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

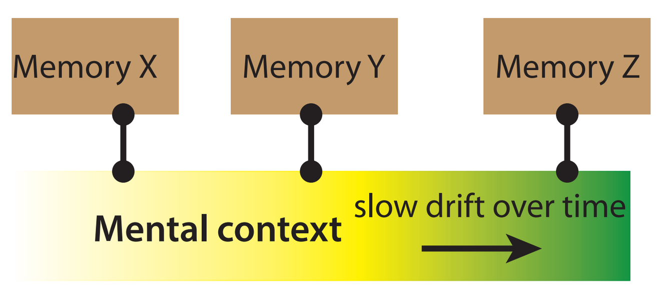
1 Introduction

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 our memories?

Background

Context-based models of memory



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 are more likely to recall memories whose mental context is most similar to the current mental context

Consequence on recall order: We are more likely to transition between two items that were associated with similar mental contexts.

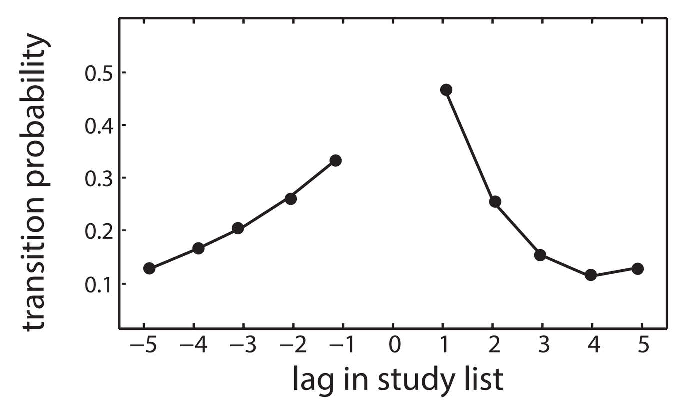
Two possibilities

What is mental context?

The semantic hypothesis: Mental context is a recency-weighted average of the meanings of our thoughts (Howard & Kahana 2002)
=> Recall transitions are more likely between items studied with similar preceding items

or

The random drift hypothesis: Mental context is a randomly-drifting signal (Estes 1955)
=> Recall transitions are more likely between items studied close in time



Approach

If we can show that semantics of prior items affect recall order, then we have provided evidence for the "semantic" hypothesis!

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

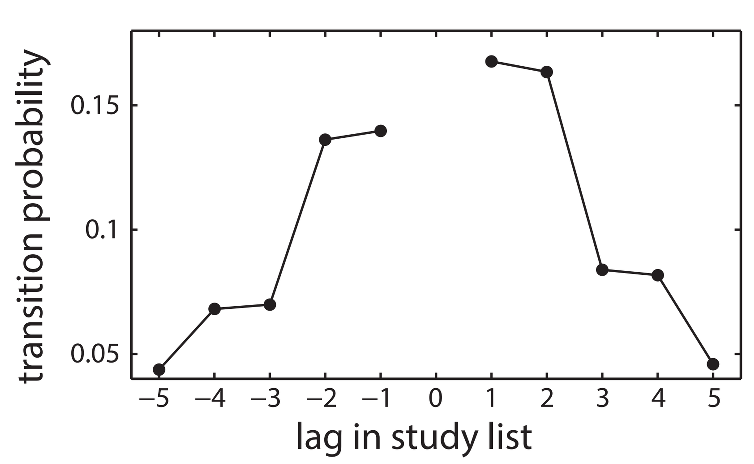
Recall transitions between landmarks with semantically similar preceding items...

...*over* temporally adjacent landmarks

3 Analysis I: Behavioral

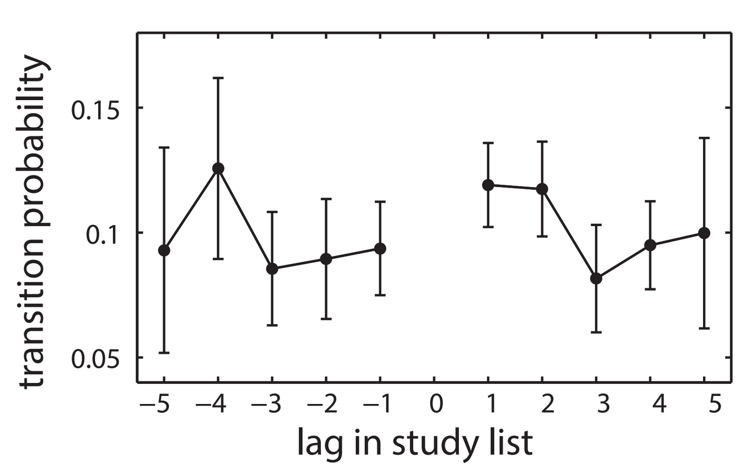
Prediction: recall transitions are biased towards Evel Knievels.

Predicted results



Superimposed on top of the temporal effects:
- Greater transition probability to the even lags

Results



+2 > +1 ?
+4 > +3 ?
-2 > -1 ?
-4 > -3 ?

p = 0.02
(n = 24)

4 Analysis II: fMRI decoding of semantics

Prediction

We think that the category of the preceding item forms the semantic content of mental context. Because these preceding items are processed in various amounts, there will be some naturally-occurring variance in the semantic similarity of context for Evel Knievel pairs.

Prediction: Evel Knievels are most likely to occur if the preceding items were both highly activated in the brain.

So, we use the following similarity score, and divide the transitions into two bins:

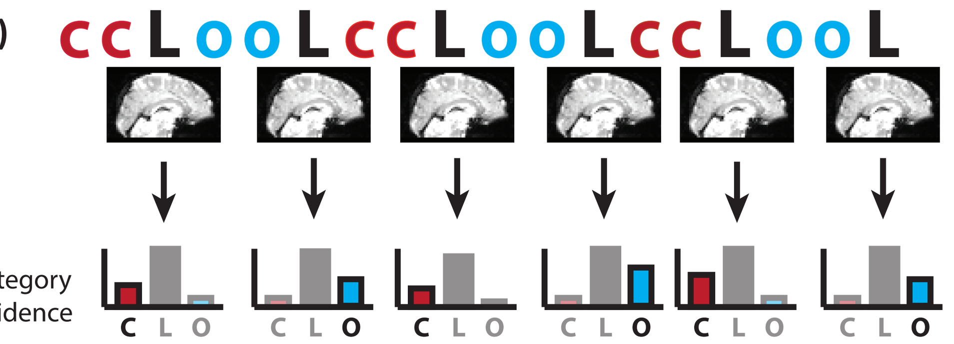
Similarity score = Activation of preceding category for 1st item x Activation of preceding category for 2nd item

		1st item		2nd item
high similarity	=	high	x	high
low similarity	=	low	x	high
		high	x	high
		low	x	low

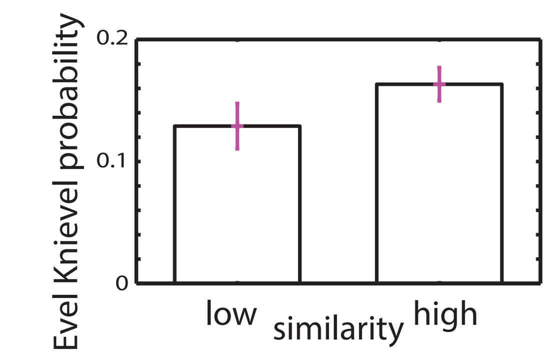
Method

Multivariate pattern analysis (MVPA)

We estimate the activation level of preceding items by applying a brain decoder to the fMRI data from study.



Results



fMRI activation of preceding items at study predicts transition likelihood at recall
(p = 0.06, n = 17)

5 Analysis III: Model-fitting

Motivation & Approach

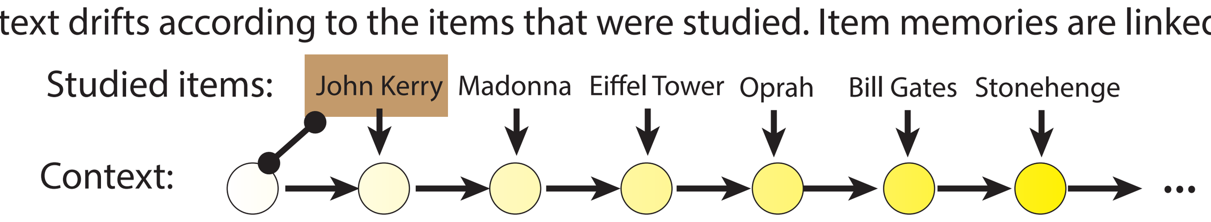
We need a principled way to discern contributions of semantic and temporal effects to recall order
=> a model!

Prediction: A context model that incorporates semantics will do better at predicting recall order than a model that does not.

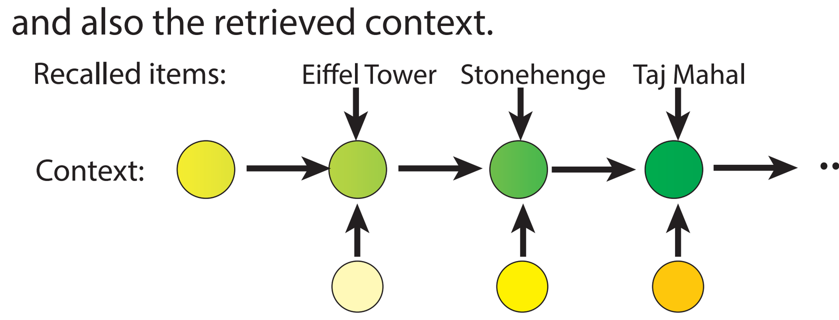
Model description

Overview
The model accepts studylists as input.
It estimates drift in mental context, and also **makes predictions of recall order.**

Context drift at study
Context drifts according to the items that were studied. Item memories are linked to this drifting context.



Context drift at recall
Drifts according to the items that were studied, and also the retrieved context.



Two kinds of context (at study and recall)

The ***semantic context*** drifts according to the categories of the item.

The ***temporal context*** treats each item as orthogonal, and therefore does not contain any semantic information.

Recall predictions

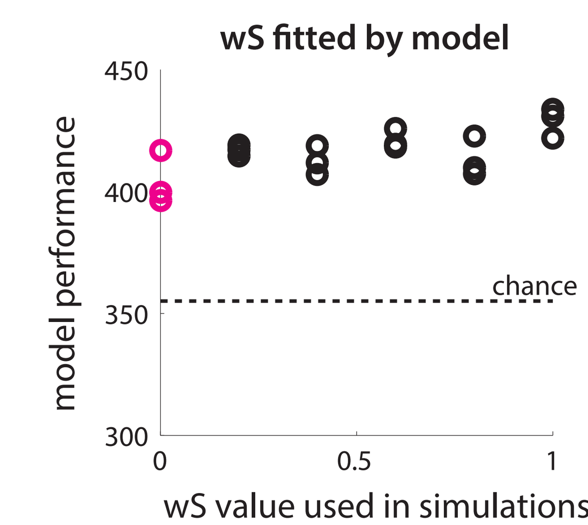
contextual affinity for item ***N*** = similarity (context at study of item ***N***, current context)

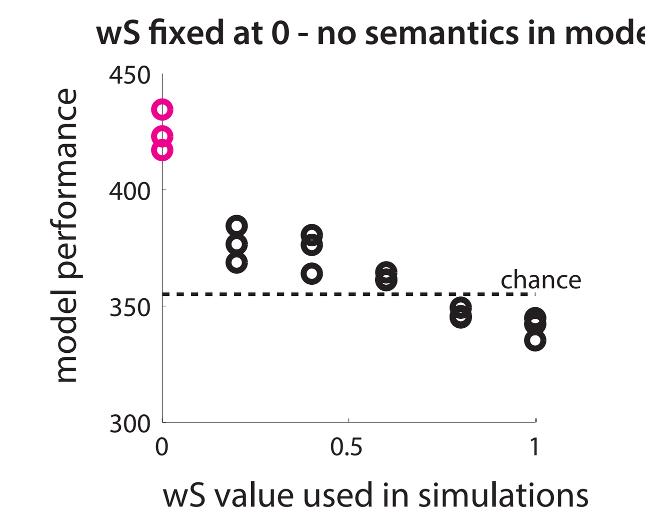
Probability of recall (item ***N***) = ***wS*** x (semantic contextual affinity) + (1 - ***wS***) x (temporal contextual affinity)

Results on simulated data

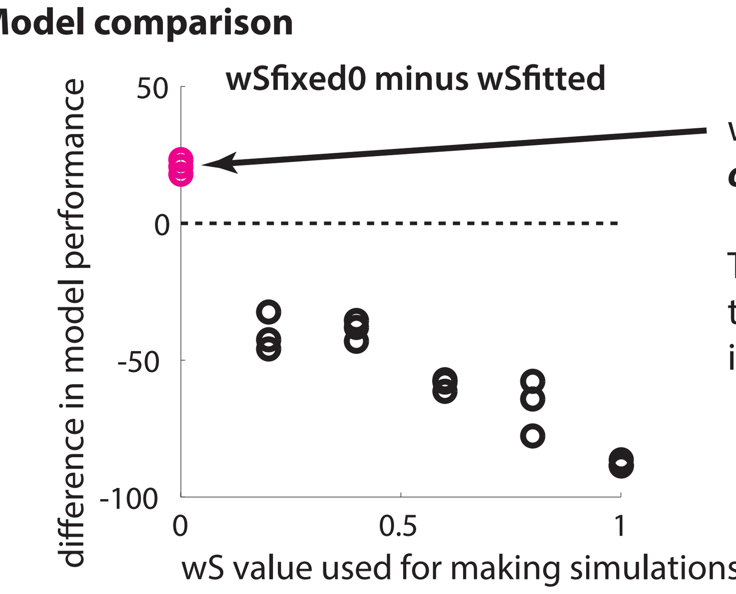
Simulations were made by varying the amount of semantics contribution to mental context, by varying ***wS*** from 0 (no semantics) to 1 (maximal semantics).

Model performance: How well does the model predict the simulated recalls?





Model comparison



wSfixed0 performs better than wSfitted ***only*** when wS is actually 0!

This indicates that we can use this measure to determine whether or wS in our real data is greater than 0.

6 Conclusions

Recall order seems to be affected by **the semantic content of the items studied immediately before** each recalled items. This is congruent with the theory that our memories our timestamped by the thoughts that are co-active at the time of the memory, and that we use these timestamps to help us retrieve our memories.

I - During recall, subjects were relatively more likely to transition between items preceded by the same semantic category (vs. items preceded by different categories).
II - During recall, subjects were more likely to transition between pairs of items preceded by more similar semantics, as estimated by an fMRI brain decoder.
III - Model-fitting will soon be applied to real data.

Acknowledgements

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

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.