1

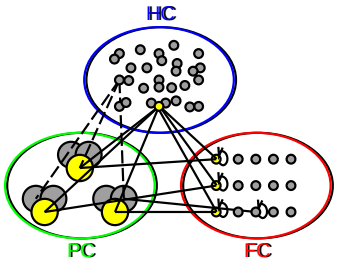
Prefrontal-Hippocampal Interactions: A Computational Perspective

Randall C. O'Reilly

Department of Psychology
Center for Neuroscience
Institute of Cognitive Science
University of Colorado Boulder

2

A Tripartite Cognitive Architecture



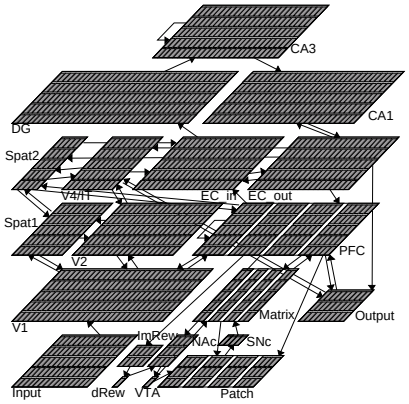
PC = Posterior cortex: graded, overlapping, slow learning of statistics in *weights* (*semantic*).
HC = Hippocampus: sparse (separated, conjunctive), rapid automatic learning in *wts* (*episodic*).
FC = Frontal cortex: robust maintenance, adaptive gating (BG), top-down control in isolated *acts* (*WM*).

Each area requires specializations for optimal performance of complementary set of functions (*avoiding tradeoffs*).

But same underlying mechanisms: neurons w/ synapses, etc just different *parameters* (continuum of function).

3

A Tripartite Cognitive Architecture



Under construction; stick to principles for now..

4

Broad Points

- Strong claim: it just has to be this way (computational necessity).
- Stronger claim: it is this way — lots of consistent data.
- Challenge: disprove it! (what are the difficult data?)

5

Specific Points

- *Automatic & rapid* learning of conjunctions is key for dissociating HC from PC.
- Organization of PFC may be according to abstraction & temporal duration of maintenance.
- *Distractors* & *proactive interference* are key for dissociating FC from HC.

6

Principled Distinctions from Computational Tradeoffs

McClelland, McNaughton & O'Reilly, 1995; O'Reilly & Rudy, 2001; Norman & O'Reilly, 2003

Goals:	Remember Specifics	Extract Generalities
Example:	Where is car parked?	Best parking strategy?

7

Principled Distinctions from Computational Tradeoffs

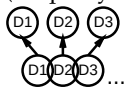
McClelland, McNaughton & O'Reilly, 1995; O'Reilly & Rudy, 2001; Norman & O'Reilly, 2003

Goals:	Remember Specifics	Extract Generalities
Example:	Where is car parked?	Best parking strategy?
Need to:	Avoid interference	Accumulate experience

8

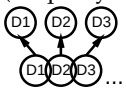
Principled Distinctions from Computational Tradeoffs

McClelland, McNaughton & O'Reilly, 1995; O'Reilly & Rudy, 2001; Norman & O'Reilly, 2003

Goals:	Remember Specifics	Extract Generalities
Example:	Where is car parked?	Best parking strategy?
Need to:	Avoid interference	Accumulate experience
	Solution:	
1.	Separate reps (keep days separate) 	Overlapping reps (integrate over days) 

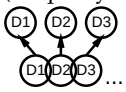
9 Principled Distinctions from Computational Tradeoffs

McClelland, McNaughton & O'Reilly, 1995; O'Reilly & Rudy, 2001; Norman & O'Reilly, 2003

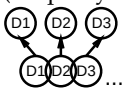
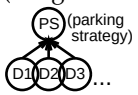
Goals:	Remember Specifics	Extract Generalities
Example:	Where is car parked?	Best parking strategy?
Need to:	Avoid interference	Accumulate experience
	Solution:	
1.	Separate reps (keep days separate) 	Overlapping reps (integrate over days) 
2.	Fast learning (encode immediately)	Slow learning (integrate over days)

10 Principled Distinctions from Computational Tradeoffs

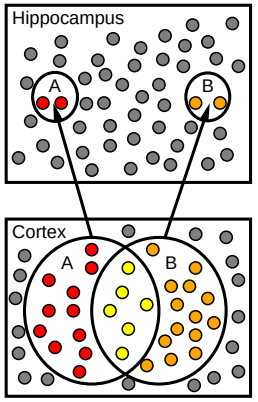
McClelland, McNaughton & O'Reilly, 1995; O'Reilly & Rudy, 2001; Norman & O'Reilly, 2003

Goals:	Remember Specifics	Extract Generalities
Example:	Where is car parked?	Best parking strategy?
Need to:	Avoid interference	Accumulate experience
	Solution:	
1.	Separate reps (keep days separate) 	Overlapping reps (integrate over days) 
2.	Fast learning (encode immediately)	Slow learning (integrate over days)
3.	Learn automatically (encode everything)	Task-driven learning (extract relevant stuff)

11 Principled Distinctions from Computational Tradeoffs

Goals:	Remember Specifics	Extract Generalities
Example:	Where is car parked?	Best parking strategy?
Need to:	Avoid interference	Accumulate experience
	Solution:	
1.	Separate reps (keep days separate) 	Overlapping reps (integrate over days) 
2.	Fast learning (encode immediately)	Slow learning (integrate over days)
3.	Learn automatically (encode everything)	Task-driven learning (extract relevant stuff)
	These are incompatible, need two different systems:	
System:	Hippocampus	Neocortex

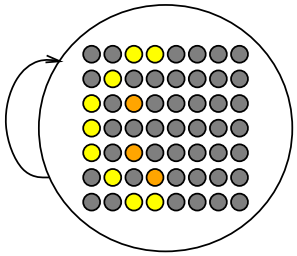
12 Hippocampal Sparseness → Separation, Conjunctions



(and makes fMRI of hippocampus so unreliable!?)

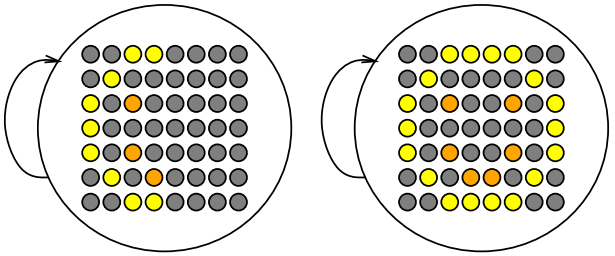
13

Recall: Pattern Completion



14

Recall: Pattern Completion



Hippocampal conjunctive representations support *pattern completion* — partial cues retrieve whole memory.

Need to bind elements together for cue to reactivate whole.

15

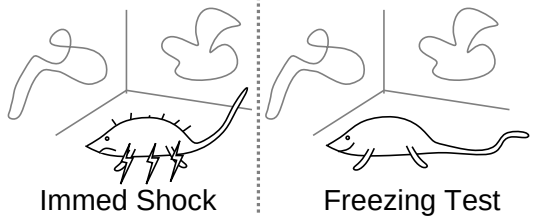
Testing the Theory: Rapid Incidental Conjunctive Memory in Rats

Fear conditioning preexposure paradigm:

- Uniquely illustrates all 3 properties of hippocampal memories.
- Provides the best animal model of human episodic memory!

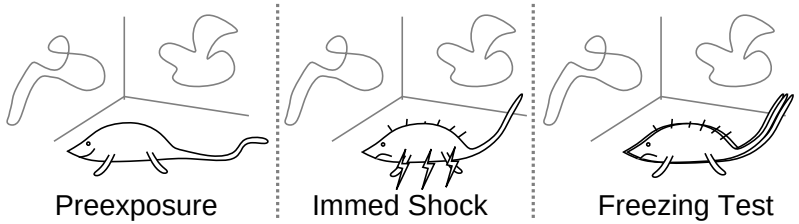
16

Immediate Shock Effect (Fanselow)



Shocking immediately in context produces no freezing at test!

17 Eliminating Immediate Shock Effect



Preexposure to context eliminates immediate shock effect.

Interpretation:

- Preexposure: incidental binding of features in hippocampus.
- Shock: conjunctive rep. is pattern completed, assoc w/ shock.
- Test: shock/fear is pattern completed from cues at test.

18 Conditioning to a Memory

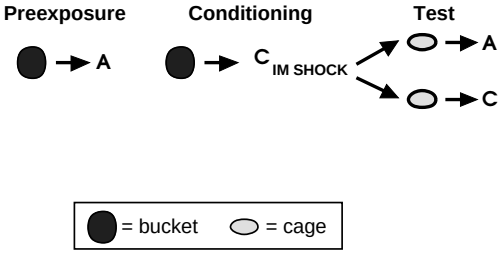
Can we trigger pattern completion to activate a memory of a context, *even when that context isn't physically present?*

1. At Shock: The hippocampus can pattern complete a memory of a context that is not present, given a retrieval cue.

2. Then, fear conditioning can occur *to that memory*, instead of to the context that is actually present.

This would provide strong evidence for hippocampal "episodic" recall in animals.

19 Conditioning to a Memory: Experimental Paradigm

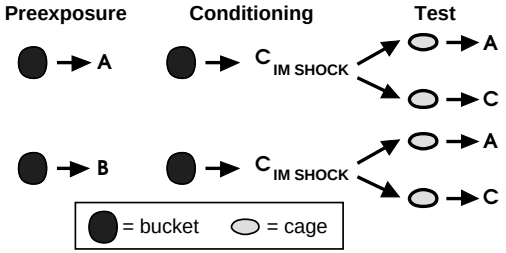


Preexposure leads to conjunction of bucket and context A.

Bucket can trigger pattern completion of A at shock (in C).

Intact rat should associate shock with memory of A *not* C!

20 Conditioning to a Memory: Experimental Paradigm

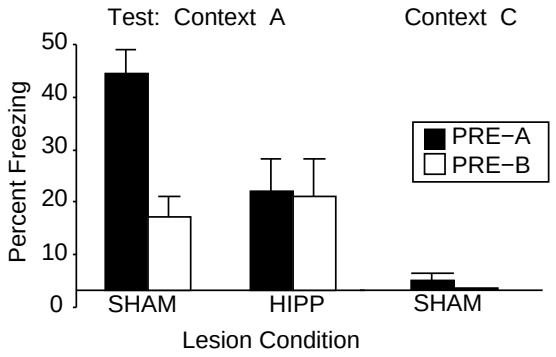


Alternate preexposure environment B serves as a control.

21

Conditioning to a Memory: Rat Data

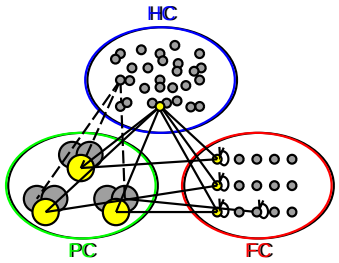
Rudy, Barrientos & O'Reilly, 2002



Note: Standard contextual fear conditioning not HC dependent!

22

A Tripartite Cognitive Architecture: HC vs. PC



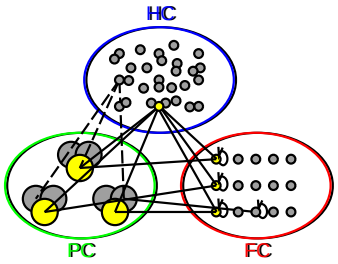
PC = Posterior cortex: graded, overlapping, slow learning of statistics in *weights* (*semantic*).
HC = Hippocampus: sparse (separated, conjunctive), rapid automatic learning in *wts* (*episodic*).
FC = Frontal cortex: robust maintenance, adaptive gating (BG), top-down control in isolated *acts* (*WM*).

Rapid & Automatic (incidental): key for *dissociating* HC from PC (PC can more slowly learn conjunctive representations if needed).

Watch out for automatic encoding during retrieval! (Stark)

23

A Tripartite Cognitive Architecture: HC vs. PC



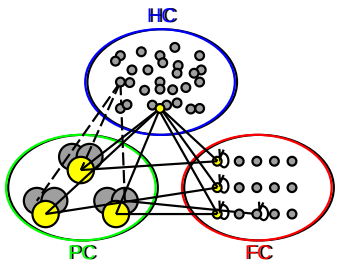
PC = Posterior cortex: graded, overlapping, slow learning of statistics in *weights* (*semantic*).
HC = Hippocampus: sparse (separated, conjunctive), rapid automatic learning in *wts* (*episodic*).
FC = Frontal cortex: robust maintenance, adaptive gating (BG), top-down control in isolated *acts* (*WM*).

Caveats: HC is also sensitive to task demands, attention, top-down input from PFC, etc (can be, but is not always, automatic).

PC can learn automatically (priming, familiarity) too:
but not *novel conjunctions*.

24

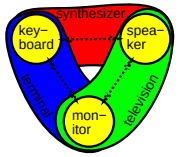
A Tripartite Cognitive Architecture: HC vs. PC



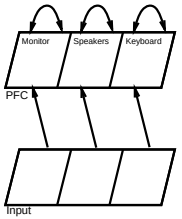
PC = Posterior cortex: graded, overlapping, slow learning of statistics in *weights* (*semantic*).
HC = Hippocampus: sparse (separated, conjunctive), rapid automatic learning in *wts* (*episodic*).
FC = Frontal cortex: robust maintenance, adaptive gating (BG), top-down control in isolated *acts* (*WM*).

25 Tradeoffs for Robust Maintenance in the PFC

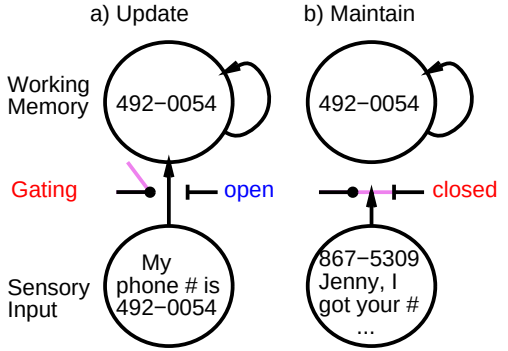
Tradeoff: overlapping, interconnected reps = spreading activation, inferences, *semantics*.
But spreading activation in *WM* w/out external input = loose memory



Solution = *isolated* representations in PFC: (e.g., stripes; Levitt et al, 1993)
& *intrinsic bistability* (up/down states)



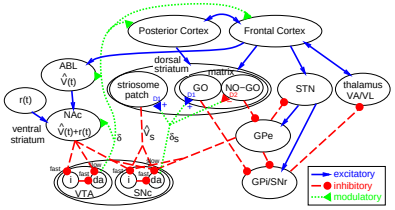
26 Tradeoffs for Robust Maintenance in the PFC

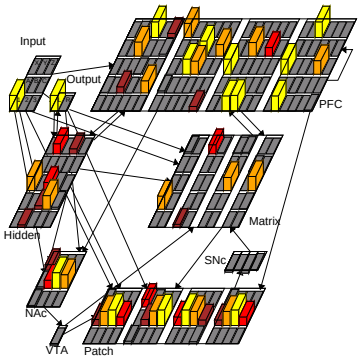


Adaptive gating required for having *both* rapid updating and robust maintenance (O'Reilly, Braver & Cohen, 1999).

27 The Basal Ganglia Support Adaptive Gating

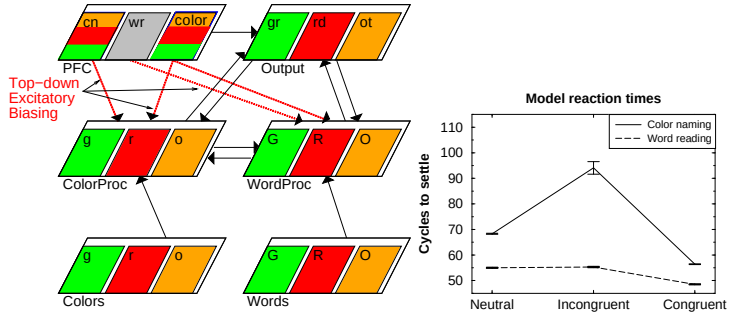
Frank, Loughry & O'Reilly, 2001; O'Reilly, submitted





28 Top-Down Excitatory Biasing: *Not* Inhibition

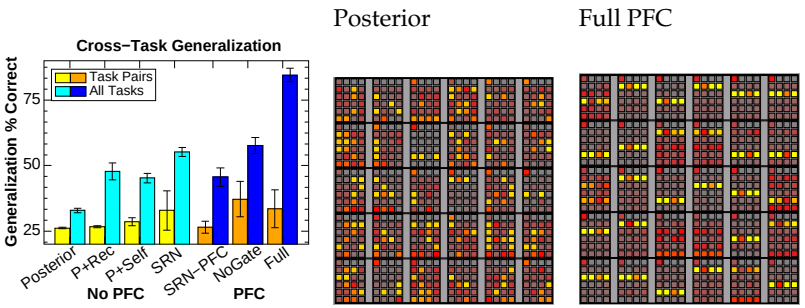
Cohen, Dunbar & McClelland, 1990; Herd, Banich & O'Reilly, submitted; c.f. Petrides, 1994



Robustly maintained PFC reps bias task-appropriate processing in posterior cortex, hippocampus (e.g., Stroop task).

29 PFC Specializations → Rule-Like Abstract Reps

Rougier, Noelle, Braver, Cohen & O'Reilly, in prep

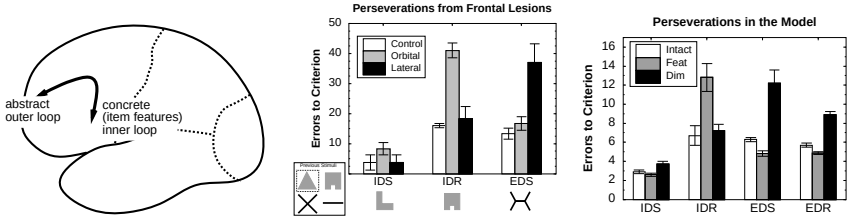


Rule = One stim dimension relevant at a time = one row.

Abstraction derives from sustained maintenance over trials!

30 Anterior-Posterior/Inferior Gradient of Abstraction

O'Reilly, Noelle, Braver & Cohen, 2002

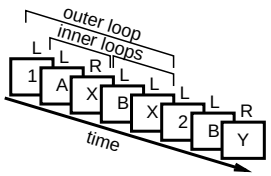


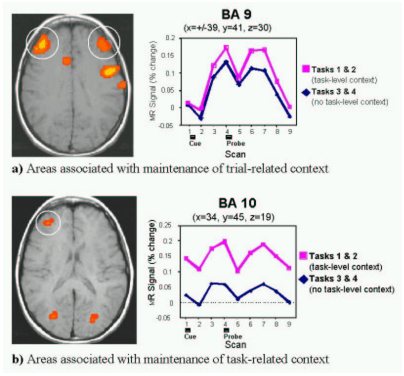
Intra-dim switching deficits = PFC feature reps lesion (inferior).

Extra-dimensional = PFC abstract dimensional reps lesion (dorsal).

31 Inner-Loop/Outer-Loop Gradient

O'Reilly, submitted; c.f., Koechhlin et al, 99; Christoff & Gabrieli, 00; Braver & Bongiolatti, 02





Target (R) =
A-X if 1, B-Y if 2

Requires independent gating of inner/outer loops (stripes, BG)

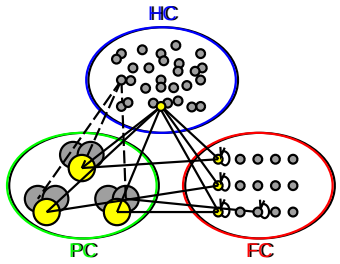
Outer-loop = longer maintenance = more abstract!

32 Application

Rugg:
Outer loop = correct/incorrect = abstract task
Inner loop = item-specific "monitoring" = concrete.

Petrides:
Outer loop = multiple item maintenance = dorsal
Inner loop = single item maintenance = ventral.

33 A Tripartite Cognitive Architecture



PC = Posterior cortex: graded, overlapping, slow learning of statistics in *weights* (*semantic*).
HC = Hippocampus: sparse (separated, conjunctive), rapid automatic learning in *wts* (*episodic*).
FC = Frontal cortex: robust maintenance, adaptive gating (BG), top-down control in isolated *acts* (*WM*).

Strong claim: it just has to be this way!

34 Interactions!

PFC top-down biasing can support encoding in hippocampus (lots of examples; nobody disagrees with this).

Hippocampal rapid, automatic encoding can support working memory-like short-term memory (Cohen & O'Reilly, 1996; O'Reilly, Braver & Cohen, 1999). (also non-controversial)

What are the challenging examples?

- Sustained WM-like activity in hippocampus.
- When can HC *not* support WM?

35 Sustained Activity in Hippocampus (Ranganath, Suzuki, Reber)

Strong constraint: Hippocampus does not have neural specializations required for *robust* maintenance.

But many posterior areas can exhibit "residual" activity, and "reflect" sustained input (top-down biasing) from PFC (consistent with fMRI data; test with lesions?)

→ Need *distractors* during delay period to test robust maintenance (ala Miller, Desimone, et al).

36 Neural Specializations in Entorhinal Cortex

- Intrinsic bistability (Egorov et al, 2002).
- Projections from basal ganglia?

Maybe EC has some of the specializations to do more robust active maintenance?

Useful for integrating stimulus information over an episode?

37 When Can HC *Not* Support WM?

Weight-based learning is subject to effects of *proactive interference* (even with hippocampal patterns separation).

So, unique vs. repeated use of stimuli within a task is a way to selectively disable HC contributions to WM function.

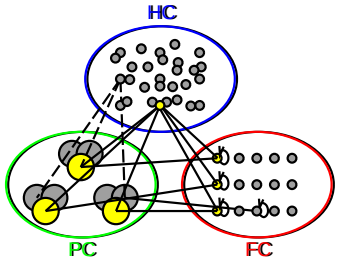
38 Interpretation of Reber

- Stimuli require novel conjunctive representations HC.
- PI builds up due to similarity of stimuli.
- Higher HC act on bad trials due to greater PI.

39 Conclusions: Specific Points

- *Automatic & rapid* learning of novel conjunctions is key for dissociating HC from PC.
- Organization of PFC may be according to abstraction & temporal duration of maintenance.
- *Distractors* & *proactive interference* are key for dissociating FC from HC.

40 A Tripartite Cognitive Architecture: Conclusions



PC = Posterior cortex: graded, overlapping, slow learning of statistics in *weights* (*semantic*).
HC = Hippocampus: sparse (separated, conjunctive), rapid automatic learning in *wts* (*episodic*).
FC = Frontal cortex: robust maintenance, adaptive gating (BG), top-down control in isolated *acts* (*WM*).

Computational models of entire system are just being developed.

Should provide more subtle predictions to test!