Behavioral/Cognitive

Neural Context Reinstatement Predicts Memory Misattribution

Samuel J. Gershman,[1,2] Anna C. Schapiro,[1,2] Almut Hupbach,[3]* and Kenneth A. Norman[1,2]*

[1]Department of Psychology and [2] Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey 08540; and [3]Department of Psychology, Lehigh University, Bethlehem, Pennsylvania, 18015

What causes new information to be mistakenly attributed to an old experience? Some theories predict that reinstating the context of a prior experience allows new information to be bound to that context, leading to source memory confusion. To examine this prediction, we had human participants study two lists of items (visual objects) on separate days while undergoing functional magnetic resonance imaging. List 1 items were accompanied by a stream of scene images during the intertrial interval, but list 2 items were not. As in prior work by Hupbach et al. (2009), we observed an asymmetric pattern of misattributions on a subsequent source memory test: participants showed a strong tendency to misattribute list 2 items to list 1 but not vice versa. We hypothesized that these memory errors were due to participants reinstating the list 1 context during list 2. To test this hypothesis, we used a pattern classifier to measure scene-related neural activity during list 2 study. Because scenes were visually present during list 1 but not list 2, scene-related activity during list 2 study can be used as a time-varying neural indicator of how much participants were reinstating the list 1 context during list 2 study. In keeping with our hypothesis, we found that prestimulus scene activation during the study of list 2 items was significantly higher for items subsequently misattributed to list 1 than for items subsequently correctly attributed to list 2. We conclude by discussing how these findings relate to theories of memory reconsolidation.

Introduction

Memory retrieval is a powerful learning event, providing an opportunity to strengthen a memory (Roediger and Butler, 2011) or update it with new information (Lee, 2008). This malleability can also be the cause of error: If new information is bound to older memories, new information may be later mistaken for old information. Striking examples of such errors have been demonstrated in recent experiments by Hupbach et al. (2007, 2008, 2009, 2011). In these experiments, participants studied two lists of items (each study session was separated by 48 h). Shortly before studying list 2 (L2), some participants were reminded of the list 1 (L1) learning episode. When asked to remember L1 items 48 h later, participants who received the reminder showed an asymmetric pattern of memory errors, misattributing a sizeable number of L2 items to L1 but not vice versa; misattribution errors were low for both lists in the no-reminder condition.

Sederberg et al. (2011) offered a theoretical explanation for the asymmetric pattern of intrusions in the reminder condition based on the Temporal Context Model (TCM; Howard and Kahana, 2002). According to TCM, items are bound in memory to the mental context that is active when the item is presented; mental context is operationalized as a running average of recently experienced items (Howard and Kahana, 2002). The effect of the reminder treatment before L2 study is to reinstate the context associated with L1 items. As a consequence, L2 items are bound to the L1 context. This binding results in participants misattributing the L2 items to L1 at test. The pattern of misattributions is asymmetric because, at study, L2 items are linked to the (reinstated) L1 context, but L1 items are not linked to the L2 context.

We sought to more directly test the context reinstatement idea by obtaining a neural measure of L1 context reinstatement. Participants studied two lists of items (separated by 48 h), followed by a source memory test 48 h later in which participants were asked to classify items as being from L1, L2, or neither list. All three sessions were conducted while participants were scanned using functional magnetic resonance imaging (fMRI). During L1 study, we presented a rapid sequence of scene images between each studied item; according to TCM, these scenes should be incorporated into the L1 context. No scenes were presented during L2 study. The scenes presented during L1 study played a role analogous to radioisotope tracers used in positron emission tomography: Scene-related processing is highly visible in fMRI (Epstein and Kanwisher, 1998); consequently, "injecting" scene-related processing into the L1 context (but not the L2 context) made it possible for us to track the emergence of the L1 context over the course of subsequent study and test episodes. In keeping with TCM, we hypothesized that scene activity during L2 study (indicating L1 contextual reinstatement) would predict which L2

Received Jan. 8, 2013; revised Feb. 24, 2013; accepted March 27, 2013.

Author contributions: S.J.G., A.C.S., A.H., and K.A.N. designed research; S.J.G. and A.C.S. performed research; S.J.G. contributed unpublished reagents/analytic tools; S.J.G. analyzed data; S.J.G., A.C.S., A.H., and K.A.N. wrote the paper.

This research was supported by National Institutes of Health R01 MH069456 (K.A.N.). We thank N. Turk-Browne for assistance with stimuli and J. Lewis-Peacock for helpful discussions.

*A.H. and K.A.N. contributed equally to this work.

The authors declare no competing financial interest.

Correspondence should be addressed to Samuel J. Gershman, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139. E-mail: sjgershm@mit.edu.

DOI:10.1523/JNEUROSCI.0096-13.2013

 0270-6474/13/338590-06$15.00/0

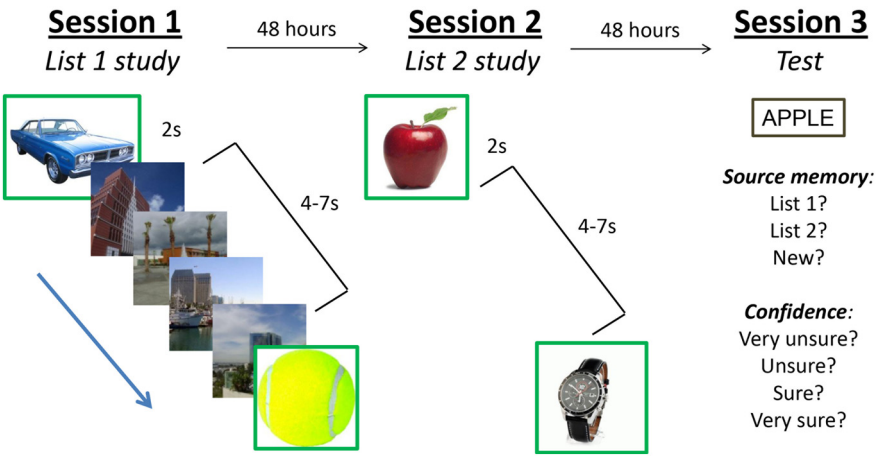

Figure 1. Experimental design. In session 1, participants studied a list of 20 items (object pictures), repeated four times. Scene images were presented during the ITIs. Forty-eight hours later, participants received a reminder of session 1, and then studied a different list of items. The ITIs during session 2 were empty. Forty-eight hours later, participants were shown the studied items (presented as words) from both lists, intermixed with novel items, and asked to make source memory judgments, as well as report their confidence.

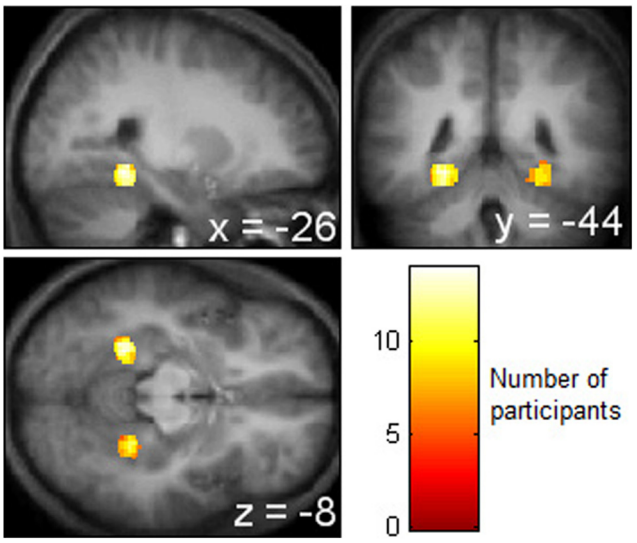

Figure 2. PPA. The average functionally defined ROI is shown in standardized (Montreal Neurological Institute) space. The color of each voxel corresponds to the number of participants whose individual ROI (in standardized space) contained that voxel. Standardization of the ROIs (by nonlinear warping to a template) was done for illustration purposes only; the ROIs used in all our analyses were in each participant's native space.

items would be subsequently misattributed to L1 on the final memory test.

Materials and Methods

Participants. Fourteen right-handed participants (age 18–30, seven female) participated in the study. All were free of neurological or psychiatric disease, and fully consented to participate. The study protocol was approved by the Institutional Review Board for Human Subjects at Princeton University.

Stimuli and task. The experiment consisted of three sessions, each separated by ~48 h. All sessions took place inside the fMRI scanner. The experimental design is shown schematically in Figure 1. During the first session, participants studied a list of 20 items (object pictures with a green frame), presented sequentially for 2 s each, followed by an intertrial interval (ITI) randomly jittered between 4 and 7 s. During session 1, the ITI was filled with a continuous sequence of random scene images (duration 1 s); during session 2, the ITI was a blank screen. Participants were informed that their memory would be tested for the object items but not the scenes. The list was presented four times in random order, each time followed by a free recall task in which participants were asked to verbally recall the names of objects studied in the list. The free recall task was performed inside the scanner, between functional scans.

Before the beginning of session 2, participants were given a "reminder" of session 1, analogous to the reminders used in previous studies (Hupbach et al., 2007, 2008, 2009, 2011). Specifically, participants were asked to recall the general procedure during session 1 (they were stopped if they began to recall particular studied items). Invariably, they described studying and recalling a list of items. The rest of session 2 proceeded in an identical manner to session 1: participants studied a new list of 20 items four times, with free recall after each list repetition. Note that, contrary to Hupbach's studies, we did not include a "no-reminder" condition. The scanner environment is itself a very strong reminder and we hypothesized that all scanned participants would recollect the session 1 context when they returned for session 2, regardless of the instructions that they were given.

During session 3, participants performed a source memory task in which they were asked to judge whether an item (presented as an object name) was studied in L1, L2, or neither (i.e., a new item). After each source memory judgment, participants were asked to rate their confidence on a 4-point scale (very unsure, unsure, sure, very sure). Responses were recorded using a button box.

Following the source memory task, we ran a scene localizer in which participants viewed alternating mini-blocks of scene and phase-scrambled scene images. Each mini-block consisted of eight images, each presented for 500 ms and separated by a 1.5 s ITI. A total of 16 mini-blocks were presented, each separated by 12 s. To keep participants attentive, they were asked to press a button each time they detected a repeated image.

fMRI data acquisition. Data were acquired using a 3 T Siemens Allegra scanner with a volume head coil. We collected four functional runs in sessions 1 and 2 and two functional runs in session 3 with a T2*-weighted gradient-echo echo-planar sequence (35 oblique axial slices, 3.5 × 3.5 mm in plane, 3.5 mm thickness; echo time 28 ms; TR 2000 ms; flip angle 71°; field of view 224 mm). We collected two anatomical runs for registration across sessions and across participants to standard space: a coplanar T1-weighted FLASH sequence and a high-resolution 3D T1-weighted MPRAGE sequence. A FLASH image was acquired for each session, while only a single MPRAGE was acquired per participant.

fMRI data preprocessing. Preprocessing was performed using Statistical Parametric Mapping software (SPM8; Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK). Images were coregistered across sessions using an affine transformation of the FLASH images and aligned to correct for participant motion. The data were then high-pass filtered with a cutoff period of 128 s. No spatial normalization or smoothing was applied to the data.

Region of interest selection. The goal of our region of interest (ROI) selection procedure was to localize the parahippocampal place area (PPA), a region that is associated with scene processing (Epstein and Kanwisher, 1998; Epstein et al., 1999). A general linear model (GLM) was fit to the localizer data for each participant using "scene" and "scrambled scene " regressors that were convolved with the canonical hemodynamic response function. The six scan-to-scan motion parameters produced during motion correction were also included as nuisance regressors in the GLM to account for residual effects of movement. A t statistic map was then created for the scene > scrambled scene contrast, thresholded at $p < 0.001$ (uncorrected). Bilateral clusters corresponding anatomically to the PPA in the posterior parahippocampal/collateral

sulcus region were selected as functional ROIs for each participant individually (Fig. 2).

Multivariate pattern analysis. For each stimulus presentation in the localizer run (scene or scrambled scene), we recorded the vector of voxel responses within the PPA ROI 6 s after stimulus onset (Polyn et al., 2005; McDuff et al., 2009; i.e., at the time of the expected peak of the hemodynamic response). We also computed resting-state patterns during the localizer run, by recording the vector of voxel responses within the PPA ROI during the second half of each interblock interval (averaging across TRs within this time period).

These patterns were entered into an L_2-regularized multinomial logistic regression classifier, trained to predict scene versus scrambled scene versus rest labels. The regularization parameter was set to 0.1, but we found that the results were insensitive to varying this parameter >3 orders of magnitude. The trained classifier was then applied to individual brain scans from the study and test phases: For each scan, the classifier provided a readout of the probability that the scan corresponded to a scene image; we will refer to this real-valued number (bounded between 0 and 1) as scene evidence. We compared scene evidence in different conditions and at different time points using paired-sample *t* tests; all tests were two-tailed.

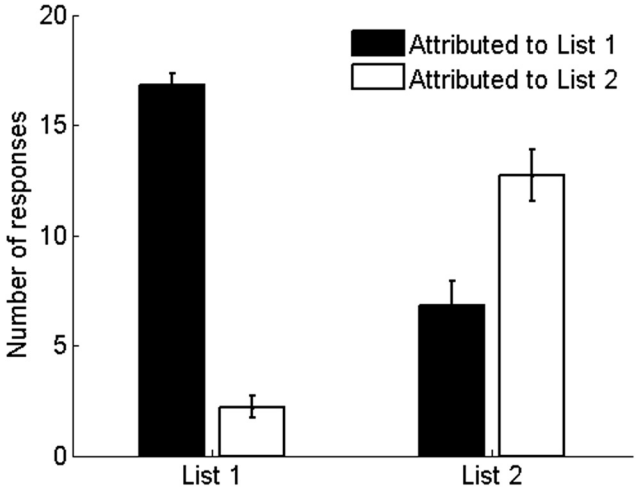

Figure 3. Behavioral results. Number of items on the source memory test attributed to L1 or L2 for each list. Error bars indicate ± 1 SEM.

Results

Behavioral results

As shown in Figure 3, we replicated the asymmetric pattern of intrusions found by Hupbach et al. (2007): L2 items were more frequently misattributed to L1 items than vice versa ($t_{(13)} = 4.00$, $p < 0.002$). The overall level of false alarms (novel items judged as old) was low (median = 2). The overall level of misses (old items judged as new) was also low (median = 1).

Neuroimaging results

TCM posits that items are bound to the context that is in place when the item is presented (Howard and Kahana, 2002). As such, we predicted that high levels of scene classifier evidence (indicating L1 contextual reinstatement) in the time points leading up to studying an L2 item would lead to misattribution of that item to L1. To test this hypothesis, we measured scene classifier evidence at six time points time locked to the appearance of L2 items in session 2. These six time points encompassed a time window ranging from 4 s before the appearance of the item to 6 s after the item appeared (Fig. 4*A*); note that the time points in the figure were not shifted to account for hemodynamic lag. We focused our analyses on $t = -4, -2$, and 0 s (unshifted) because these were the only time points that unambiguously reflected prestimulus activity. All subsequent time points ($t = 2, 4, 6$ s) were potentially contaminated by the evoked response to the stimulus (these later time points were only included in Fig. 4*A* for comparability with the test-phase analyses described below). For each item and time point, we computed the average level of scene evidence across the four presentations of that item during session 2. These average scene evidence values were sorted according to whether L2 items were subsequently correctly attributed to L2 (red line) or misattributed to L1 (blue line). The results show that, at 2 s before trial onset (unshifted), scene evidence was significantly higher for L2 items subsequently misattributed to L1 compared with those subsequently correctly attributed to L2 ($t_{(13)} = 2.71$, $p < 0.02$; $p = 0.05$, using a Bonferroni correction for multiple comparisons across $t = -4, -2$, and 0 s). The finding that a significant difference was present at $t = -2$ s in the unshifted response suggests that scene activity ~6 s before stimulus onset was predictive of subsequent source misattributions. No significant differences in scene evidence were found at other time points.

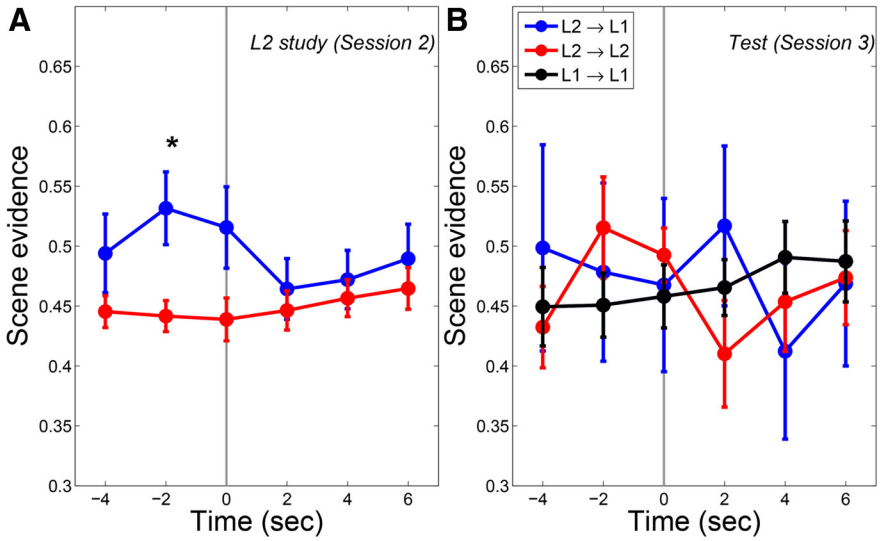

Figure 4. Neuroimaging results. ***A***, Time course of scene classifier evidence during study of L2 items in session 2 (note that 0 s = the time of the L2 item's appearance, not adjusted for the hemodynamic response). The blue line represents L2 items subsequently misattributed to L1, and the red line represents L2 items correctly attributed to L2. The vertical line represents the trial onset. Scene evidence was operationalized as the classifier's estimate of the probability that the mental state of "scene" was present. ***B***, Time course of scene evidence during the source memory test (session 3). The black line represents L1 items correctly attributed to L1. Error bars indicate ± 1 SEM. Asterisks indicate time points at which L2 → L1 scene evidence is significantly different from L2 → L2 scene evidence ($p \leq 0.05$, two-tailed *t* test, using a Bonferroni correction for multiple comparisons).

When an L2 item is presented on the source memory test, our theory posits that the L2 item will be misattributed to L1 if it evokes reinstatement of the L1 context. As such, we predicted that high levels of scene activation in the time points following the presentation of an L2 item would be associated with enhanced levels of misattribution. To test this hypothesis, we measured scene evidence time locked to the appearance of items on the final

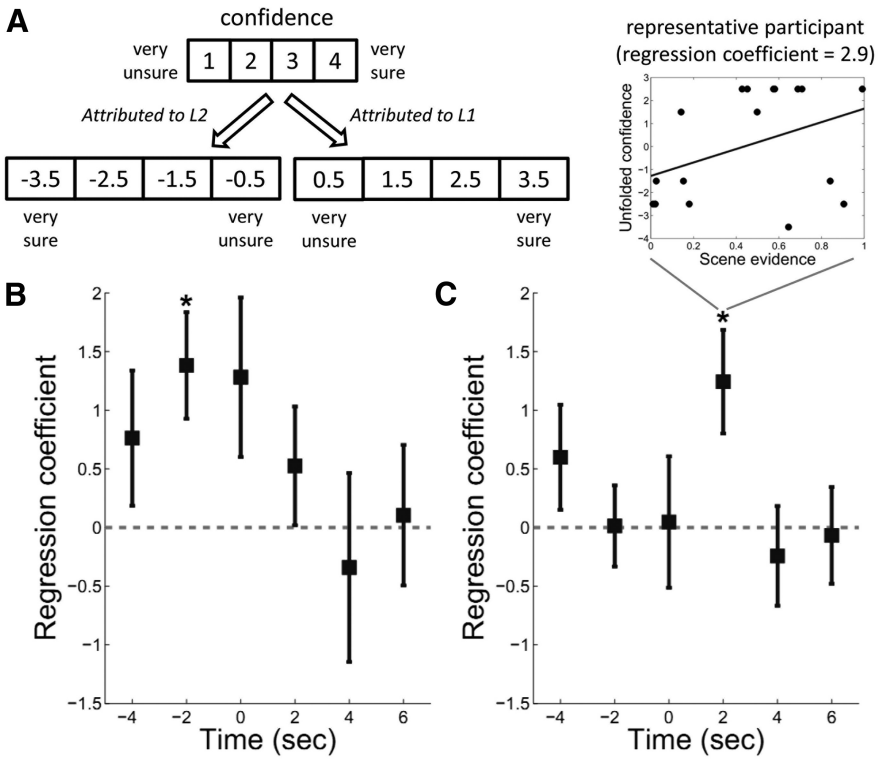

Figure 5. ***A***, Schematic of the confidence ratings analysis. The confidence ratings were "unfolded" into an 8-point scale ranging from "very sure L2" to "very sure L1." A representative participant's data are shown on the right. Unfolded confidence was regressed onto scene evidence for each participant separately, and the regression coefficients were then entered into a *t* test. ***B***, Strength of relationship between scene evidence during each time point of L2 study trials and unfolded confidence for L2 items. ***C***, Strength of relationship between scene evidence during each time point of L2 test trials and unfolded confidence for L2 items. Error bars indicate ± 1 SEM. Asterisks indicate coefficients that are significantly different from 0 ($p \leq 0.05$, two-tailed *t* test, using a Bonferroni correction for multiple comparisons).

test (this time, in addition to including L2 items that were attributed to L1 and L2, we also included L1 items subsequently attributed to L1; the number of L1 items attributed to L2 was too small to include this condition in the analysis); these results are shown in Figure 4*B*. For this analysis, we focused on the scans collected 2, 4, and 6 s after stimulus onset (unshifted). The 2 s time point captures the onset of the stimulus-evoked hemodynamic response and the 4 and 6 s time points capture the peak of this evoked response (the $t = -4, -2$, and 0 s time points are included in Fig. 4*B* purely for comparability with Fig. 4*A*). There was a numerical trend in the predicted direction at 2 s after item presentation (i.e., greater scene evidence for L2 items attributed to L1 vs L2 items attributed to L2), but none of the differences between conditions were significant.

To obtain a more fine-grained picture of how scene evidence related to memory performance, we next examined whether scene evidence was predictive of parametric differences in response confidence. We first converted the confidence ratings to an "unfolded" scale, where −3.5 represents "sure L2" and +3.5 represents "sure L1" (Fig. 5*A*). We then fit, for each time point during the trial, a linear regression model with scene evidence as the predictor and unfolded confidence as the response variable. This regression analysis was done separately for each participant. To get an overall measure of the strength of the relationship between scene evidence and response confidence at a particular time point, we averaged these regression coefficients across participants.

Figure 5*B* shows the results for L2 study: unfolded confidence increased as a function of scene evidence for prestimulus time points, as indicated by positive regression coefficients for these time points. Thus, the stronger the reinstatement of L1 context (during L2 study, before stimulus onset), the more confident participants were that L2 items were studied in L1. To assess the significance of these results, we performed a *t* test on the regression coefficients at each time point against zero. This analysis revealed a significant effect at 2 s before trial onset ($t_{(13)} = 3.04$, $p < 0.01$; $p < 0.03$, using a Bonferroni correction for multiple comparisons across $t = -4, -2$, and 0 s).

At test (Fig. 5*C*), scene evidence was significantly correlated with unfolded confidence 2 s after trial onset ($t_{(13)} = 2.82$, $p < 0.02$; $p < 0.05$, using a Bonferroni correction for multiple comparisons across $t = 2$, 4, and 6 s). Thus, for this more sensitive analysis, the predicted relationship was observed at test: Greater levels of scene evidence during the poststimulus time period were associated with greater confidence that the item was studied as part of L1. This finding of a relationship between classifier evidence and behavior at 2 s poststimulus-onset is consistent with the results of a previous source memory study from our lab (McDuff et al., 2009). The fact that the effect appeared at 2 s (i.e., before the expected peak of the hemodynamic response to the cue) suggests that, in this task, source confidence was more closely related to the onset time of the evoked hemodynamic response than to the peak amplitude of this response. Alternatively, the early timing of this effect also fits with the possibility that prestimulus scene activity was driving the effect; that is, prestimulus scene activity may have biased participants to attribute items to L1 at test (for other studies showing that prestimulus activity can influence memory test responding, see Quamme et al., 2010; Addante et al., 2011).

Discussion

In this study, we investigated a mechanistic explanation for the asymmetric pattern of memory misattributions observed by Hupbach et al. (2007, 2008, 2009, 2011). Participants studied two lists (L1 and L2) on separate days, with scene images presented during the ITIs of L1 but not L2. Before studying L2, participants were reminded of the L1 study session. Using scene-related activation as a proxy for L1 mental context reinstatement, we found that the amount of scene classifier evidence during L2 study predicted which items would be subsequently misattributed to L1 in a source memory test. The amount of scene evidence, during both L2 study and test, also parametrically predicted the confidence with which these memory misattributions were made. These findings are consistent with a theoretical explanation of retrieval-induced memory misattribution set forth by Sederberg et al. (2011).

It is conspicuous that the amount of scene evidence before trial onset (during L2 study) predicted subsequent misattribu-

tion errors, but the amount of scene evidence following trial onset did not predict these errors. Furthermore, the parametric relationship between scene activation and memory confidence was only statistically significant 2 s before trial onset. These findings are predicted by the TCM, which posits that items are bound to contextual features that are active in the time period leading up to item presentation (Howard and Kahana, 2002).

In this paper, we have argued that memory misattribution errors are caused by contextual reinstatement (of L1 context during L2 study). However, other explanations are possible. For example, if participants were simply mind wandering (Mason et al., 2007) during some of the L2 study trials, then those items would be poorly encoded and hence would be more likely to be misattributed to L1 during test (note that this explanation is not mutually exclusive with the contextual-reinstatement explanation). The fact that we trained our classifier on a third "rest" category gives us a means of evaluating this mind-wandering hypothesis: if we assume that more mind wandering was taking place during rest periods in the functional localizer (compared with "scene" and "scrambled scene" periods), then we can use rest classifier evidence as an index of how much participants were mind wandering during L2 study. To the extent that mind wandering was responsible for misattribution errors, we would expect to see a difference in rest classifier evidence during L2 study for subsequently misattributed (vs correctly attributed) items. This pattern was not observed (all p values for the time point-specific comparisons were >0.17), suggesting that mind wandering was not a major cause of misattribution errors in this study (note also that there is nothing about the mind-wandering explanation that would predict asymmetric misattributions; in principle, mind wandering could cause misattribution of L1 items to L2 as well as misattribution of L2 items to L1).

A number of recent studies have used pattern classifiers to track memory retrieval. For example, Polyn et al. (2005) showed that category-specific neural activation precedes the recall of items and can predict the category of item recalled. Johnson et al. (2009) used this method to show that retrieval of encoding-task information occurs during recognition tests, and Kuhl et al. (2012) used pattern classifiers to track competition between paired-associate memories (for review, see Rissman and Wagner, 2012). Note that all of the aforementioned studies tracked retrieval of the properties of the to-be-remembered items (e.g., their category, or the encoding task performed on the item). Our study is different. Instead of directly tracking features of the to-be-remembered items, we tracked reinstatement of incidentally presented scene information that was not onscreen at the same time as the to-be-remembered objects. In the Polyn et al. (2005) study, it was ambiguous whether the information tracked by the classifier reflected contextual properties (i.e., remembering the "face study context") or item properties (i.e., remembering a specific face). In the present study, we used different types of information (scene vs object) for context versus items, thereby making it possible to specifically track contextual reinstatement.

Our study also adds to the growing literature on the neural basis of memory misattribution errors. For example, Stark et al. (2010) recently conducted an fMRI study using a variant of the Loftus et al. (1978) misinformation paradigm (Okada and Stark, 2003). In the Stark et al. (2010) study, participants viewed pictures of an event during an initial encoding phase, and then were given a mix of correct and incorrect information about the initial event (presented auditorily) during a second phase; when participants were asked about the initial event, they sometimes misattributed false information from the second phase to the initial encoding phase. Mirroring other fMRI studies of false memory formation (for review, see Schacter and Slotnick, 2004), the Stark et al. (2010) study mainly focused on the role of retrieved sensory information, both in driving false memories and in distinguishing between true and false memories. Specifically, Stark et al. (2010) found that activity was higher in both auditory cortex and early visual cortex at test for falsely attributed items (misinformation items that were misattributed to the original event) compared with correct rejections; they also found that, on average, activity was higher in early visual cortex for true than false memories. Our study differs from the Stark et al. (2010) study in two ways: First, as stated above, our study focused on retrieval of sensory properties of the context, as opposed to sensory properties of the to-be-remembered stimuli. Second, our study focused on how retrieval of original-event information during the second event affects subsequent list memory, whereas Stark et al. (2010) focused on how brain activity on the final test related to memory accuracy.

Finally, we consider how these results relate to memory reconsolidation. According to reconsolidation theory, retrieving a memory makes its molecular substrate malleable; when the memory is in this malleable state, it can be changed or even erased (for review, see Lee, 2009). The original Hupbach et al. (2007, 2008, 2009, 2011) papers explained the asymmetric pattern of intrusions in terms of reconsolidation: They argued that, when participants were reminded of L1 before L2 study, this made the L1 memory malleable, whereupon it was "updated" with subsequently presented L2 items. Sederberg et al. (2011) argued that these findings could instead be explained in terms of item-context associations, without making any reference to cellular-level reconsolidation. Having said this, we should emphasize that our results do not rule out a reconsolidation account. It may be the case that the contextual reinstatement that we are measuring in our study (using the scene pattern classifier) triggers a cellular reconsolidation process that promotes integration of L2 items into the L1 memory.

In conclusion, our findings provide further empirical constraints on mechanistic theories of memory misattribution. A context reinstatement account (Sederberg et al., 2011) seems to provide the most parsimonious explanation of our data, but further experiments will be required to draw any decisive conclusions (e.g., relating to the relevance of reconsolidation). There is a rich theoretical and experimental literature on the role of temporal context in shaping memory retrieval (Polyn and Kahana, 2008). Now that we have a simple neural "tracer" for mental context reinstatement, we can begin to investigate in more detail how the dynamics of context reinstatement influence memory errors.

References

Addante RJ, Watrous AJ, Yonelinas AP, Ekstrom AD, Ranganath C (2011) Prestimulus theta activity predicts correct source memory retrieval. Proc Natl Acad Sci U S A 108:10702–10707. CrossRef Medline

Epstein R, Kanwisher N (1998) A cortical representation of the local visual environment. Nature 392:598–601. CrossRef Medline

Epstein R, Harris A, Stanley D, Kanwisher N (1999) The parahippocampal place area: recognition, navigation or encoding? Neuron 23:115–125. CrossRef Medline

Howard MW, Kahana MJ (2002) A distributed representation of temporal context. J Math Psychol 46:269–299. 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

Hupbach A, Hardt O, Gomez R, Nadel L (2008) The dynamics of memory: context-dependent updating. Learn Mem 15:574–579. CrossRef Medline

Hupbach A, Gomez R, Nadel L (2009) Episodic memory reconsolidation: updating or source confusion? Memory 17:502–510. CrossRef Medline

Hupbach A, Gomez R, Nadel L (2011) Episodic memory updating: the role of context familiarity. Psychon Bull Rev 18:787–797. 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

Kuhl BA, Bainbridge WA, Chun MM (2012) Neural reactivation reveals mechanisms for updating memory. J Neurosci 32:3453–3461. CrossRef Medline

Lee JL (2008) Memory reconsolidation mediates the strengthening of memories by additional learning. Nat Neurosci 11:1264–1266. CrossRef Medline

Lee JL (2009) Reconsolidation: maintaining memory relevance. Trends Neurosci 32:413–420. CrossRef Medline

Loftus EF, Miller DG, Burns HJ (1978) Semantic integration of verbal information into a visual memory. J Exp Psychol Hum Learn Mem 4:19–31. CrossRef

Mason MF, Norton MI, Van Horn JD, Wegner DM, Grafton ST, Macrae CN (2007) Wandering minds: the default network and stimulus-independent thought. Science 315:393–395. CrossRef Medline

McDuff SG, Frankel HC, Norman KA (2009) Multivoxel pattern analysis reveals increased targeting and reduced use of retrieved details during single-agenda source monitoring. J Neurosci 29:508–516. CrossRef Medline

Okado Y, Stark C (2003) Neural processing associated with true and false memory retrieval. Cogn Affect Behav Neurosci 3:323–334. CrossRef Medline

Polyn SM, Kahana MJ (2008) Memory search and the neural representation of context. Trends Cogn Sci 12:24–30. CrossRef Medline

Polyn SM, Natu VS, Cohen JD, Norman KA (2005) Category-specific cortical activity precedes recall during memory search. Science 310:1963–1966. CrossRef Medline

Quamme JR, Weiss DJ, Norman KA (2010) Listening for recollection: a multi-voxel pattern analysis of recognition memory retrieval strategies. Front Hum Neurosci 4. pii:61. Medline

Rissman J, Wagner AD (2012) Distributed representations in memory: insights from functional brain imaging Annu Rev Psychol 63:101–128. CrossRef

Roediger HL 3rd, Butler AC (2011) The critical role of retrieval practice in long-term retention. Trends Cogn Sci 15:20–27. CrossRef Medline

Schacter DL, Slotnick SD (2004) The cognitive neuroscience of memory distortion. Neuron 44:149–160. CrossRef Medline

Sederberg PB, Gershman SJ, Polyn SM, Norman KA (2011) Human memory reconsolidation can be explained using the Temporal Context Model. Psychon Bull Rev 18:455–468. CrossRef Medline

Stark CE, Okada Y, Loftus EF (2010) Imaging the reconstruction of true and false memories using sensory reactivation and the misinformation paradigms. Learn Mem 17:485–488. CrossRef Medline