Neural context reinstatement predicts memory misattribution

801.13/ CCC25

Samuel J. Gershman,[1,2] Anna C. Schapiro,[2,3] Almut Hupbach,[3*] & Kenneth A. Norman[1,2*]
[1]Department of Psychology, [2] Princeton Neuroscience Institute, Princeton University
[3]Department of Psychology, Lehigh University *equal contribution

ABSTRACT

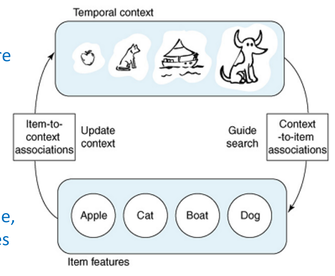
What causes new information to be mistakenly attributed to an old experience? The reconsolidation account posits that the reactivation of a memory re-introduces fragility, allowing for memory modification. Presenting new information after a prior experience has been reactivated can result in the new information becoming associated with the prior, reactivated experience. Hupbach et al. (2007) experimentally demonstrated that, when subjects were reminded of a list of previously encoded items (L1) prior to learning a new list (L2), a significant subset of L2 items was later misattributed to L1. The present study attempted to probe the neural correlates of this effect using fMRI and multi-voxel pattern analysis. Specifically, L1 items were accompanied by a stream of scene images during the intertrial interval, whereas the interval between L2 items was not. This allowed us to measure scene activation during the L2 intertrial intervals as an index of L1 context reinstatement. We found that L2 items subsequently misattributed to L1 were associated with significantly higher scene activation than correctly attributed L2 items. These results support the mechanistic interpretation of asymmetric misattributions offered by TCM (Sederberg et al., 2011), and they are also compatible with a memory reconsolidation account.

1. Background

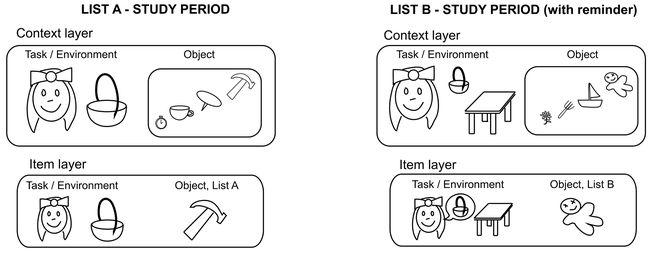
- Why do we misattribute items from one source to another?
- One idea is that reactivating a memory causes it to be associated with new information at the time of reactivation. TCM (Howard & Kahana, 2002; see below) is a computational instantiation of this idea.
- Hupbach et al. (2007) asked subjects to study two lists (L1 & L2) of objects, separated by 48 hours. A subtle reminder of L1's context prior to L2 study caused subjects to have an asymmetric pattern of intrusions, misattributing L2 items to L1, but not vice versa.
- Sederberg et al. (2011) proposed an explanation of these findings in terms of TCM.

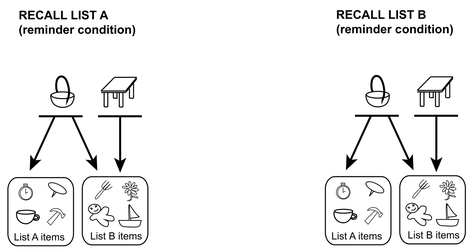
The Temporal Context Model (TCM)

- Items are represented as feature vectors
- The temporal context is a superposition of recently experienced items
- Items are bound to temporal context via Hebbian learning
- Context is used as a retrieval cue, and retrieving an item reinstates its context



What happens in TCM:
•L1 items are bound to L1 context
•After the L1 reminder prior to L2 study, L2 items are bound to *both* L1 and L2 contexts

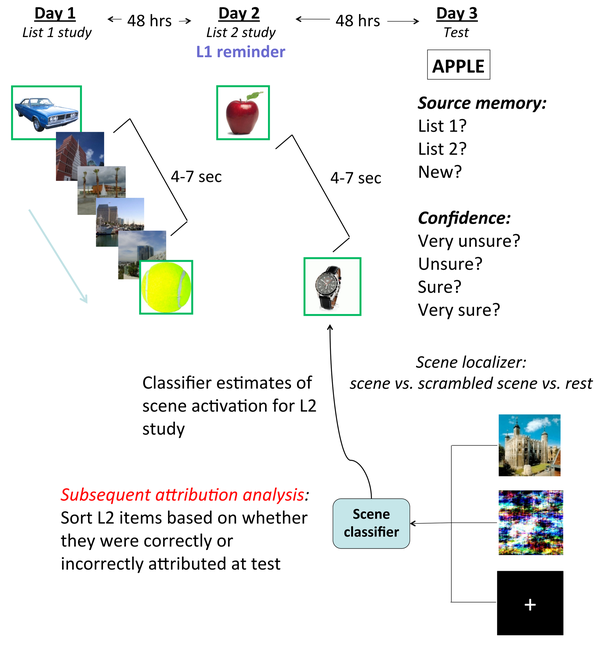




At retrieval, L1 context serves as a retrieval cue for L1 & L2 items, whereas L2 context is a retrieval cue only for L1 items, thus producing the misattribution asymmetry.

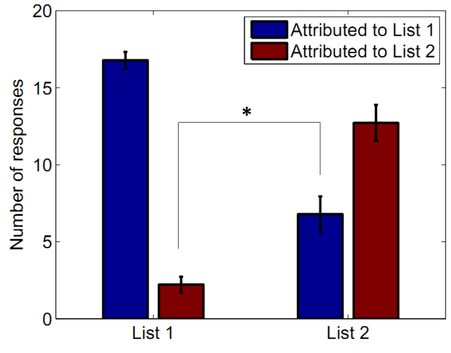
This account asserts that it is the reinstatement of L1 context that leads to misattributions. We set out to test this assertion by measuring the neural reinstatement of L1 context during L2 study. We displayed scene images during L1 study but not L2 study; during L2 study we used scene activation as a neural proxy for L1 context reinstatement.

2. Experimental design



- Each list consisted of 20 items (object pictures) presented one at a time for 2 sec.
- ITI during L1 was filled with a continuous stream of scene images.
- Participants (N = 14) studied each list 4 times (random order of items), followed by free recall.
- All phases done in the scanner.
- L1 reminder prior to L2 study: "Describe the general procedure during the previous study session."
- We used an L2-regularized multinomial logistic regression classifier. "Rest" class was defined as a the last 6 seconds of the 12-sec interval between scene / scrambled scene blocks.

3. Behavioral results



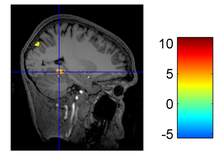
We replicated Hupbach's asymmetric misattribution effect: **significantly more L2→L1 misattributions than L1→L2 misattributions** ($p < 0.002$).

Number of false alarms (novel items judged as old) was low: median = 2.

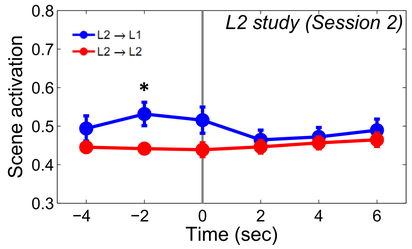
* = $p < 0.05$

4. Neural results

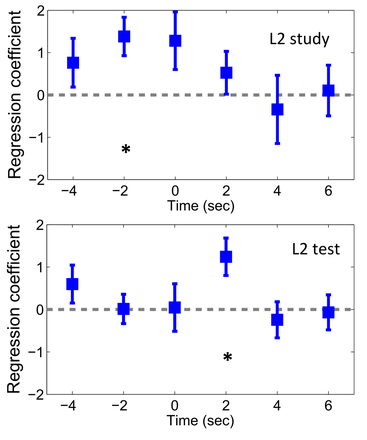
Functional localizer used to constrain classifier input to the parahippocampal place area (**PPA**)



TCM posits that items are bound to the context that is in place when the item is presented. So we expect that high levels of scene activation (indicating L1 contextual reinstatement) in the time points **leading up** to studying an L2 item will lead to misattribution of that item to L1.



Pre-stimulus scene activation during L2 study was higher for items subsequently misattributed to L1 than for items that were subsequently correctly attributed to L2.



We constructed "unfolded" confidence ratings, ranging from -3.5 ("very sure L1") to +3.5 ("very sure L2"). We then performed a linear regression predicting unfolded confidence from scene activation at each point in the trial.

Confidence was significantly predicted by scene activation during L2 study (2 sec before stimulus onset). This fits with the idea that parametric variations in scene reinstatement determine how strongly L2 items will be associated with L1 context.

Confidence was also predicted by scene activation 2 sec after stimulus onset at test. The test phase results fit with the idea that scene reinstatement triggered by the test item causes subjects to make attributions to L1.

5. Conclusions

- Replicated asymmetry of memory misattributions in Hupbach et al.'s (2007) reminder paradigm.
- Memory misattributions are predicted by context reinstatement, measured by decoding scene activation from brain patterns.
- Supports TCM's context-binding explanation of memory misattributions in this paradigm.

Acknowledgments. This research was supported by NIH RO1 MH069456 awarded to Kenneth A. Norman. Samuel J. Gershman was supported by a graduate research fellowship from the NSF.

References

Howard, MW, & Kahana, MJ (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46*, 269-299.

Hupbach, A., Gomez, R., Hardt, O., Nadel, L. (2007). Reconsolidation of episodic memories: a subtle reminder triggers integration of new information. *Learning & Memory, 14*, 47-53.

Sederberg, P.B., Gershman, S.J., Polyn, S.M., & Norman, K.A. (2011). Human memory reconsolidation can be explained using the Temporal Context Model. *Psychonomic Bulletin & Review*, *18*, 455-468.