [CrossRef](#) [Medline](#)
- Zeithamova D, Schlichting ML, Preston AR (2012b) The hippocampus and inferential reasoning: building memories to navigate future decisions. *Front Hum Neurosci* 6:70. [Medline](#)