Complementary learning systems within the hippocampus: A neural network modeling approach to reconciling episodic memory with statistical learning

Anna C. Schapiro[1,2], Nicholas B. Turk-Browne[1], Matthew M. Botvinick[3], & Kenneth A. Norman[1]

[1]Dept. of Psychology / Princeton Neuroscience Institute, Princeton University; [2]The Center for Sleep and Cognition, Beth Israel Deaconess Medical Center / Harvard Medical School; [3]Google DeepMind

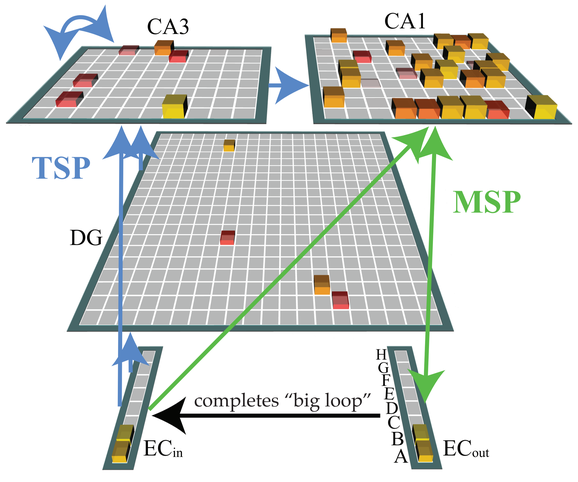
Introduction

• The hippocampus is critical for encoding individual episodes, using separate neural populations to avoid interference, while the cortex specializes in slowly extracting statistics across experiences, using overlapping representations to promote generalization (McClelland, McNaughton, & O'Reilly, 1995, *Psychol Rev*).

• But how do we extract statistics *rapidly*? Evidence is building that the hippocampus is responsible for this as well (Schapiro & Turk-Browne, 2015, *Brain Mapping: An Encyclopedic Reference*).

• How can one system provide both these functions?

Model architecture



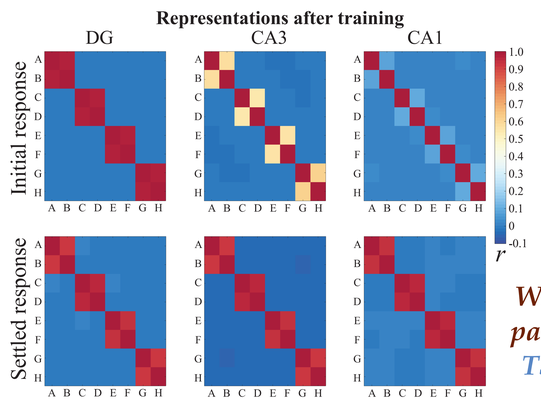
• **Trisynaptic pathway (TSP):** fast learning rate, sparse representations

• **Monosynaptic pathway (MSP):** slower learning rate, overlapping representations

Ketz, Morkonda, & O'Reilly, 2013, *PLOS Comput Biol*

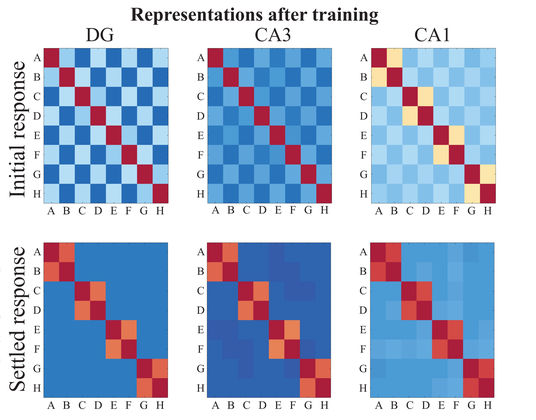
Pair structure: episodes vs. statistics

A B G H C D E F G H A B C D G H E F

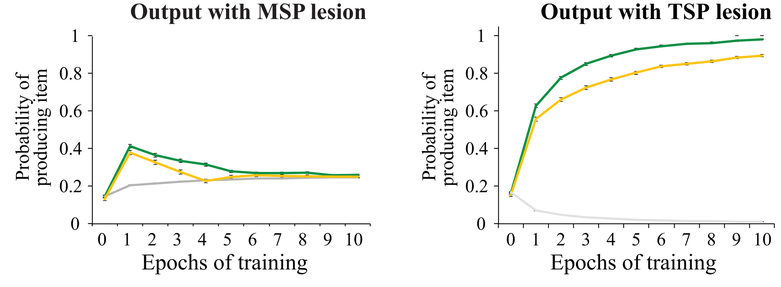


With episode demarcations, pairs represented strongly in TSP but not MSP in initial response

A B G H C D E F G H A B C D G H E F

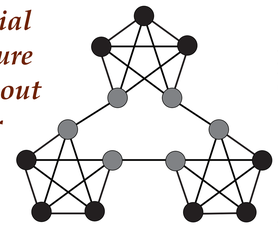


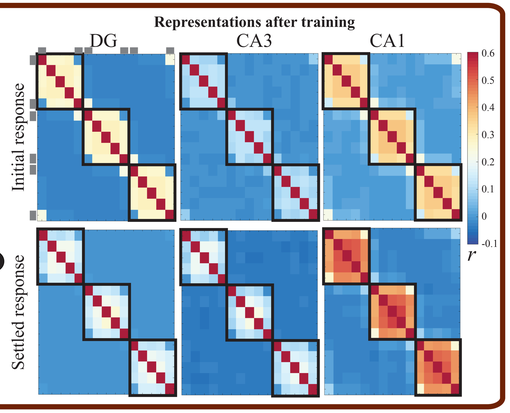
When statistical learning required, pairs represented strongly in MSP but not TSP in initial response



Higher level structure

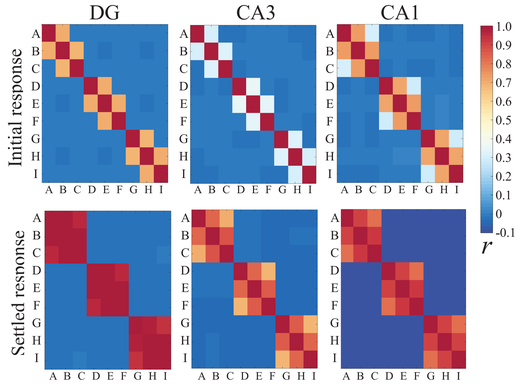
MSP is sensitive to higher level structure in initial response; structure stronger throughout network after big-loop recurrence





Associative inference

AB BC ... AC?



MSP uses static representational similarity to link A and C; TSP can also do linking via recurrence

(see also Kumaran & McClelland, 2012, *Psychol Rev*)

Predictions

• Over the course of a trial, representations of associated items should appear *first* in CA1, then EC, then DG/CA3.

• Lesions to the TSP should not affect representational structure in MSP or statistical learning behavior, assuming enough time for MSP learning. (This may explain why infants, who have undeveloped TSPs, are prodigious statistical learners.)

• Lesions to the MSP should leave intact some ability to make transitive associations.

Conclusions

The model suggests there may be complementary learning systems within the hippocampus itself, the TSP and MSP, allowing it to accomplish both episodic memory and statistical learning.

Paper in press at *Philosophical Transactions of the Royal Society B.* Contact: annacschapiro@gmail.com. Funded by NIH R01EY021755 (N.B.T.-B.); NIH R01MH069456 (K.A.N.); John Templeton Foundation (all authors).