Connectionist modeling of source memory phenomena

582.5

Sean M. Polyn, Kenneth A. Norman & Jonathan D. Cohen

Department of Psychology, Princeton University, Princeton, NJ, USA

Introduction

In recent years, several researchers have argued that the hippocampus routinely contributes to item recognition memory (Zola & Squire, 2001; Norman & O'Reilly, in press).

More specifically, hippocampus is thought to contribute by recollecting specific studied details (Norman & O'Reilly, in press; Yonelinas, 2002).

There is extensive evidence from behavioral studies that this (hippocampal) recollection process is context-sensitive (Macken, 2002; Mayes, 1988; Yonelinas, 2002; Godden & Baddeley, 1975; Smith, 1988).

If the hippocampus contributes to recognition, and it is context-sensitive, recognition should be context-sensitive. However, this result isn't always obtained (Smith, 1988).

This dilemma is made explicit in a study by Dodson & Shimamura (2000) that examined item recognition and source recall.

They showed, in the same memory test, a significant context change effect for source recall (presumably hippocampally-driven), and a null context change effect for item recognition.

We used the Complementary Learning Systems model (McClelland, McNaughton & O'Reilly, 1995; O'Reilly & Rudy, 2001; Norman & O'Reilly, in press) to address this puzzling state of affairs.

The CLS model consists of two parts: a medial temporal lobe cortex (MTLC) system which computes a familiarity signal and a hippocampal (HCMP) system which actively recalls detailed episodes; both systems contribute to recognition memory.

The model shows that it is possible to resolve this controversy without abandoning the idea that hippocampus routinely contributes to recognition memory.

The paradigm

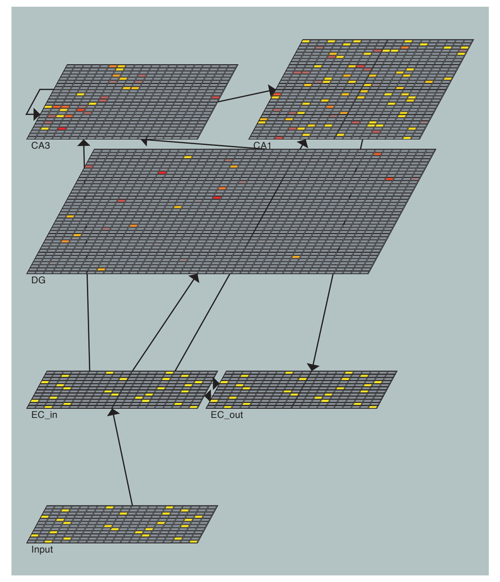
At study, subjects heard words spoken by one of two speakers, Eric and Julie.

At test, the same speakers presented old and new words. Subjects were asked whether they had heard the word before (item recognition); if they had, they were asked who said it originally (source recall).

Here, the word is the item and the speaker is the source. The speaker is part of the overall context for the word, and when a different speaker presents a word at test, that is a context change.

Words could be presented by the same speaker that said them during study (the matching source condition), by the other speaker (the mismatching source condition), or could simply be presented visually (the empty source condition).

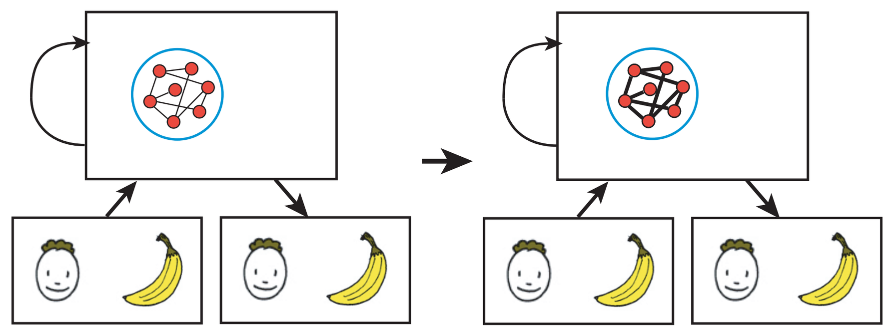
The hippocampal model



•Creates distinct representations
•Binds together features of the episode
•Supports pattern completion and detailed recollection
•Context-sensitive

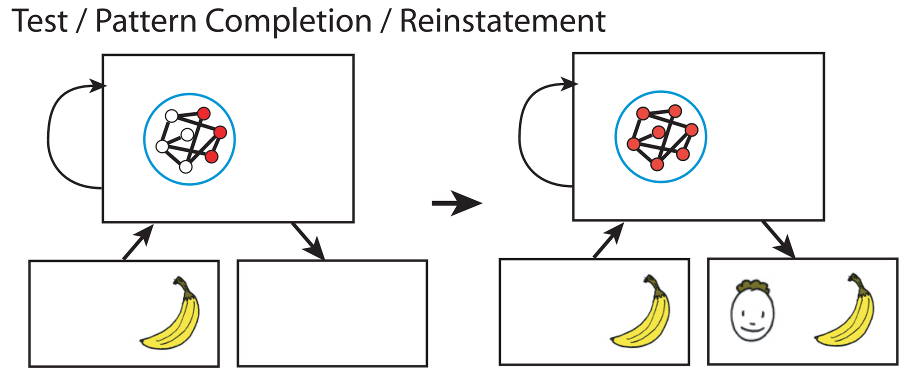
Overview of areas:
Input patterns are represented in EC-in. Pattern separation takes place in DG. Binding takes place in CA3. CA1 allows for reinstatement of learned patterns in EC-out.

Study / Binding



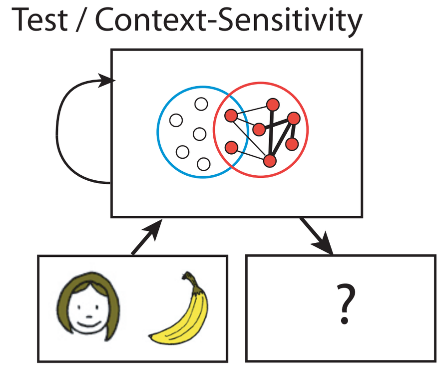
At study, Eric says the word 'banana'. This activates a set of units in CA3, the connections between which are strengthened with a Hebbian rule.

Test / Pattern Completion / Reinstatement



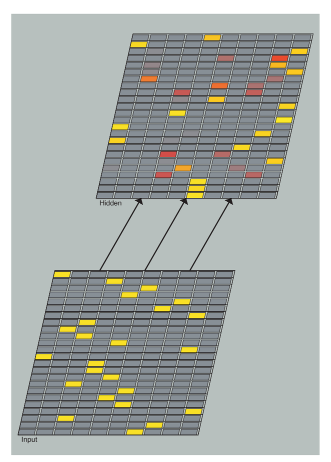
At test, the word 'banana' could be presented alone. This activates a subset of the original units. The strengthened weights between the units allows for pattern completion to the original attractor.

Test / Context-Sensitivity

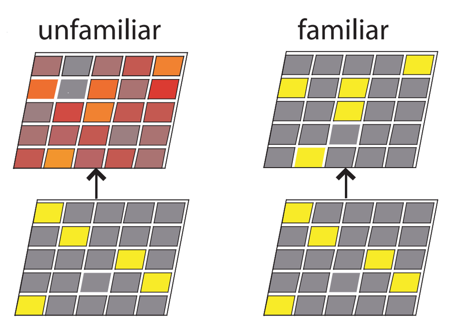


If at test, Julie says the word 'banana', then a different, but overlapping set of units is activated in CA3. To the extent that this is an imperfect memory cue for banana, the model will make more errors.

The MTLC model

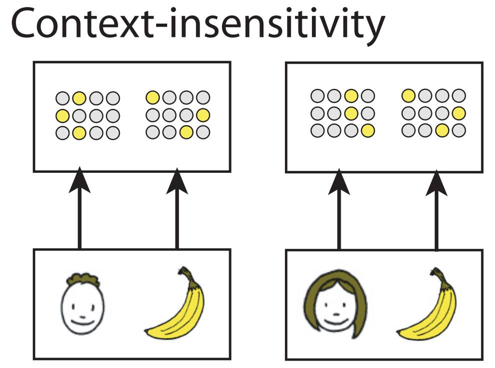


•Overlapping representations for similar events.
•Computes a scalar familiarity signal.
•Context insensitive.



Familiar stimuli have sharper MTLC representations than unfamiliar stimuli. Hebbian learning tunes some units to respond strongly to the stimulus, and these strongly responding units inhibit other units.

Context-insensitivity



Behavioral data (Vargha-Khadem et al, 1997) suggests that MTLC can only support association of items processed within a single cortical area. To capture this data point we introduced a channeled structure to our MTLC model, whereby item and source information was processed by separate pools of units. Here, MTLC provides a strong familiarity signal to both Eric-banana and Julie-banana, because both Eric and Julie are familiar faces by testing.

Simulation Methods

The input pattern is divided up into "item" features and "context" features; all items presented by a given voice get the same "context tag".

We vary the amount of attention paid to item information by varying the proportion of the input pattern devoted to item versus context information.

We are attempting to understand why context change effects are only sometimes seen in recognition memory. We forced the MTLC system to be insensitive to context changes to a familiar context. Thus, we focus here on the hippocampal model to explain why context sensitivity sometimes is (and sometimes is not) obtained, and we do not present MTLC simulation results.

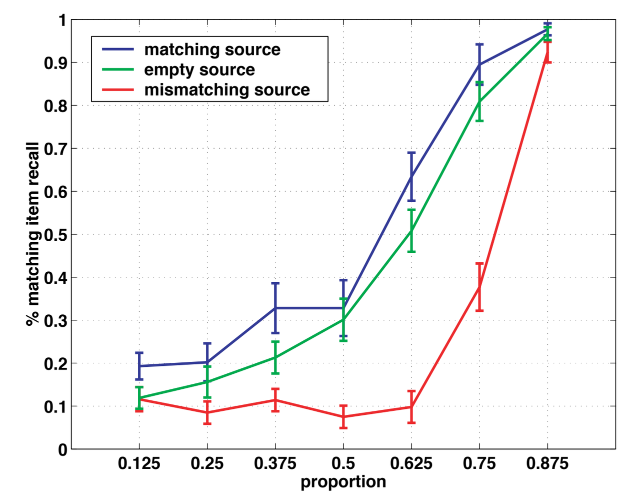
Applying the hippocampal model to item recognition. At test, the item pattern retrieved in EC-out is compared to the pattern in EC-in. The number of mismatching details is subtracted from the number of matching details. If this number exceeds a threshold, the item is reported as "old", otherwise the item is reported as "new".

Applying the hippocampal model to source recall. Retrieved details in the source channel are compared to a template for each source, using the same match minus mismatch calculation described above; the source that receives the larger score is the response given (in case of a tie, a guess is made).

Simulation results

First we examine the performance of the model on item recognition and source recall, while varying the strength of item information.

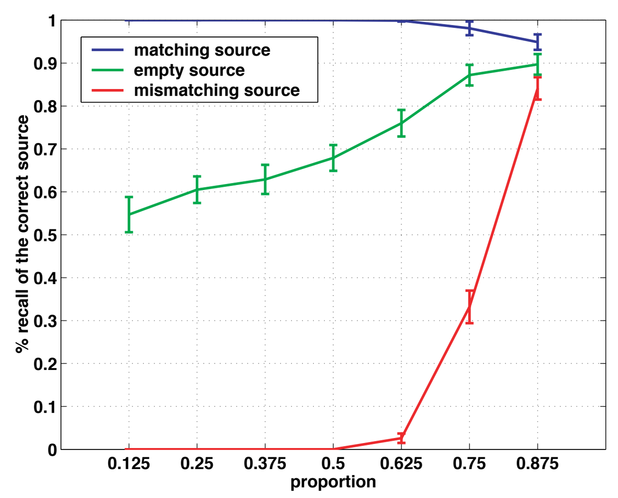
The percent of correct item recall for studied items as amount of item info is varied



← less item info   more item info →

When item and source information are balanced in the input layer (the center of the graph) there is a clear context change effect (blue vs red). However, on both the left and right-hand sides, the context effect goes away. (note: matching recall of item information is at floor for lures in the hippocampal model)

The percent of correct source recall as amount of item info is varied



Source recall shows a different behavior. There is a large context change effect for most of the range of item strength; the effect decreases when context information is weak in the input layer.

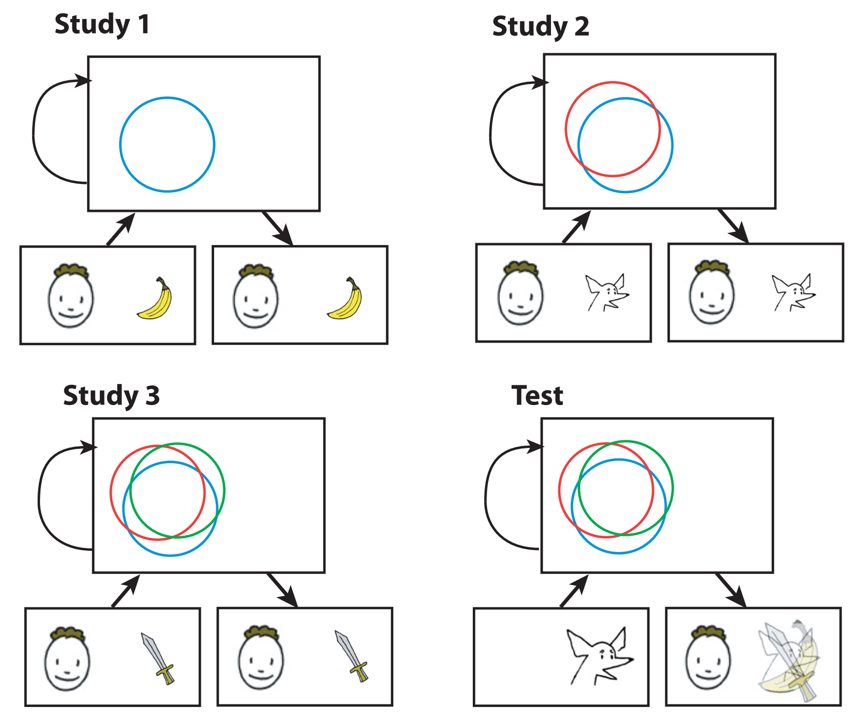
•Item recognition and source recall show differential sensitivity to changes in context, depending on the relative strength of the two types of information.

When does hippocampus show a null context change effect?

The model predicts that there are two states in which the hippocampus will show null context-change effects for item recall.

•When item information is strong, the null context change effect is due to source information being ignored.
•When item information is weak, there is a more interesting failure state: Item recognition is at floor, but source recall is not.

What is happening in this "weak item information" failure state?

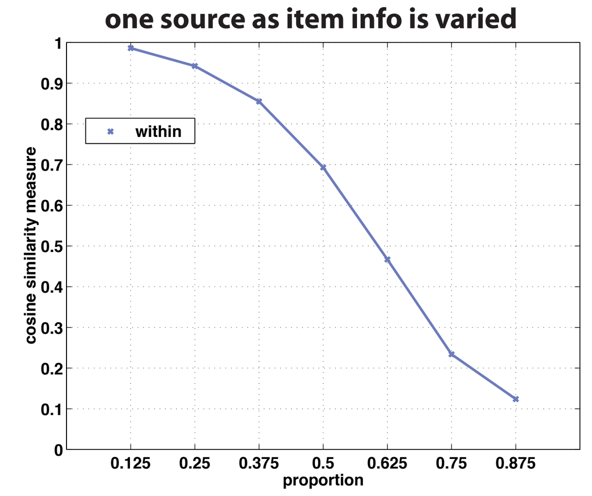


During the study phase of the experiment, Eric presents a number of words. Subjects are instructed to attend to Eric's voice, and are unaware of an impending memory test. This is modeled by devoting a small proportion of input units to item information. Since the subject is focusing on the information common to the set of items, the items receive overlapping representations in CA3. Each item representation has very few units that distinguish it from the others.

This causes the creation of an attractor in CA3 which is associated with many items but only one source. At test, an old item is presented. The hippocampal model is unable to retrieve meaningful item information, but is able to retrieve the appropriate source information.

We now examine the similarity between attractors created in the CA3 portion of the model, as a function of proportion of item information in the input layer.

Average cosine similarity of all items associated with one source as item info is varied

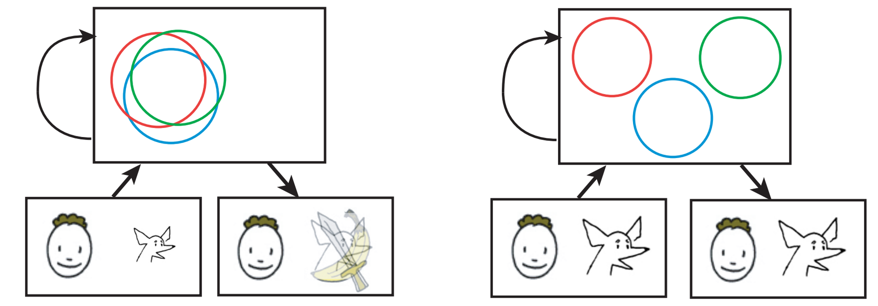


The blue line is the average similarity of all CA3 representations associated with one of the sources - as the relative amount of item information increases, the representations get less similar.

In our hypothesized failure state, the model is towards the left of this curve. The reason item information cannot be retrieved is that the CA3 representations for the items are too similar!

So, hippocampus is not retrieving item information - subjects must rely on the context-insensitive MTLC to perform item recognition. Then, they turn to the context-sensitive hippocampus to perform source recall.

Obtaining a context change effect



The model makes a prediction about how to get context change effects in the Dodson and Shimamura (2000) paradigm for item recognition. We believe that an item recognition context change effect will be found if subjects can be made to attend more to item information.

Focusing on distinct item information should force a balance of item and context information in the hippocampus. This will cause items to be assigned distinct CA3 representations, putting the system in a state that robust context change effects should be seen.

Conclusions

The CLS model shows that there can be a situation in which the hippocampus recalls source information but not item information.

This explains the puzzling finding of Dodson & Shimamura (2000) in which a null context change effect was found for item recognition but a significant context change effect was found for source recall.

In this case, item recognition responses are based on the familiarity signal of context-insensitive MTLC, and source recall responses are based on context-sensitive hippocampus.

Some theories of memory suggest that an effortful association of item and context information is necessary to get context change effects (Murnane, Phelps & Malmberg, 1999). Our research suggests that hippocampus performs its binding operation automatically, and that one only needs to balance attention to item and context information to ensure a context change effect.

Future research will determine whether the predictions of the model appropriately explain the behavioral data.

Future Directions

This research stands as part of a broader endeavor, that of using a single computational model with fixed underlying parameters to account for both behavioral and neuroscientific memory data. The model has already been used to address puzzles in the cognitive memory literature (e.g., when are list strength interference effects obtained) and in the neuroscientific memory literature (e.g., when does hippocampal damage impair recognition; see Norman & O'Reilly, in press).

One question that we have not addressed thus far is how prefrontal cortex (PFC) contributes to episodic memory. Prefrontal cortex has been implicated in source memory, free recall, and more generally, in episodic memory tasks that involve some kind of strategic processing (Shimamura, 2002). One could argue that, by insuring that the cues fed into the model are well-specified, we are including an "implicit" PFC.

We are presently working on a version of the model that includes an explicit PFC component. Several researchers have argued that PFC implements an internal "temporal context" representation that allows us to mentally jump back in time to the target episode (Tulving, 2002; Schacter, 1987). We are currently exploring how computational mechanisms of PFC such as gating and maintenance (Rougier & O'Reilly, 2002) give rise to this context vector, and how PFC context interacts with the medial temporal memory system.

References

Dodson & Shimamura (2000) Differential effects of cue dependency on item and source memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 26(4) 1023-1044.

Godden DR & Baddeley AD (1975) Context-dependent memory in two natural environments: on land and underwater. *British Journal of Psychology*, 66, 325-332.

Mayes AR (1988) Amnesia and memory for contextual information. In GM Davies & DM Thomson (Eds.), *Memory in Context: Context in Memory* (pp. 193-213). New York: Wiley.

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. Psychological Review, 102(3), 419-457.

Murnane K, Phelps MP & Malmberg K (1999) Context-dependent recognition memory: the ICE theory. *Journal of Experimental Psychology: General*, 128(4), 403-415.

Norman KA & O'Reilly RC. Modeling hippocampal and neocortical contributions to recognition memory: a complementary learning systems approach. *Psychological Review*, in press.

O'Reilly RC & Rudy JW (2001) Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychological Review*, 108(2), 311-345.

Rougier NP & O'Reilly RC (2002) Learning representations in a gated prefrontal cortex model of dynamic task switching. *Cognitive Science*, 26(4): 503-520.

Schacter DL (1987) Memory, amnesia, and frontal lobe dysfunction. *Psychobiology*, 15(1), 21-36.

Shimamura AP (2002) Memory retrieval and executive control processes. In DT Stuss & RT Knight (Eds.), *Principles of Frontal Lobe Function* (ch. 13). Oxford University Press: Oxford.

Smith SM (1988) Environmental context-dependent memory. In GM Davies & DM Thomson (Eds.), *Memory in Context: Context in Memory* (pp. 13-34). New York: Wiley.

Tulving E (2002) Chronesthesia: conscious awareness of structured time. In DT Stuss & RT Knight (Eds.), *Principles of Frontal Lobe Function* (ch. 20). Oxford University Press: Oxford.

Vargha-Khadem F, Gadian DG, Watkins KE, Connelly A, Van Paesschen W, Mishkin M (1997) Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277, 376-380.

Yonelinas AP (2002) The nature of recollection and familiarity: a review of 30 years of research. *Journal of Memory and Language*, 46, 441-517.

Zola SM & Squire LR (2001) Relationship between magnitude of damage to the hippocampus and impaired recognition memory in monkeys. *Hippocampus*, 11(2):92-98.