Recall order is predicted by category-specific neural activity of preceding items at study

Stephanie C.Y. Chan[1] Marissa C. Applegate[1] Neal W. Morton[2] Sean M. Polyn[2] Kenneth A. Norman[1]

[1] Princeton University [2] Vanderbilt University

PRINCETON Neuroscience INSTITUTE

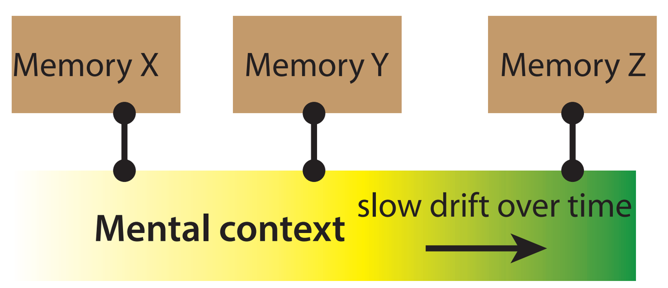
Princeton Computational Memory Lab

1 Introduction

Question

Do we timestamp our memories using the thoughts that are co-active at the time of encoding? Specifically, do we use the *meanings* of those thoughts to timestamp and organize our memories?

Background



Memories are timestamped by the concurrent state of "mental context", which drifts slowly over time.

We use this signal to help retrieve memories. Specifically:
- When we retrieve a memory, we reinstate the mental context associated with that memory
- We use that retrieved mental context to cue for other memories.

Consequence: During recall, we are more likely to transition between items that were associated with similar mental contexts.

Two possibilities

What is mental context?

The semantic hypothesis: Mental context is a recency-weighted average of the meanings of our thoughts (Howard & Kahana 2002)

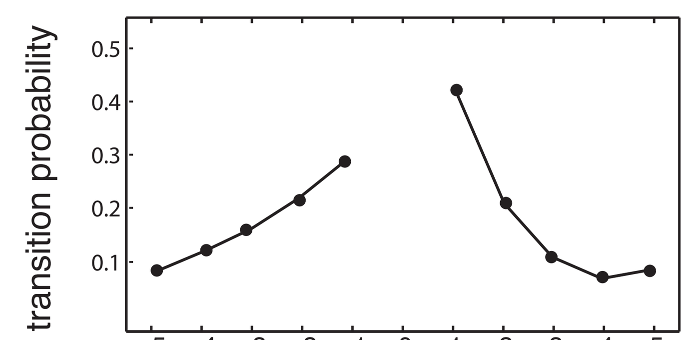
or

The random drift hypothesis: Mental context is a randomly-drifting signal (Estes 1955)

Approach

Show that the **semantics of preceding items** affect recall order.

Recall behavior on lists of random words



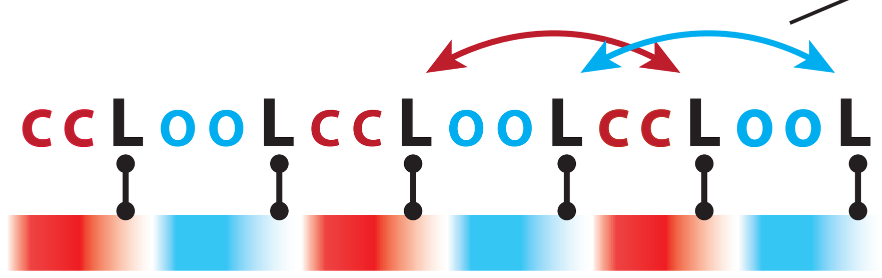
2 Our "Evel Knievels" free recall task

Study List → **Categorical Recall**

Subjects are presented with a list of 18 items. The items belong to one of three categories: celebrities, landmarks, or objects.

Subjects recall as many *landmarks* as possible, in any order.

Study-list structure



Evel Knievels (EK)

Recall transitions between landmarks with semantically similar preceding items...

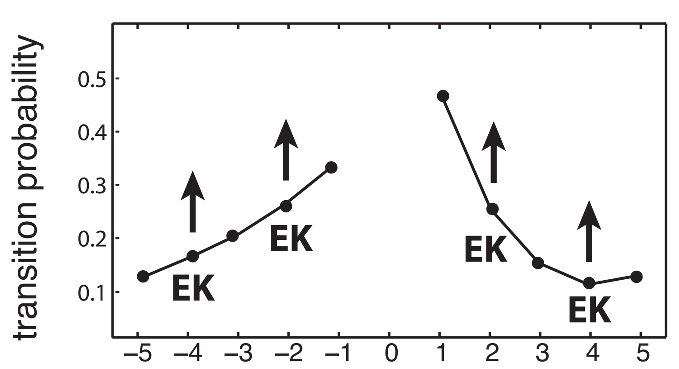
...*over* temporally adjacent landmarks

Items are timestamped with alternating semantic context.

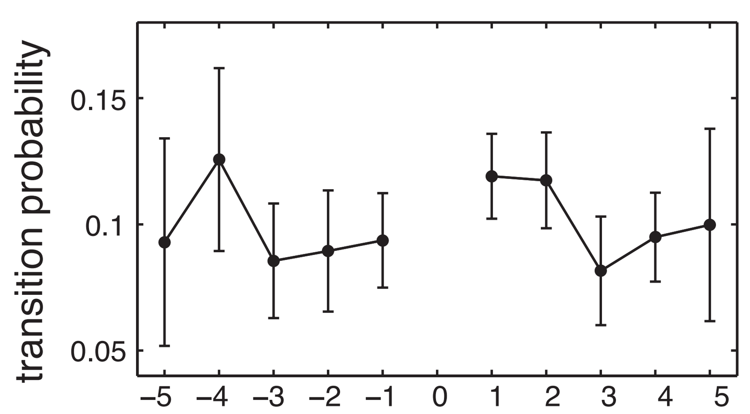
(categories are counterbalanced across lists)

3 Behavioral evidence

What we might expect



Our data



Lots of variance-- not strong evidence.

n = 24

4 fMRI decoding of semantic context

Prediction

We think that the category of the preceding item enters into mental context. Because these preceding items are processed and linger to varying degrees, there will be some naturally-occurring variance in contextual similarity for Evel Knievel pairs.

We can use an fMRI pattern classifier to track this variance.

Prediction: Evel Knievel recall transitions are more likely to occur when (according to brain data) the two items were studied in similar semantic contexts

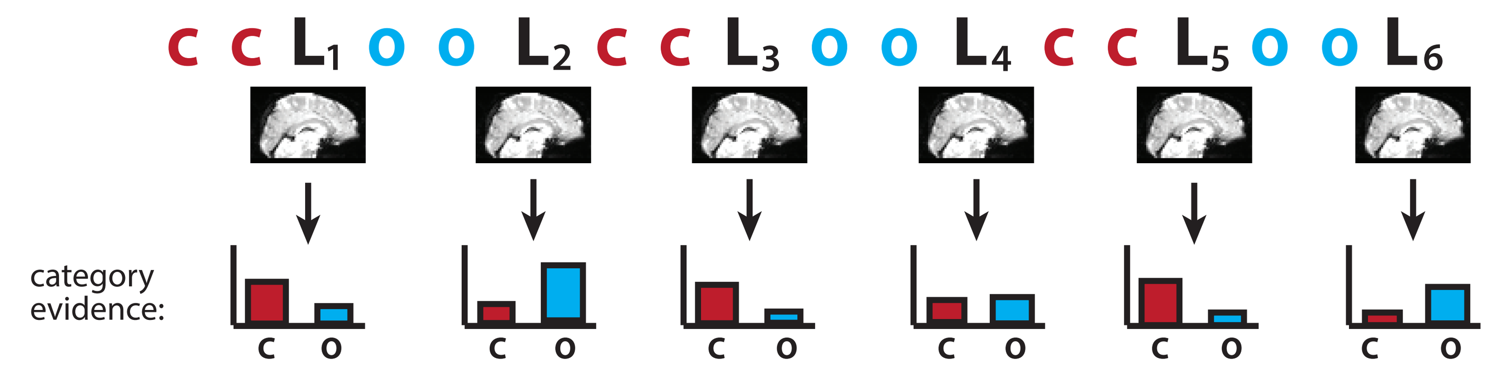
Method

Multivariate pattern analysis (MVPA)

We estimate the semantic context for each item by applying a brain decoder to the fMRI data from study.

The brain decoder was trained to identify the ***preceding*** category

performance = 0.64 (greater than 0.50, $p < 10^{-8}$)



EK transition scores - computed using brain activity from study

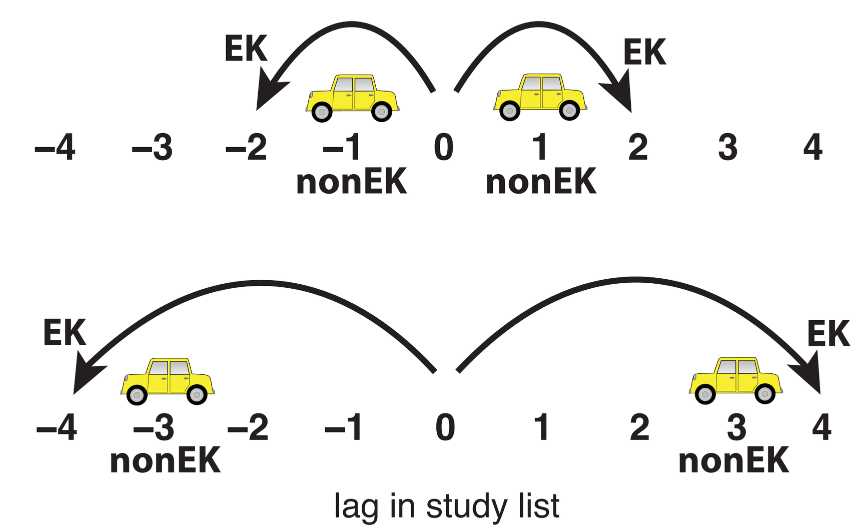
We use category evidence from the MVPA brain decoder to evaluate the similarity in semantic context between two items:

Semantic context similarity score	*+1*	*+1*	*-1*	*-1*
MVPA outputs for Item 1	C > O	O > C	C > O	O > C
MVPA outputs for Item 2	C > O	O > C	O > C	C > O

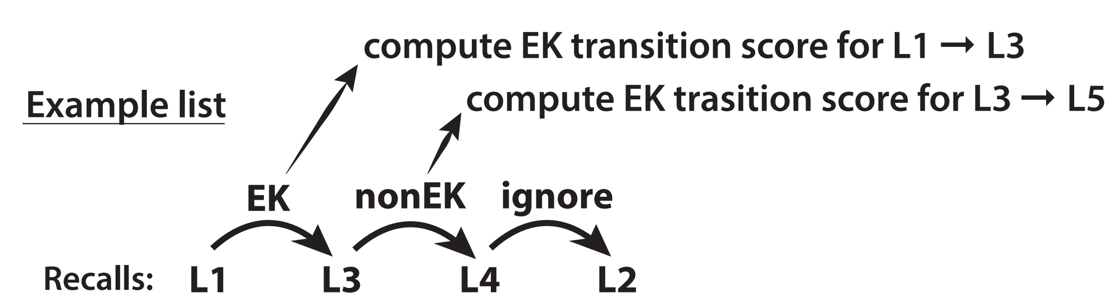
Analysis Strategy - can EK transition scores predict recall order?

We compare the situations in which
(A) an EK transition was made over a slightly closer nonEK
(B) a nonEK transition was made instead of a slightly farther EK

To be fair, we do not include those situations in which the slightly closer/farther transition was not possible.



Is the EK transition score higher for situations in which the EK actually occurred?

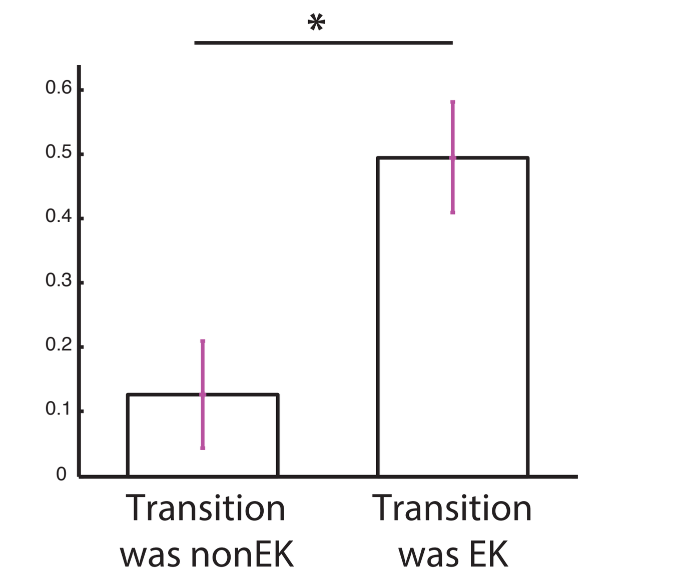


Results

Neural decoder trained on study data to identify *preceding* category predicts transition likelihood at recall.

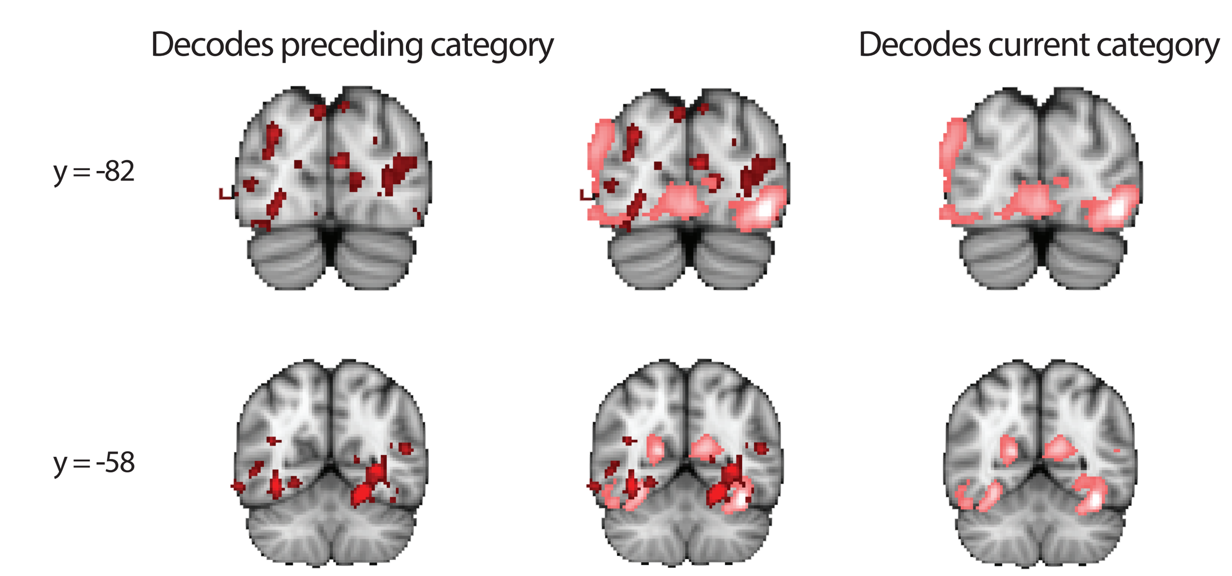
mean semantic context similarity score for the potential EK

* random effects $p < 0.001$, $n = 17$



This is a robust effect that can be obtained using different measures of semantic context similarity.

Importance maps for brain decoder for category "Celebrity"



5 Conclusions

Recall order seems to be affected by **the semantics of the items studied immediately before** each recalled item.

This is congruent with the theory that we timestamp our memories using the meanings of the thoughts that are co-active at the time of encoding, and that we use these timestamps to help us retrieve our memories.

I - Behavioral analysis
Effects of prior-item semantics are difficult to observe in a behavioral-only analysis.

II - MVPA analysis
During recall, subjects were likely to transition between a pair of items if they were encoded in a similar semantic context, as estimated by an fMRI brain decoder.

Acknowledgements

This research was supported by the NSF/NIH Collaborative Research in Computational Neuroscience Program, grant # NSF IIS-1009542, and the NIH grant # 2T32MH065214.

Citations

Estes, W.K. 1955. Statistical theory of spontaneous recovery and regression. Psychol. Rev. 62: 145 -154.

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