An Introduction to Learning

Lecture 9

Todd M. Gureckis, Lila Davachi
Department of Psychology
New York University
{todd.gureckis,lila.davachi}@nyu.edu

Agenda for Today

- The medial temporal lobe and memory..
  - Distinct forms of memory subserved by regions in the medial temporal lobe
  - Recall and familiarity memory reliant on different encoding and retrieval computations
  - Evidence for pattern separation in hippocampal subregions

The Medial Temporal Lobe and Memory

- review of relevant anatomy (how hippo connects to cortex)

- how the hippocampus works

- effects of large MTL lesions

- effects of more focal hippocampal lesions

- roles of particular hippocampal substructures

The Medial Temporal Lobe and Memory

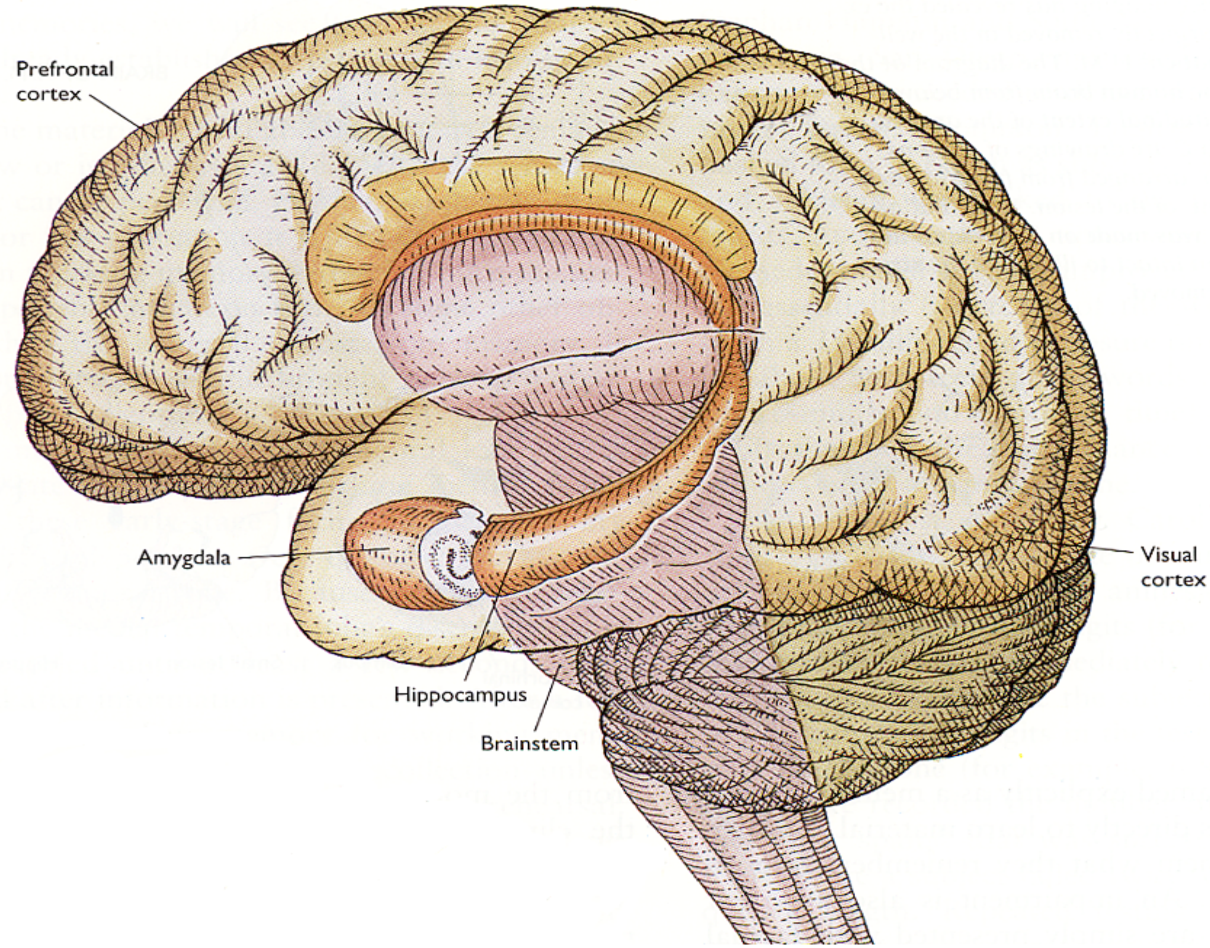
- review of relevant anatomy (how hippo connects to cortex)

- how the hippocampus works

- effects of large MTL lesions

- effects of more focal hippocampal lesions

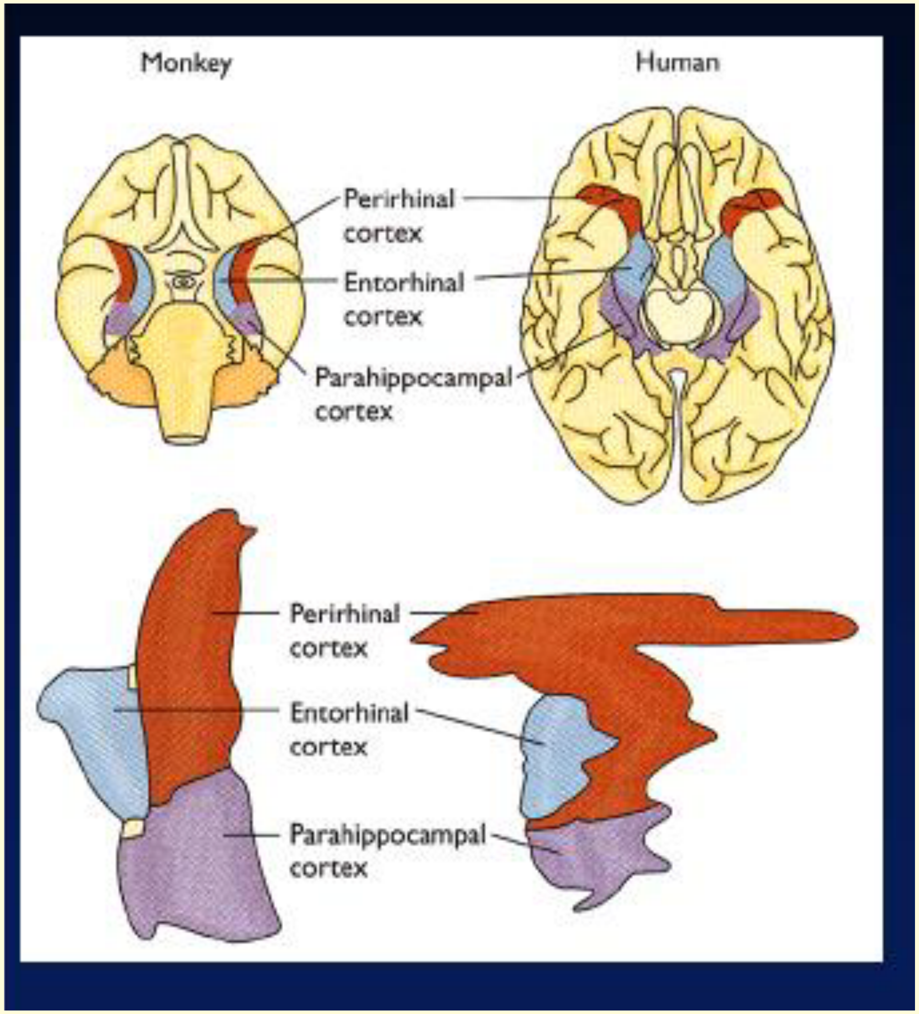
- roles of particular hippocampal substructures

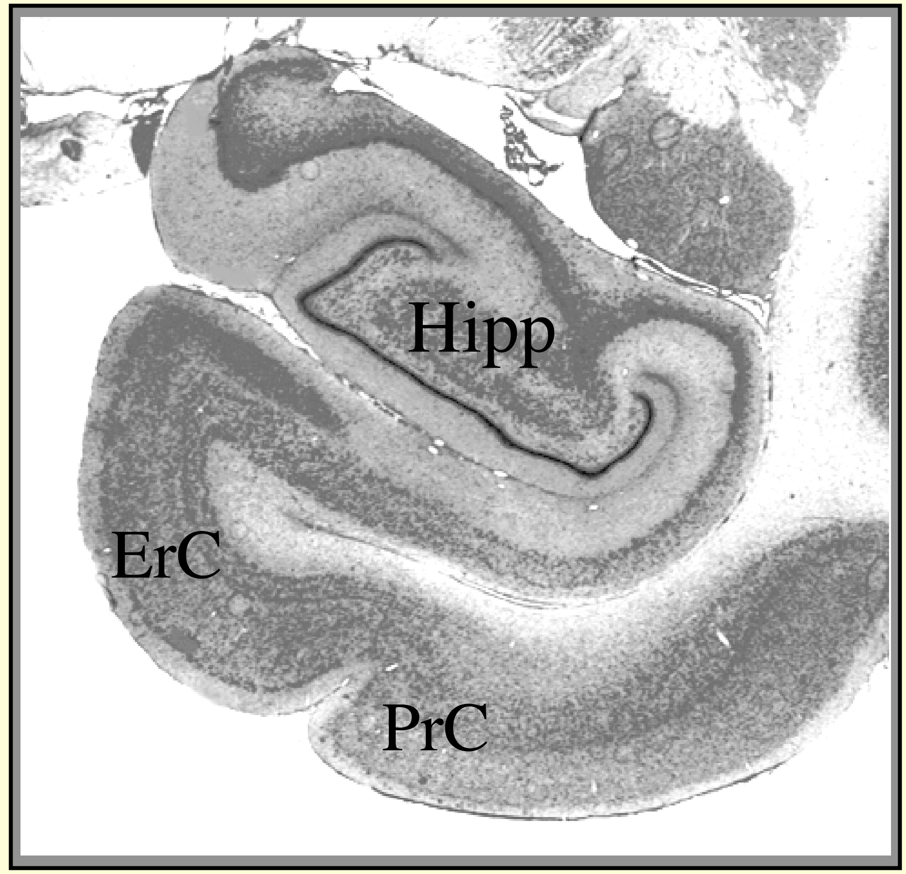


from Squire & Kandel (1999)

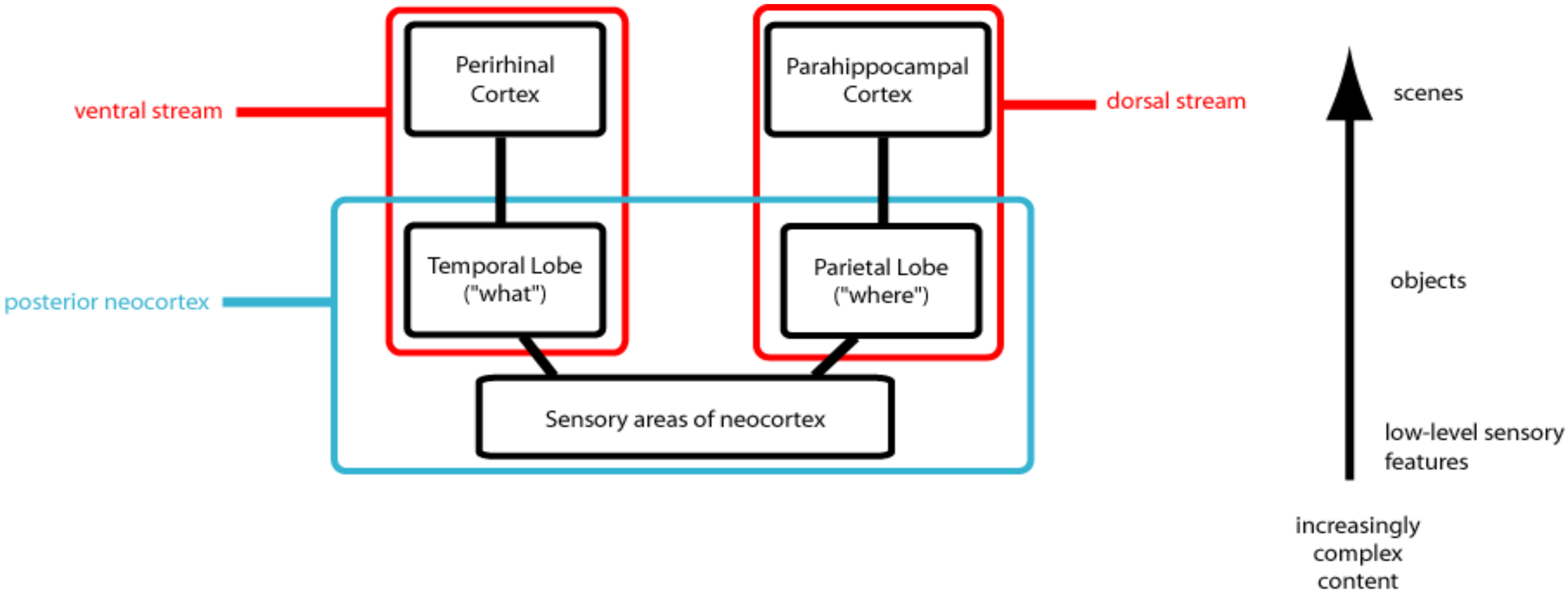
from DeArmond,
Fusco, & Dewey (1989)

Medial Temporal-lobe Memory System

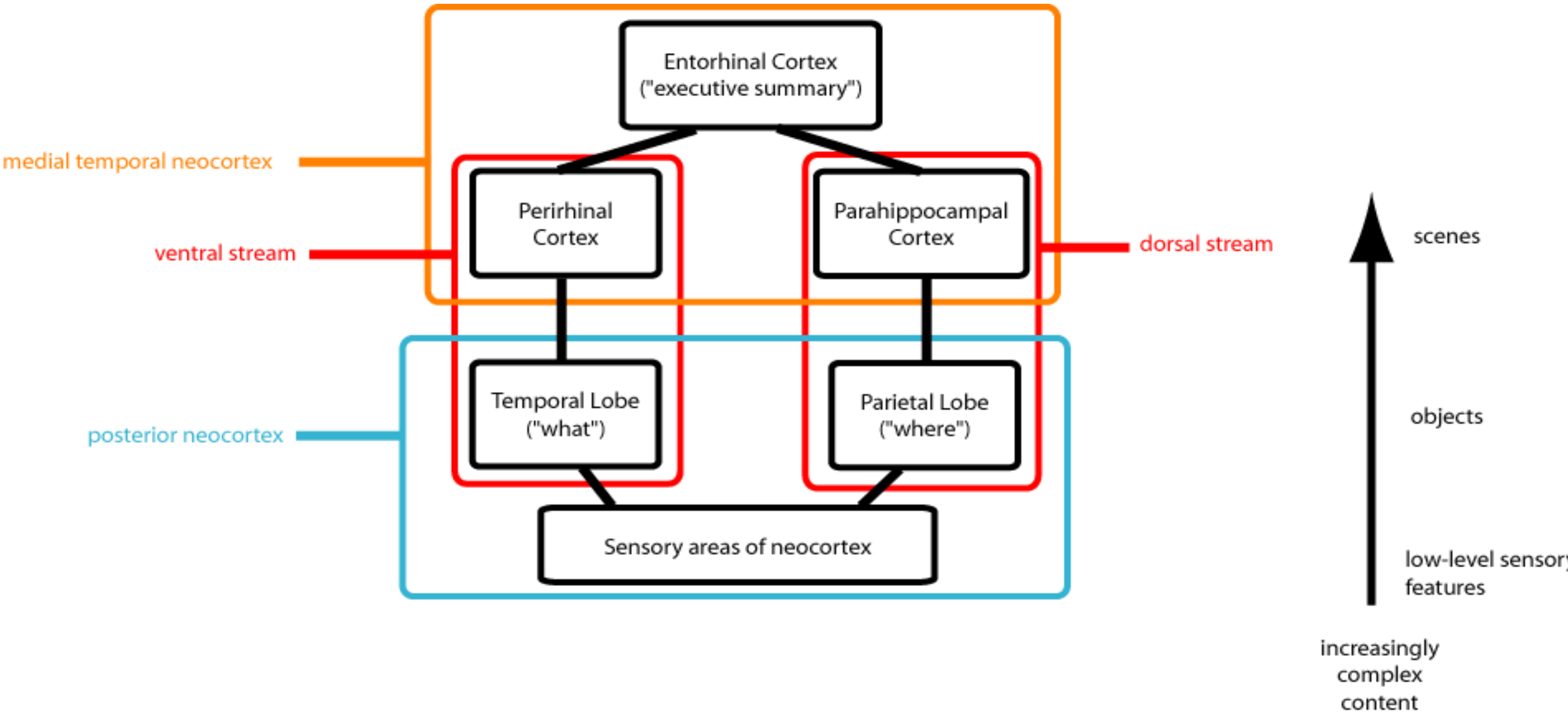


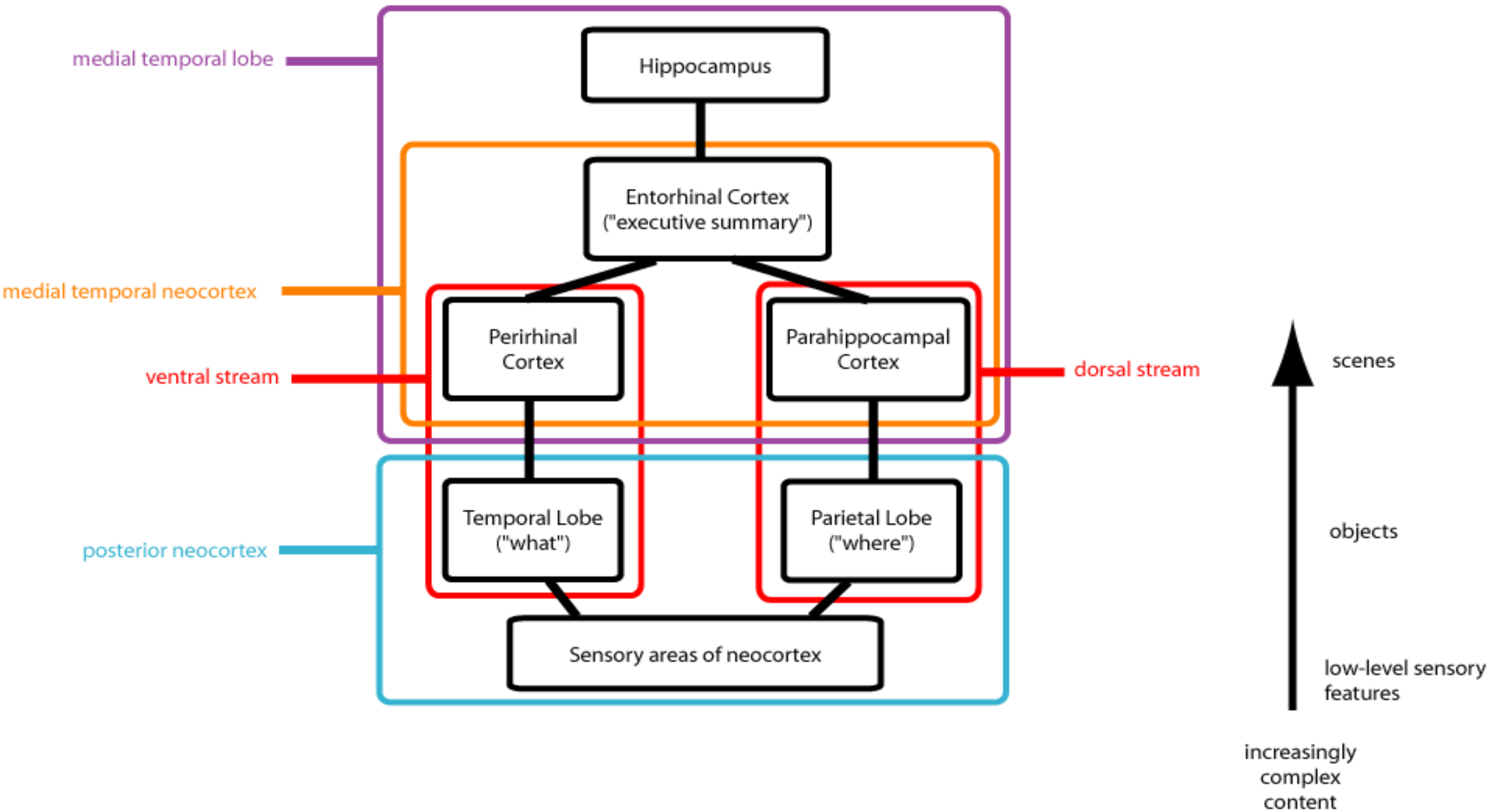


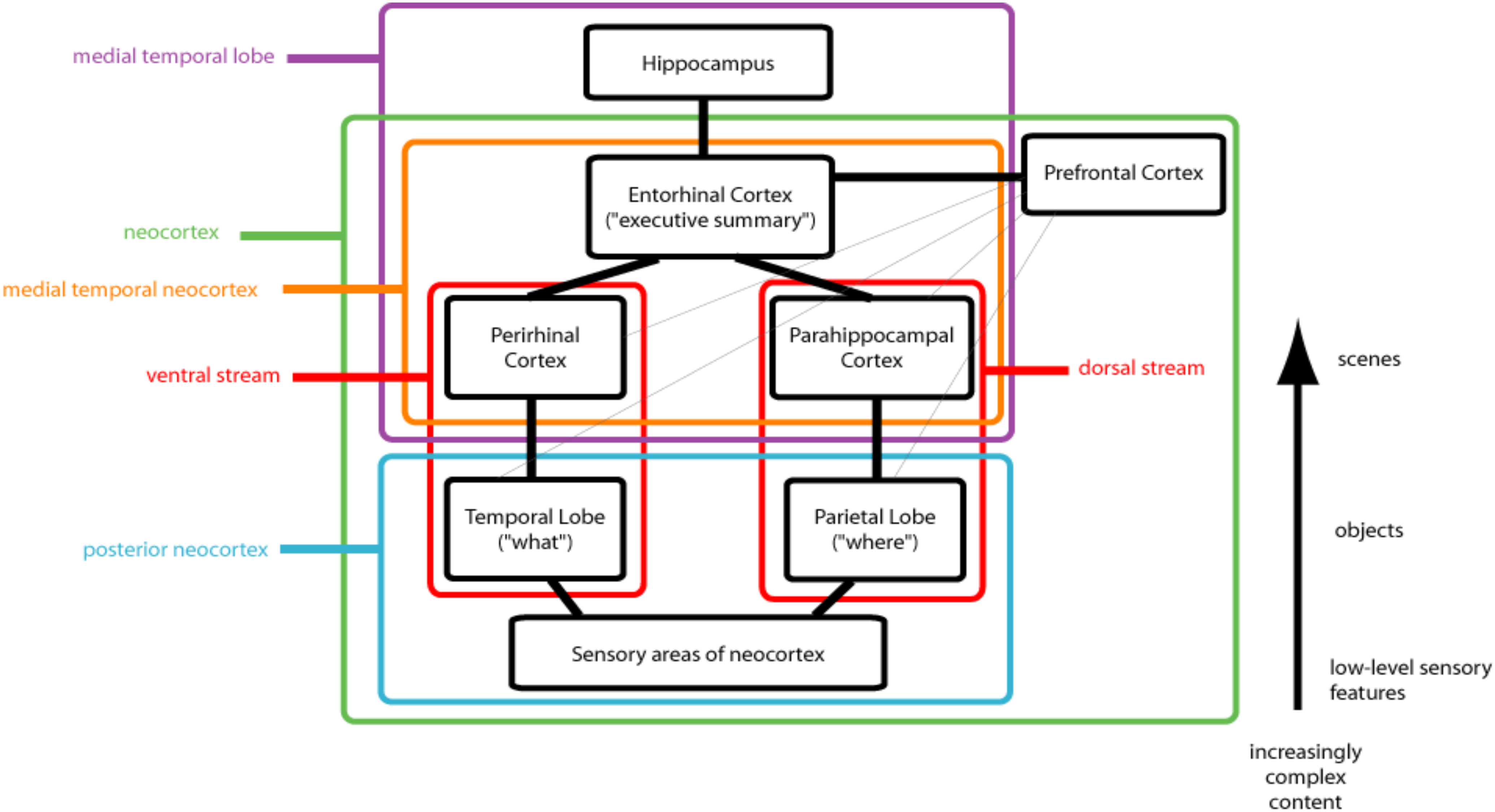
How does this system contribute to episodic memory formation?



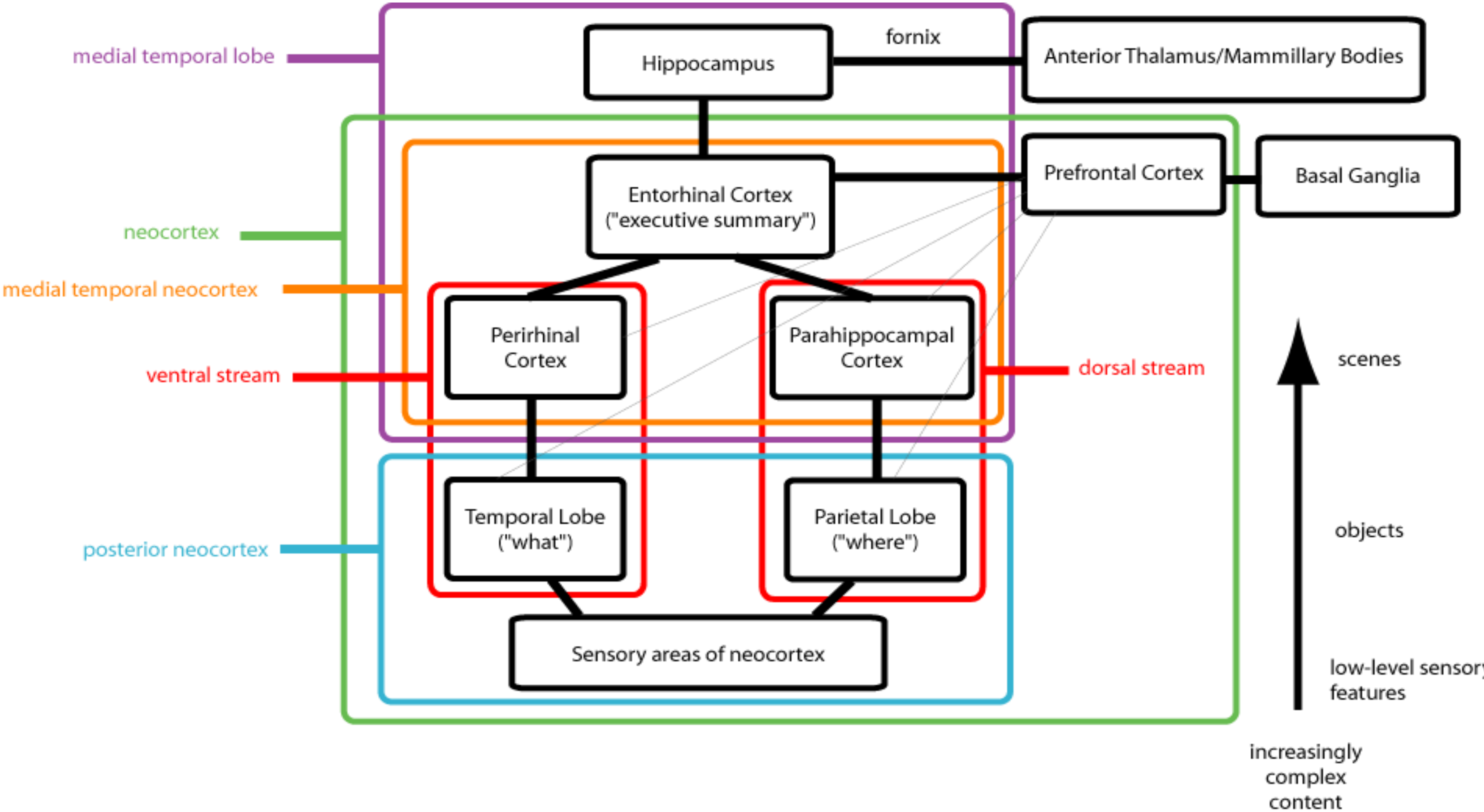
Perirhinal Cortex
Parahippocampal Cortex
ventral stream
dorsal stream
scenes
Temporal Lobe ("what")
Parietal Lobe ("where")
posterior neocortex
objects
Sensory areas of neocortex
low-level sensory features
increasingly complex content


Entorhinal Cortex
("executive summary")
medial temporal neocortex
Perirhinal
Cortex
Parahippocampal
Cortex
ventral stream
dorsal stream
scenes
objects
Temporal Lobe
("what")
Parietal Lobe
("where")
posterior neocortex
Sensory areas of neocortex
low-level sensory
features
increasingly
complex
content

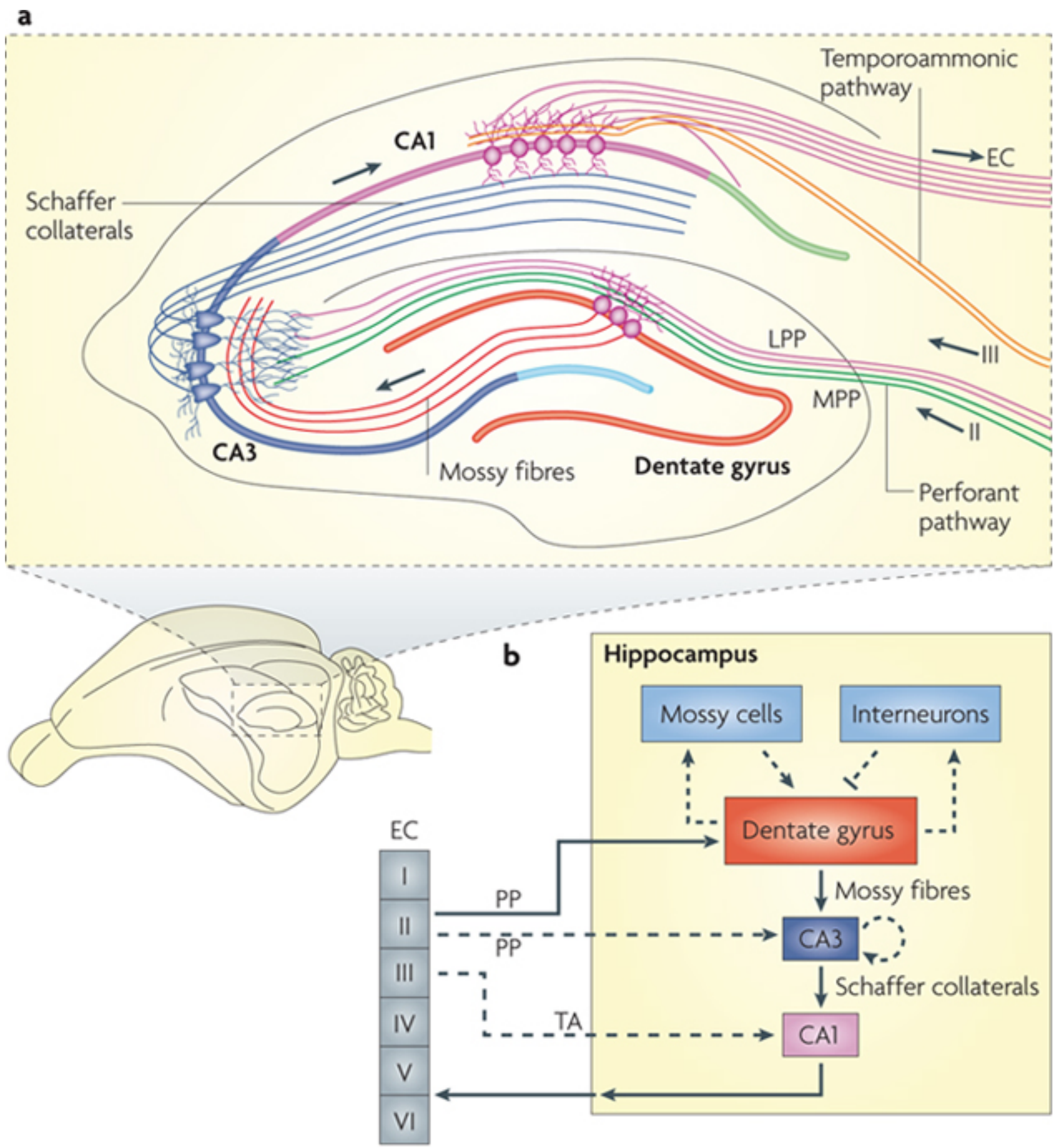

medial temporal lobe
medial temporal neocortex
ventral stream
posterior neocortex
Hippocampus
Entorhinal Cortex
("executive summary")
Perirhinal
Cortex
Parahippocampal
Cortex
dorsal stream
Temporal Lobe
("what")
Parietal Lobe
("where")
Sensory areas of neocortex
scenes
objects
low-level sensory
features
increasingly
complex
content



medial temporal lobe
neocortex
medial temporal neocortex
ventral stream
posterior neocortex
Hippocampus
Entorhinal Cortex
("executive summary")
Prefrontal Cortex
Perirhinal
Cortex
Parahippocampal
Cortex
dorsal stream
Temporal Lobe
("what")
Parietal Lobe
("where")
Sensory areas of neocortex
scenes
objects
low-level sensory
features
increasingly
complex
content


medial temporal lobe
neocortex
medial temporal neocortex
ventral stream
posterior neocortex
Hippocampus
fornix
Anterior Thalamus/Mammillary Bodies
Entorhinal Cortex
("executive summary")
Prefrontal Cortex
Basal Ganglia
Perirhinal
Cortex
Parahippocampal
Cortex
dorsal stream
Temporal Lobe
("what")
Parietal Lobe
("where")
Sensory areas of neocortex
scenes
objects
low-level sensory
features
increasingly
complex
content

The real deal...



The Medial Temporal Lobe and Memory

- review of relevant anatomy (how hippo connects to cortex)

- how the hippocampus works

- effects of large MTL lesions

- effects of more focal hippocampal lesions

- roles of particular hippocampal substructures

How does the Hippocampus Work?

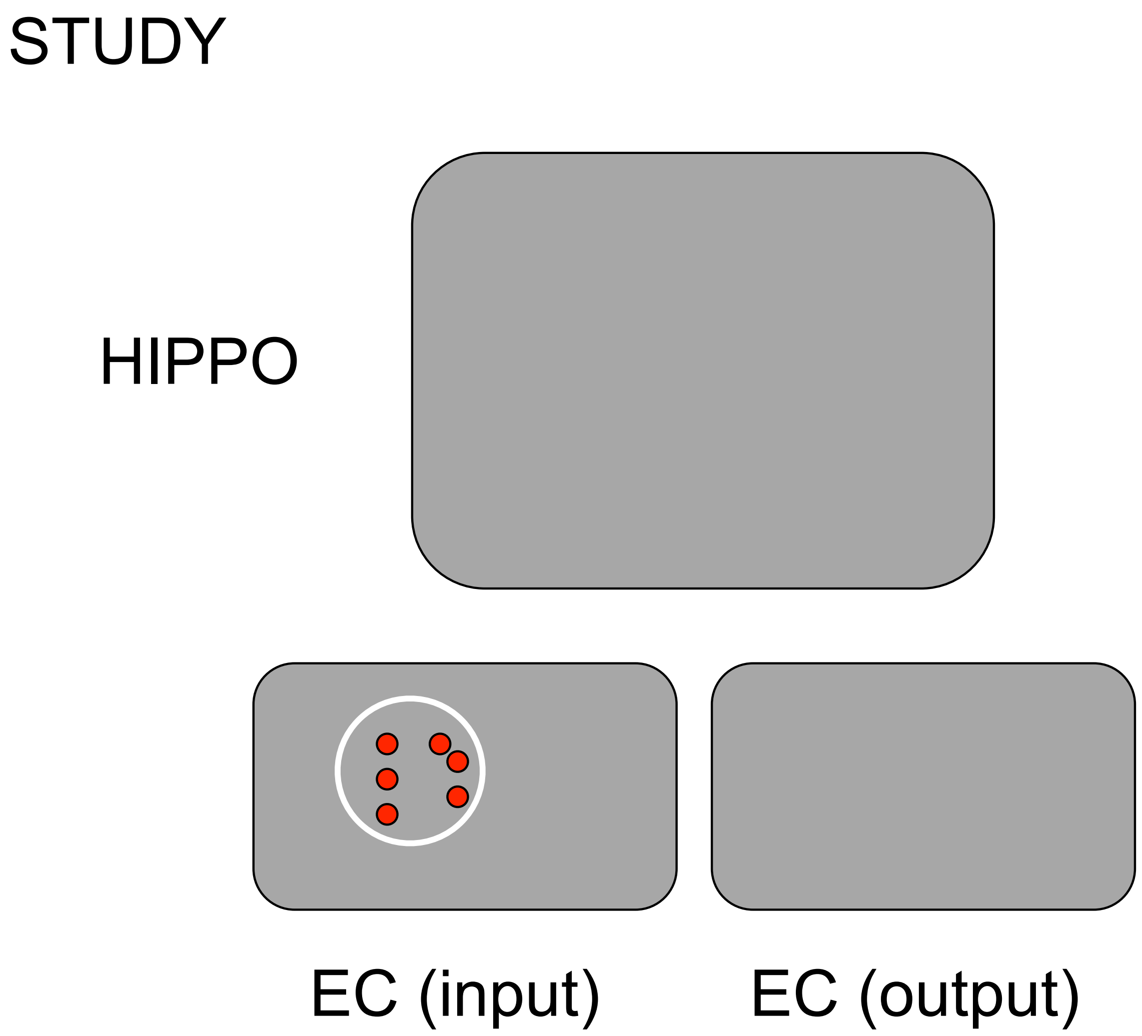
The job of the hippocampus is to rapidly memorize patterns of cortical activity (in entorhinal cortex?) so they can be recalled, given partial cues..

Key property: hippocampus tries to assign a distinct set of neurons to each memory

This **pattern separation** property allows the hipp to learn rapidly without suffering catastrophic interference

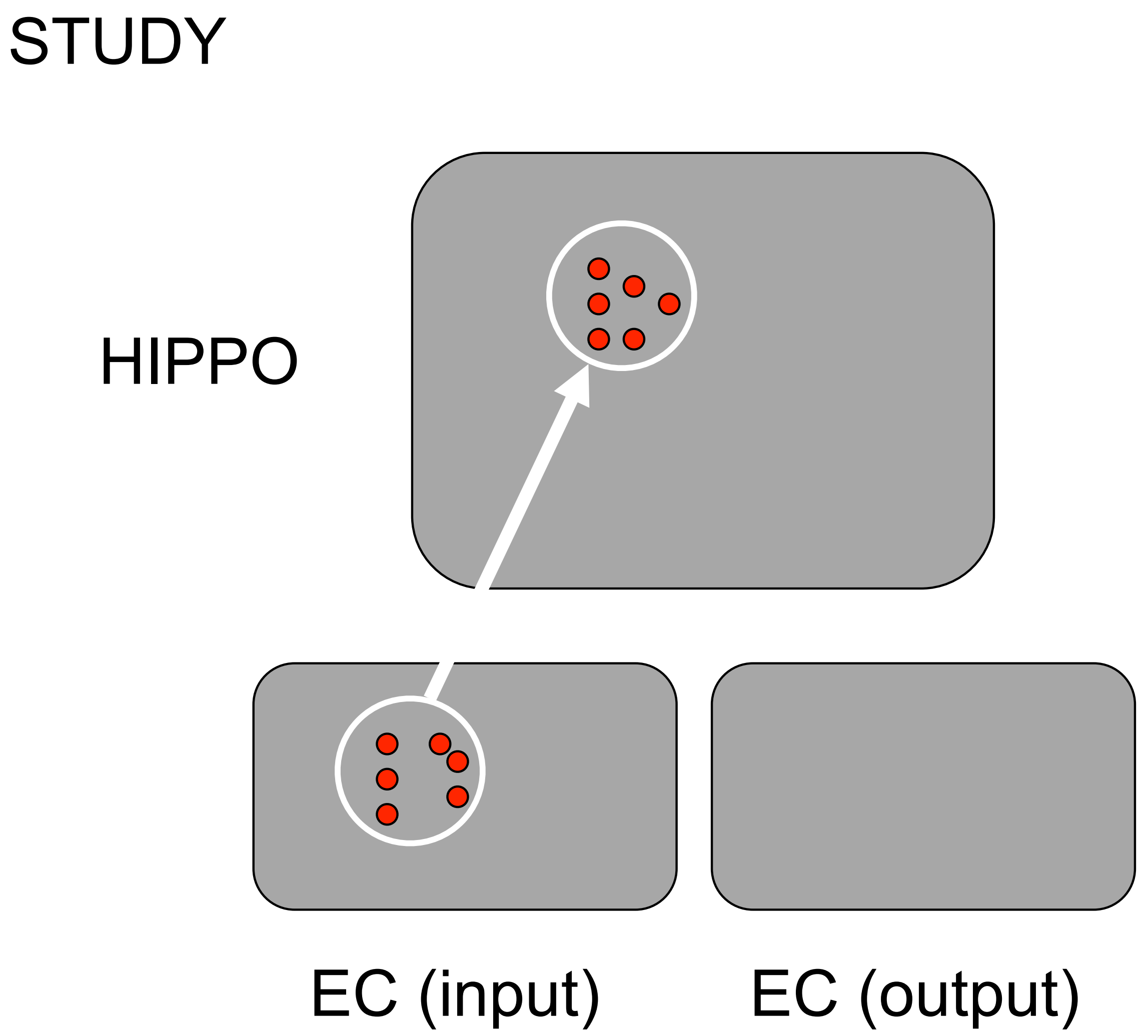
Goal for the first part of this lecture: Overview of how the hippocampus stores & retrieves memories, and how it accomplishes pattern separation

Hippocampal binding



Hippo binding

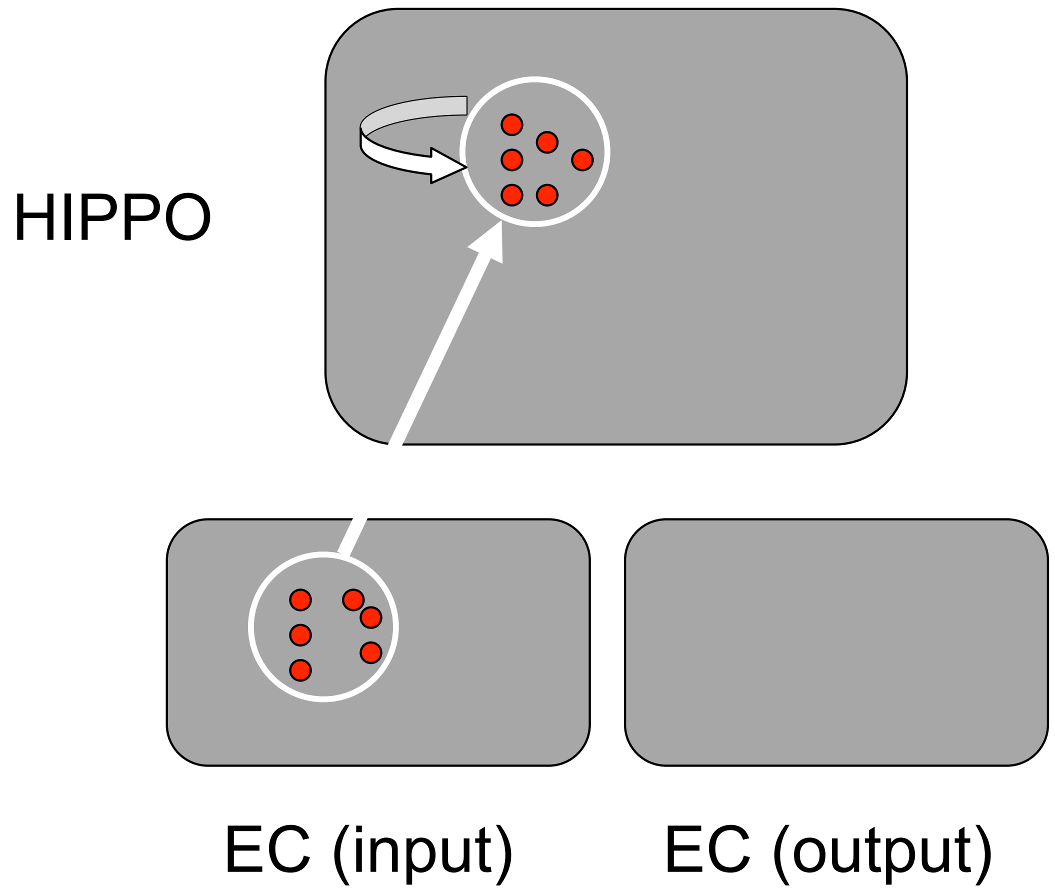
Activation Spreads from EC to Hipp (both directly and indirectly)



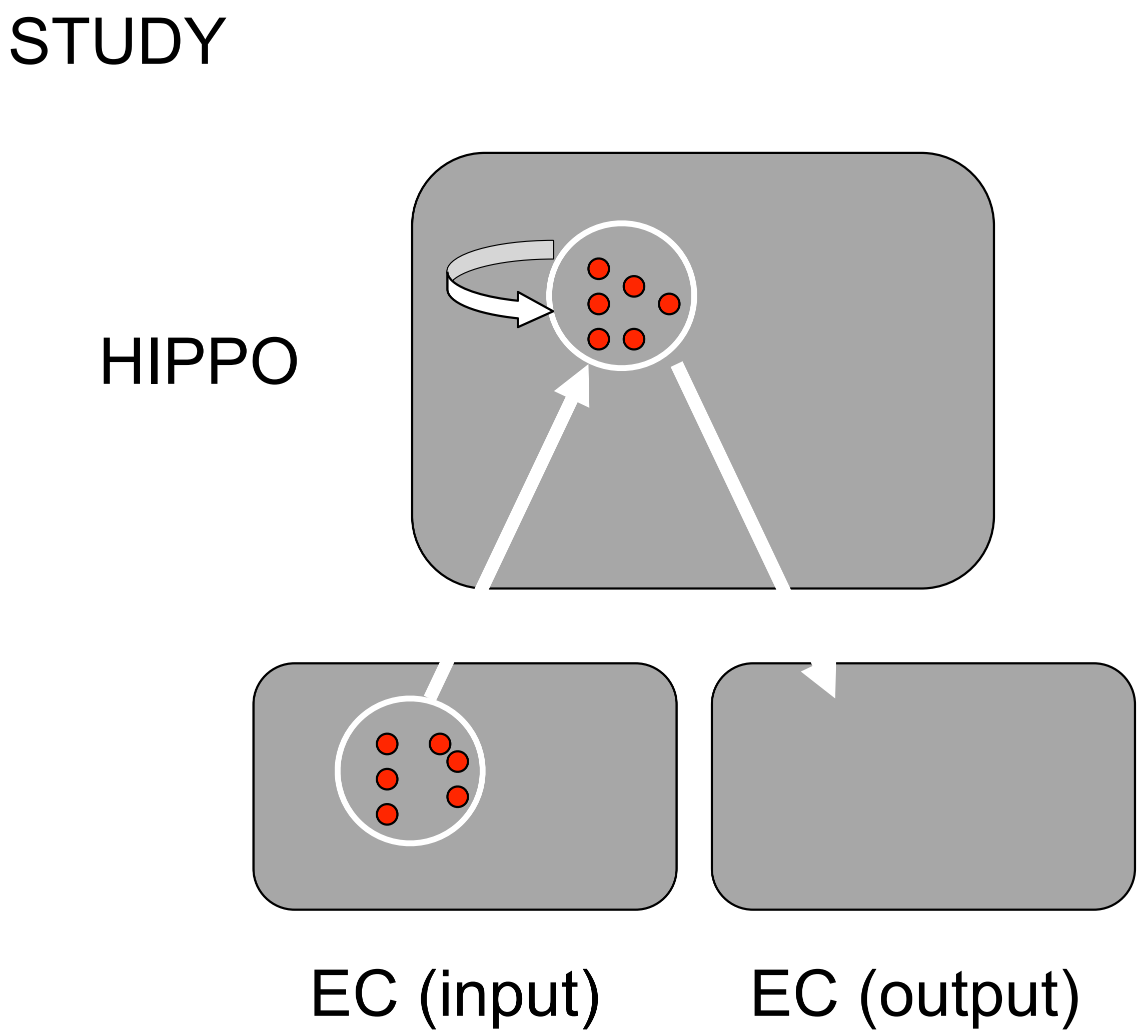
+ Hippo binding

Recurrent connections in HIPP allow for Hebbian mechanisms of synaptic strengthening to occur

STUDY

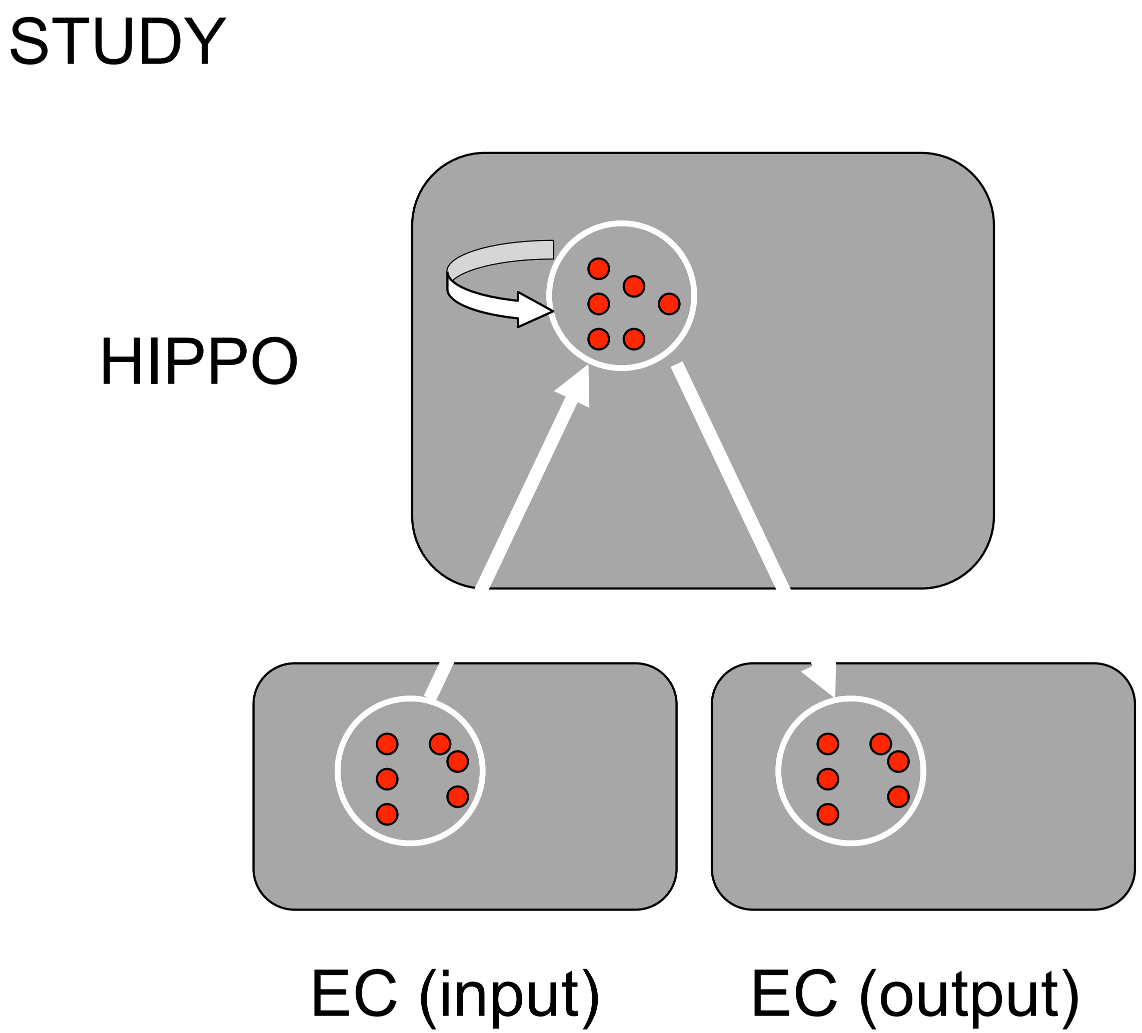


Hippo binding



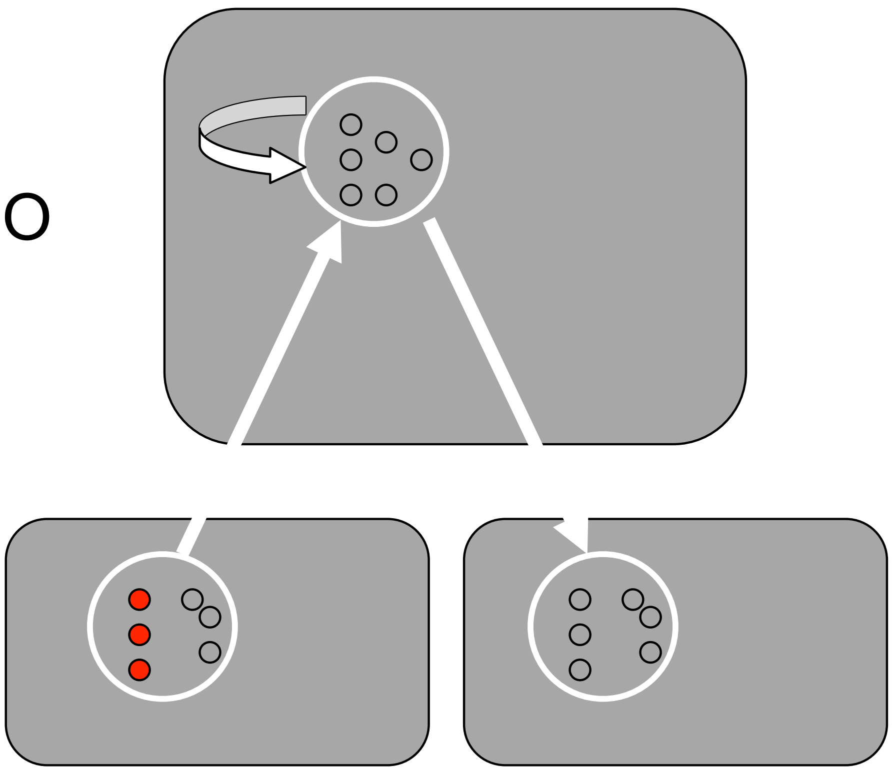
Hippo binding

Hippocampal episodic representation can then be read out into an output layer in EC



Pattern Completion

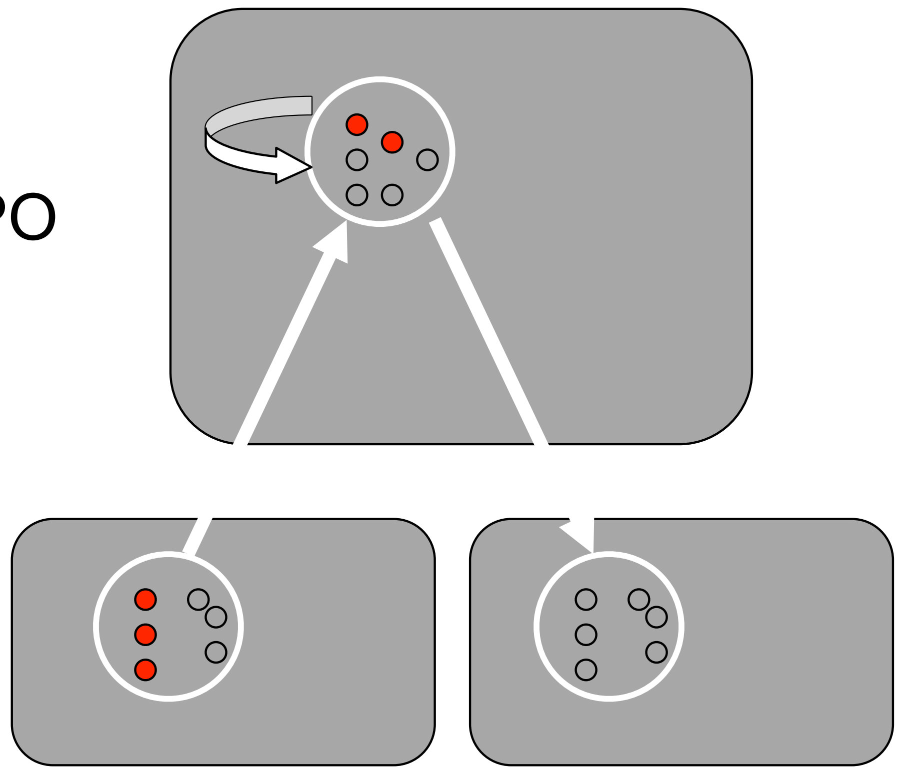
TEST



Pattern Completion

TEST

HIPPO

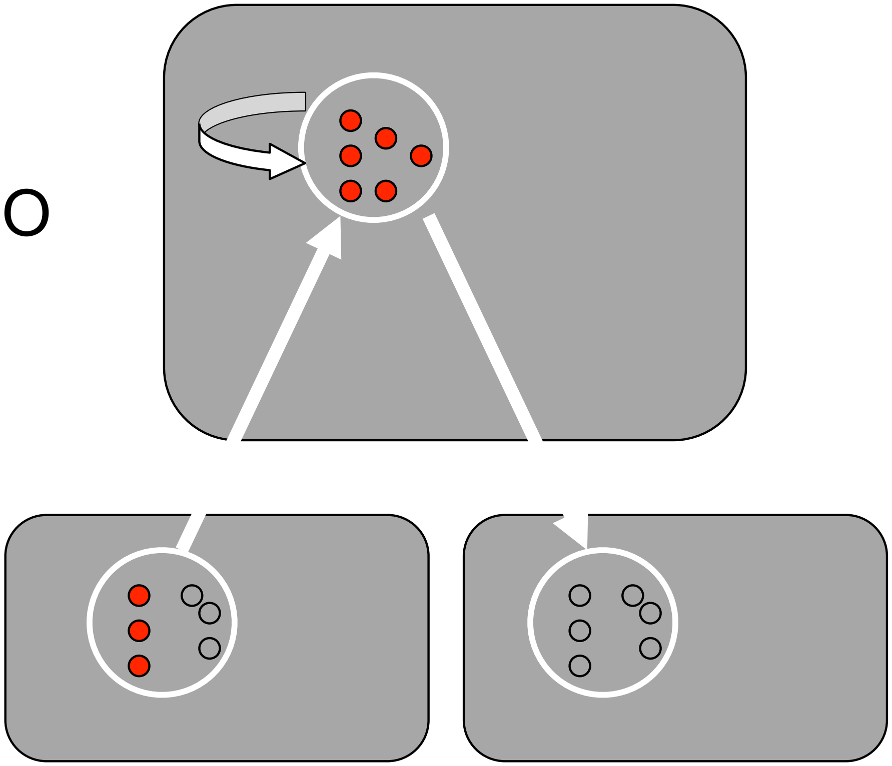


EC (input) EC (output)

+ Pattern Completion

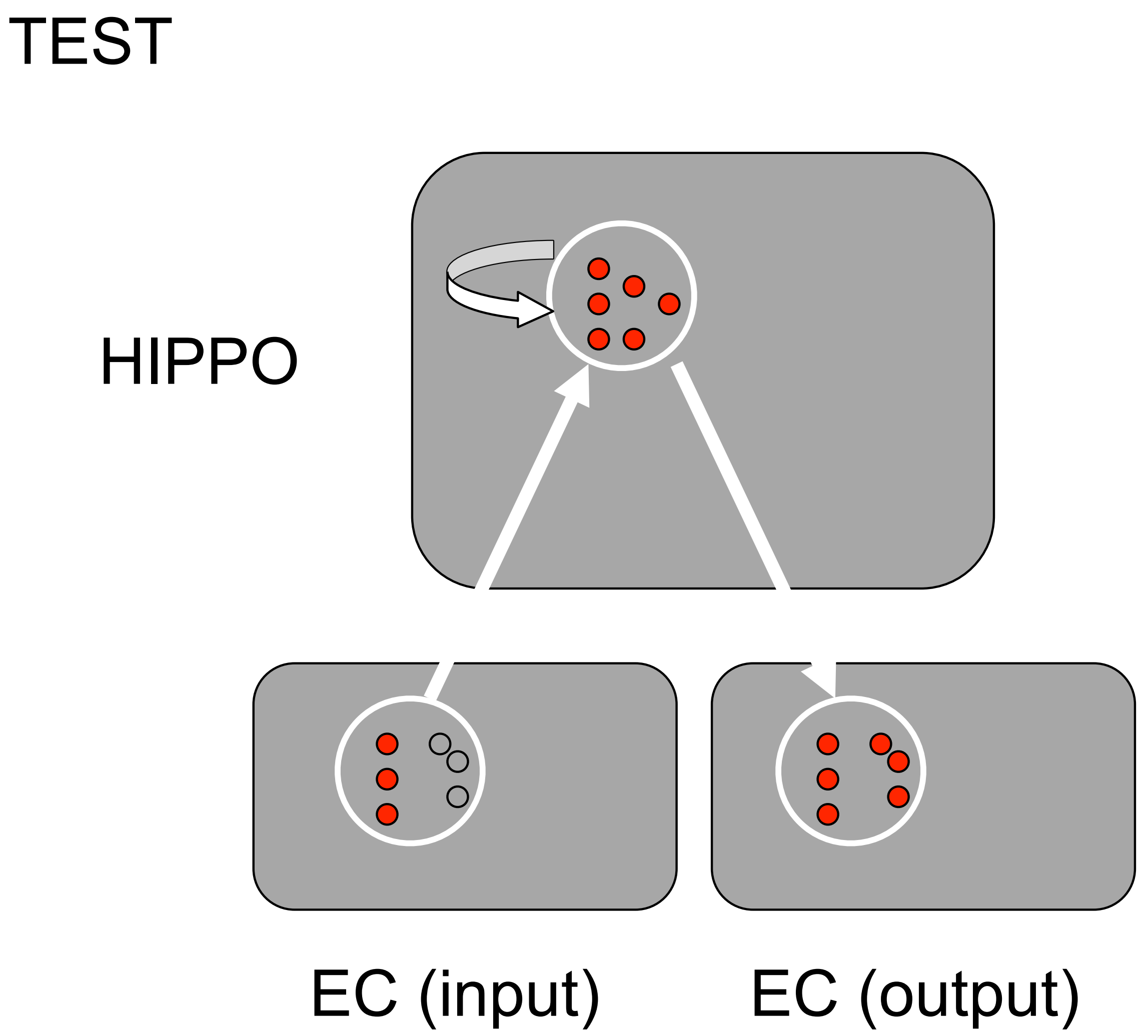
TEST

HIPPO

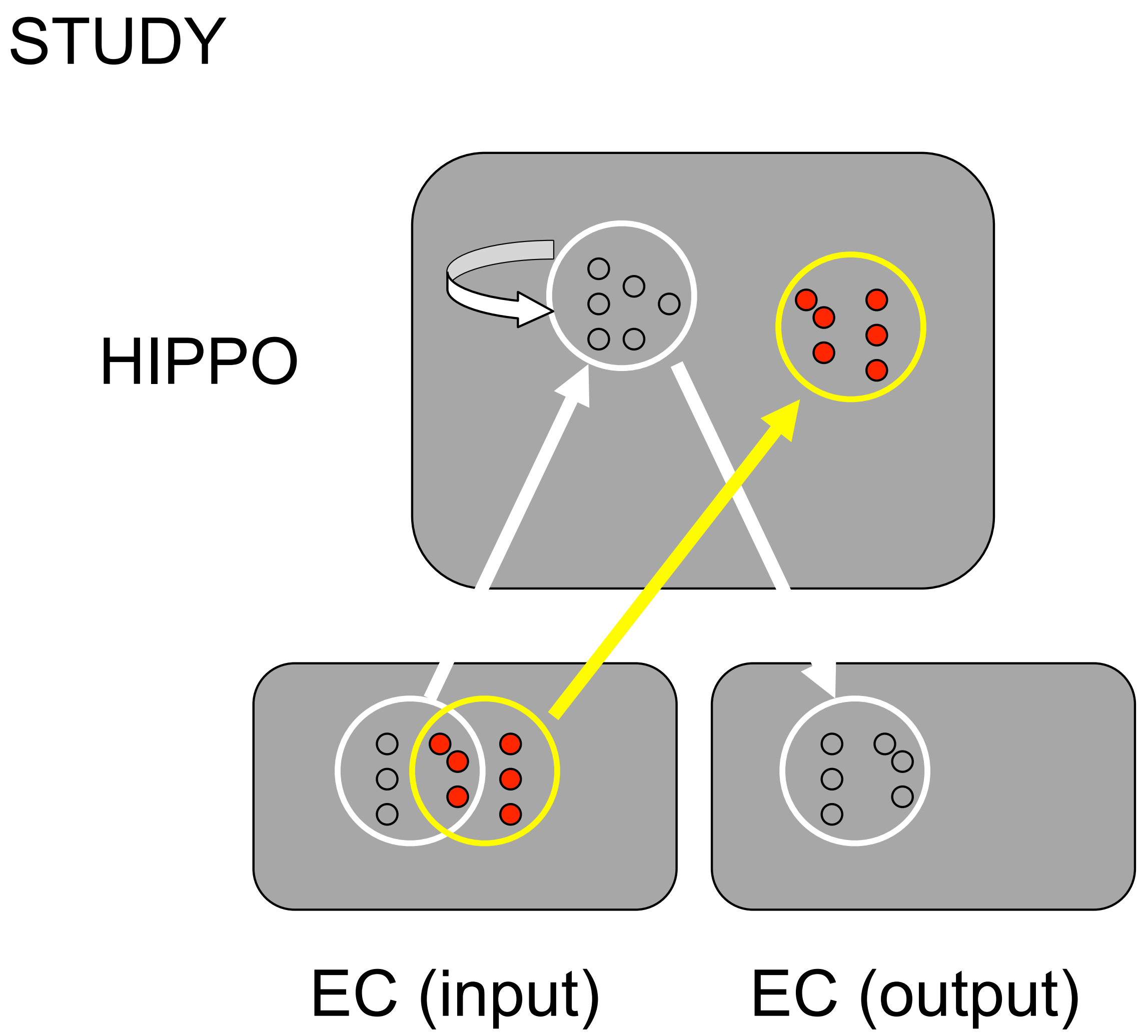


EC (input) EC (output)

Pattern Completion



Pattern Separation



Explaining Pattern Separation

How does the hippocampus manage to assign distinct representations to similar inputs?

Answer: Sparse activity

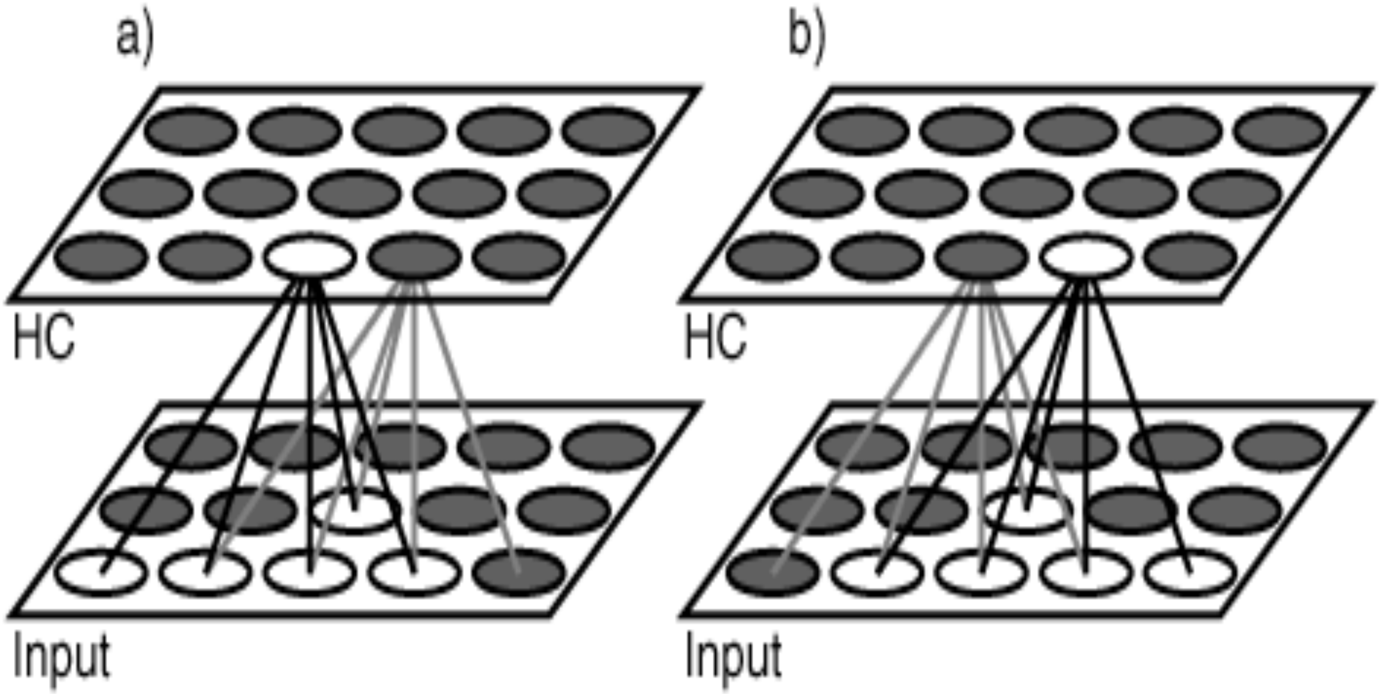
Only a small number of neurons are allowed to be active in hippocampus (relative to cortex)

Sparse activity is enforced by inhibitory interneurons

Inhibitory competition is especially fierce in the hippocampus

When activity is sparse, small changes in the input pattern can have large effects on the hippocampal representation

+ Pattern Separation and Sparse Activity



Here, hippocampal units are connected to 5 inputs; only one hippocampal unit is allowed to be active at a time

In the left-hand picture, the middle unit wins because all 5 of its inputs are active (whereas only 4 of its neighbor's inputs are active)

Changing one input unit causes a different hippo unit to win!

The Medial Temporal Lobe and Memory

- review of relevant anatomy (how hippo connects to cortex)

- how the hippocampus works

- effects of large MTL lesions

- effects of more focal hippocampal lesions

- roles of particular hippocampal substructures

Large MTL damage in humans..like H.M.

- Everything we talked about before?

- PLUS, spatial memory, implicit memory for associations, novel relational imagery

The Medial Temporal Lobe and Memory

- review of relevant anatomy (how hippo connects to cortex)

- how the hippocampus works

- effects of large MTL lesions

- effects of more focal hippocampal lesions

- roles of particular hippocampal substructures

Focal Hipp Lesions

All of the amnesia data we have discussed thus far relates to big lesions, encompassing several MTL brain structures (e.g., perirhinal cortex) in addition to the hippocampus

Recently, researchers have started to look at the effects of smaller lesions, limited to the hippocampus

Outline

The goal of this next section is to gain a better understanding of the respective contributions of the **hippocampus** vs. **surrounding medial temporal lobe structures** (perirhinal, parahippocampal, entorhinal cortex)

Focus on 3 areas

- Fact learning
- Item recognition
- Association formation

Vargha-Khadem et al. (1997)

Vargha-Khadem et al. (1997)- looked at three patients (including “Patient Jon”) who suffered hippocampal damage in childhood due to oxygen deprivation; surrounding cortical regions were spared

Recall of once-presented stimuli is terrible

study: window-reason; test: window-?

However, these patients do OK in school....?

Episodic/semantic

Vargha-Khadem et al. (1997) talk about these findings in terms of a dissociation between episodic memory and semantic memory

Episodic memory = recall & recognition of events
Semantic memory = recall & recognition of facts

Based on OK school performance, they argue that semantic learning ability may be intact

However, this conclusion is premature ..

Well-controlled studies of semantic learning in these patients have revealed that semantic learning is impaired also (e.g., Gardiner, Brandt, Baddeley, Vargha-Khadem, & Mishkin, 2008, *Neuropsychologia*)

Episodic/semantic

Episodic and semantic memory are closely intertwined

Contrary to claims made in the Vargha-Khadem article, it is impossible to knock out episodic memory without there being **some** effect on semantic memory

All fact memories start out as (hippocampally-dependent) episodic memories

If a person is exposed to the fact many times (either because of many "real world" exposures, or because of hippocampal replay), the fact eventually sinks in to cortex and becomes a semantic memory.

Explaining the Vargha-Khadem Results

Overall: The patients' good performance in school validates the idea that cortex is an enormously powerful learning system when given repeated exposure to study materials...

Studies of patients with big lesions showed that, if you have **some** sparing of cortex, you can do **some** memorization of new facts

The Vargha-Khadem study shows that, if you have a **lot** of sparing of cortex, you can do a **lot** of memorization of new facts

Item Recognition

In the patients studied by, Vargha-Khadem et al. (1997), item recognition is OK, even after 1 study exposure

study:
    banana
    moose
    telephone

test:
    eardrum?
    banana?

Animal studies also support the idea that focal hippocampal damage does not disrupt item recognition

In animals, recognition memory is typically tested using a **delayed nonmatching to sample (DNMS)** procedure

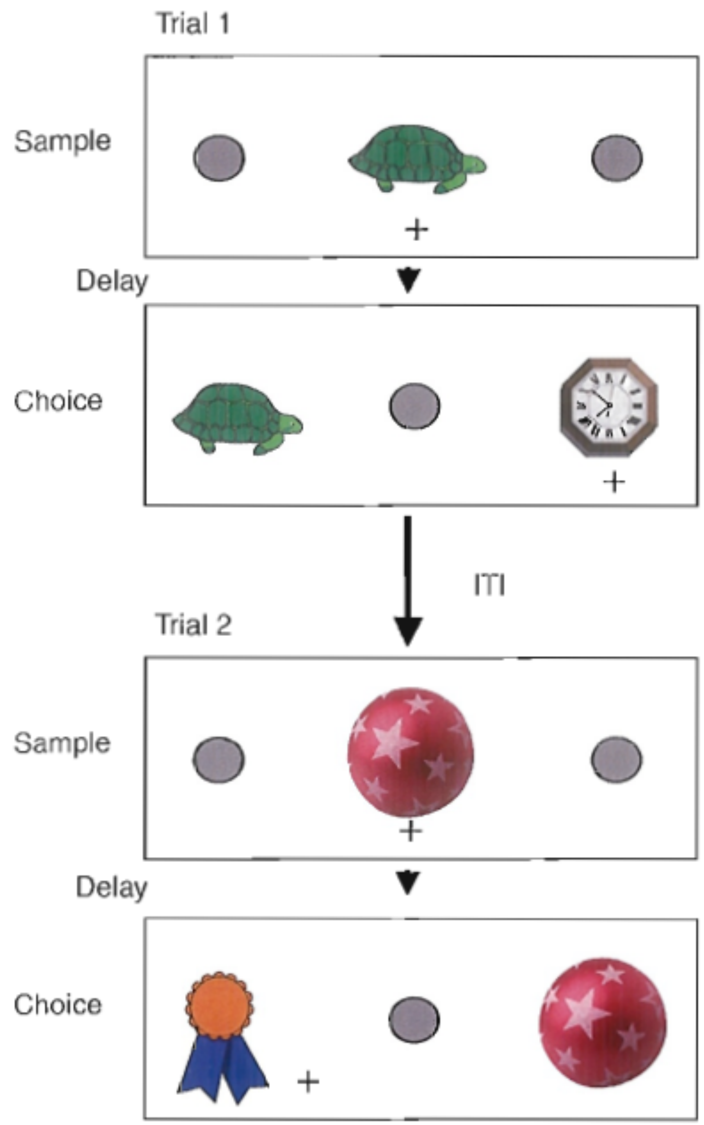


Figure 3 Delayed nonmatching-to-sample (DNMS) task. Trial 1: The monkey is presented with a sample object covering a food reward (+) in the central well of the testing tray. After a delay, during which the test tray is obscured, the sample object and a novel choice object are presented simultaneously covering the lateral wells, but the novel object is now rewarded (i.e., nonmatch-to-sample). After a brief intertrial interval (ITI), the next trial begins. Trial 2: For this and all subsequent trials, a new pair of objects serves as the sample and choice (i.e., trial unique).

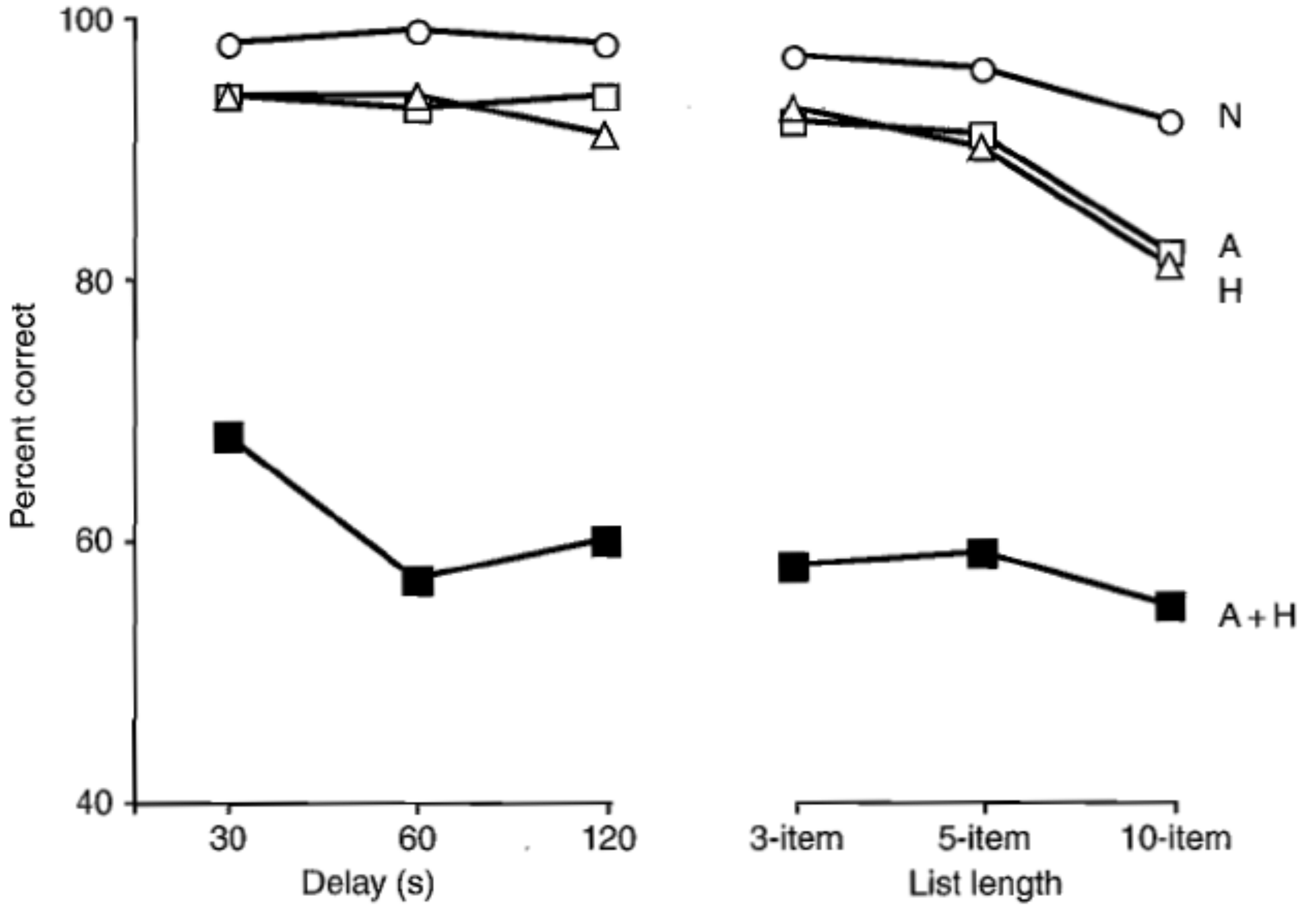


Figure 4 Performance of monkeys with damage to the medial temporal lobe on DNMS, when delays are increased from 30 to 120 s or the number of to-be-remembered objects is increased from three to ten. N, unoperated controls; A, animals with aspiration lesions of the amygdala; H, animals with aspiration lesions of the hippocampal formation; A + H, animals with combined amygdala and hippocampal lesions. Adapted from Mishkin M (1978) Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus. *Nature* 273(5660): 297–298, with permission from Macmillan Journals Ltd.

Item Recognition

- If the hippocampus is not the key structure for item recognition memory, which structure(s) is most responsible for item recognition?
- **Perirhinal cortex!**

Big lesions that encompass perirhinal cortex lead to poor recognition memory; subjects with spared perirhinal cortex show better item recognition memory

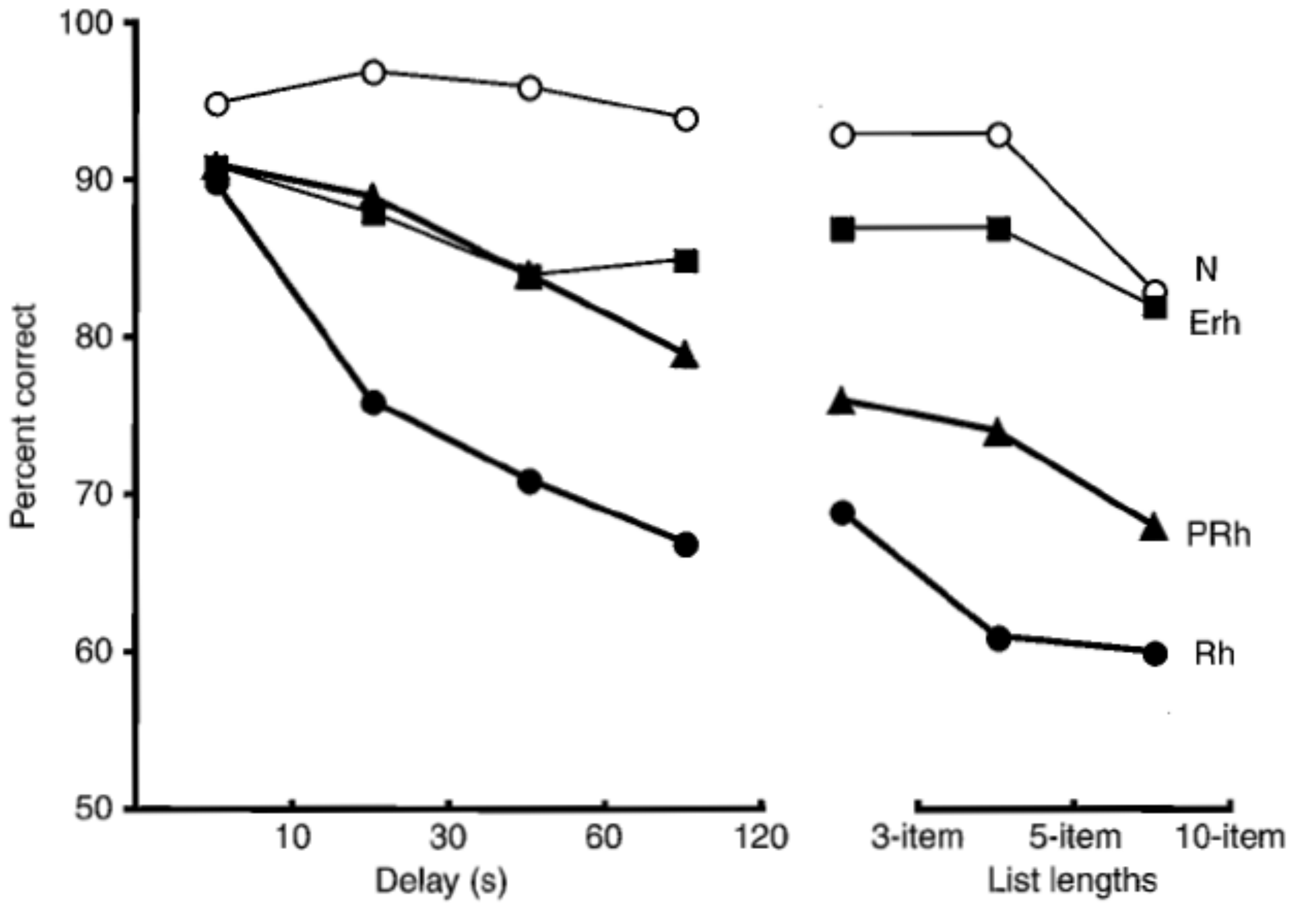


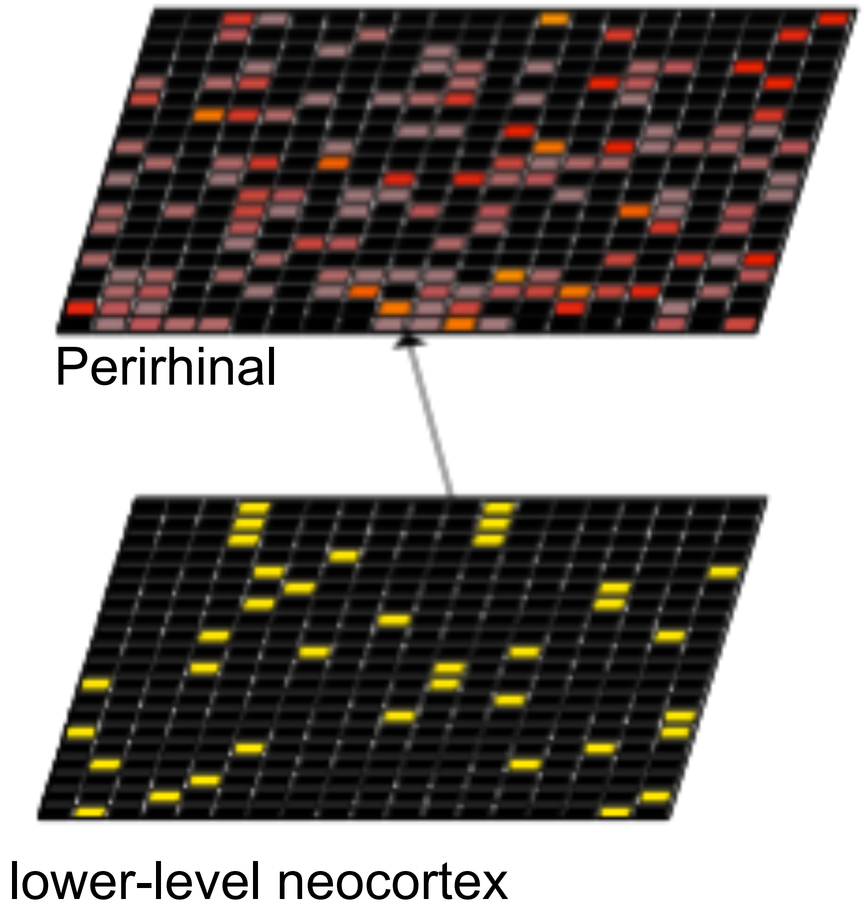
Figure 5 Effects of entorhinal (Erh), perirhinal (PRh), and rhinal (Rh; Erh + PRh) lesions on DNMS performance as compared to unoperated (N) animals. Adapted from Meunier M, Bachevalier J, Mishkin M, and Murray EA (1993) Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. *J. Neurosci.* 13: 5418–5432, with permission from the Society for Neuroscience.

Item Recognition

Puzzle: How do we explain the finding of good recognition memory (after a single study exposure) in people and animals with focal hippocampal damage?

The cortex learns incrementally

Good recognition in these subjects after one study trial seems to contradict this “incremental learning” principle



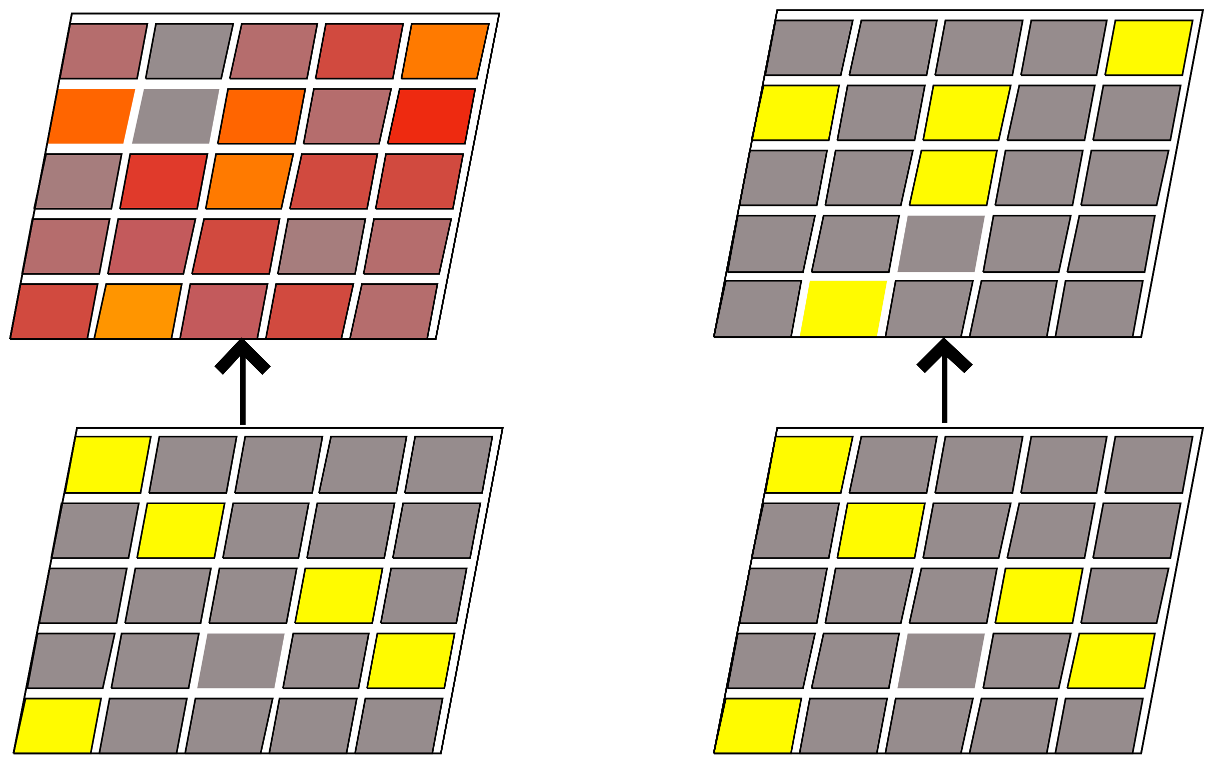
Even though cortex learns incrementally, **repeated presentation** of a stimulus leads to predictable changes in the overall structure of cortical representations...

Cortical Recognition

Repeating stimuli makes their representations **sharper**

=> novel stimuli **weakly** activate a **large** number of units
=> repeated stimuli **strongly** activate a **small** number of units

novel -->



<-- repeated

There is less overall activity in the upper (perirhinal) layer for repeated vs. novel stimuli

Recognition in Cortex

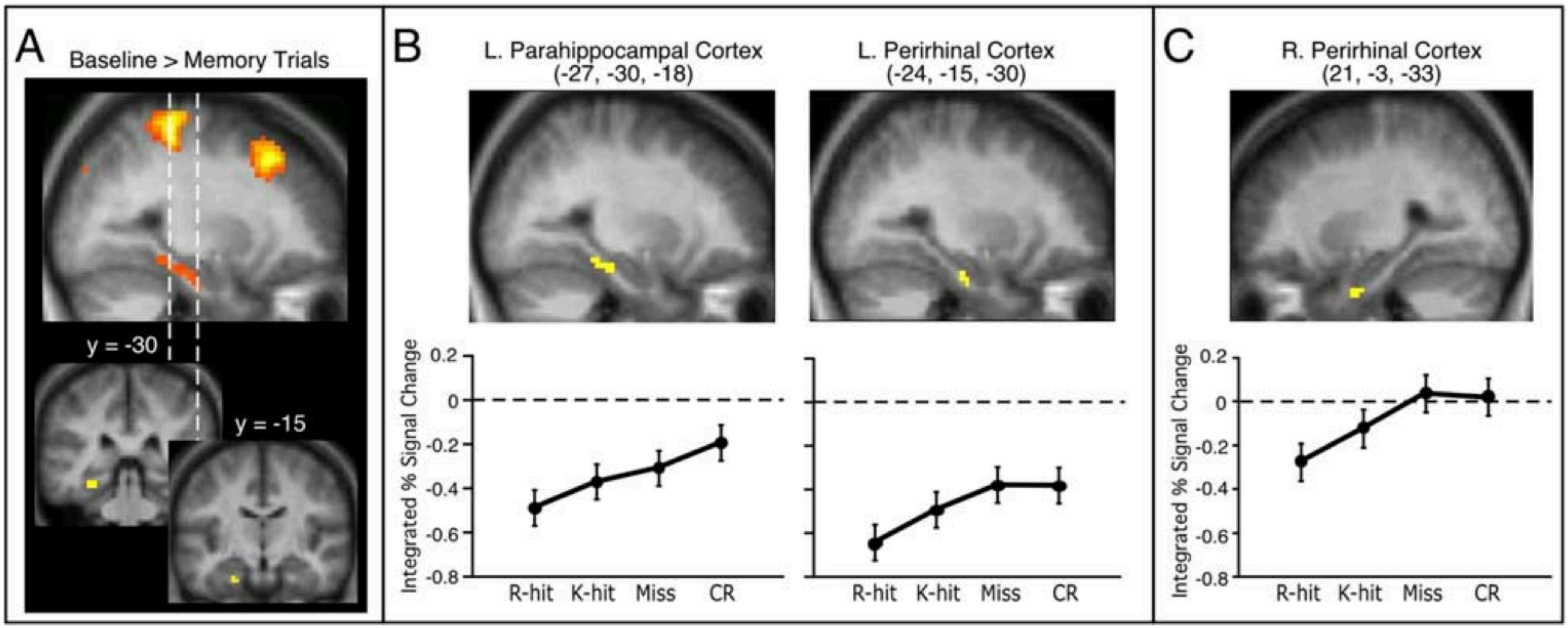
The model's prediction of less activity (overall) in perirhinal cortex for repeated vs. novel stimuli has been validated by neuroimaging studies and single-cell-recording studies (e.g., Xiang & Brown, 1998)

This finding of less activity for repeated stimuli implies that we can make recognition judgments by **reading out the overall amount of activity in perirhinal cortex**

if activity is low, say "studied"
if activity is high, say "nonstudied"

Recognition in Cortex

The idea that subjects base their recognition judgments on perirhinal activity is supported by a fMRI study conducted by Gonsalves et al. (2005, *Neuron*)



Relational Memory

If the hippocampus isn't directly responsible for item recognition memory, what *is* the hippocampus responsible for?

Eichenbaum and Cohen have argued that the hippocampus is crucial for **relational memory**: Rapidly forming a network of links between stimuli that can be **flexibly** traversed (in mind) by subjects...(think Tolman!)

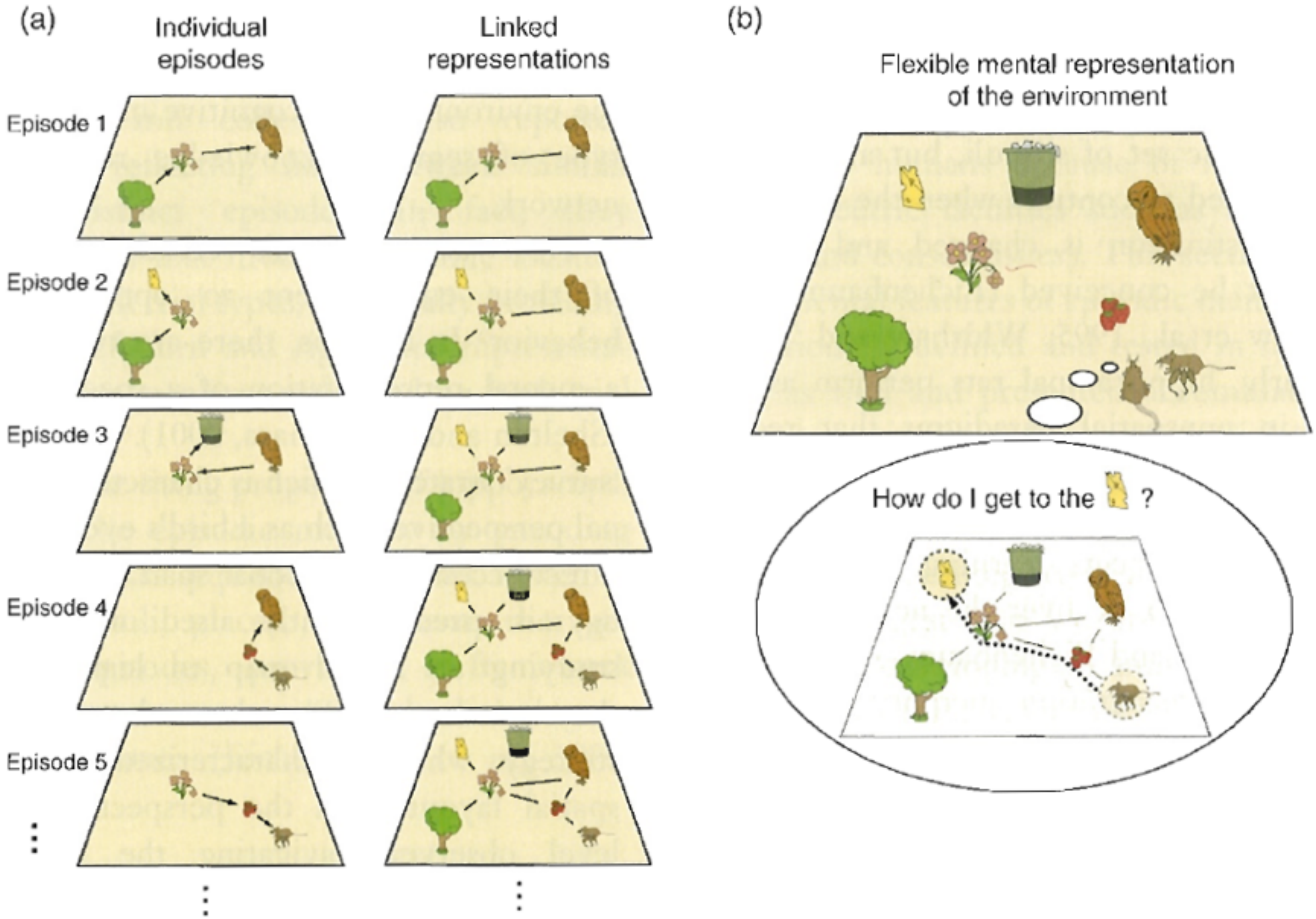


Figure 13 Cognitive map of the environment built out of individual episodes. (a) The column on the left represents individual episodes, each consisting of a particular trajectory in the animal's environment. The column on the right shows the representation that gradually forms as the common features between individual episodic memories (in this case, locations) are linked together. (b) After extensive experience in the environment, the linked representations can support flexible navigation by allowing the animal to determine a novel trajectory between two known points.

Konkel et al. (2009)
Relational Memory Study

The goal of this study was to simultaneously test item memory and different forms of relational memory in patients with focal hippocampal damage, patients with larger MTL lesions, and controls

Predictions:

Patients with focal hipp damage will show impaired relational memory and relatively spared item memory

Patients with larger MTL lesions will show impaired relational AND item memory

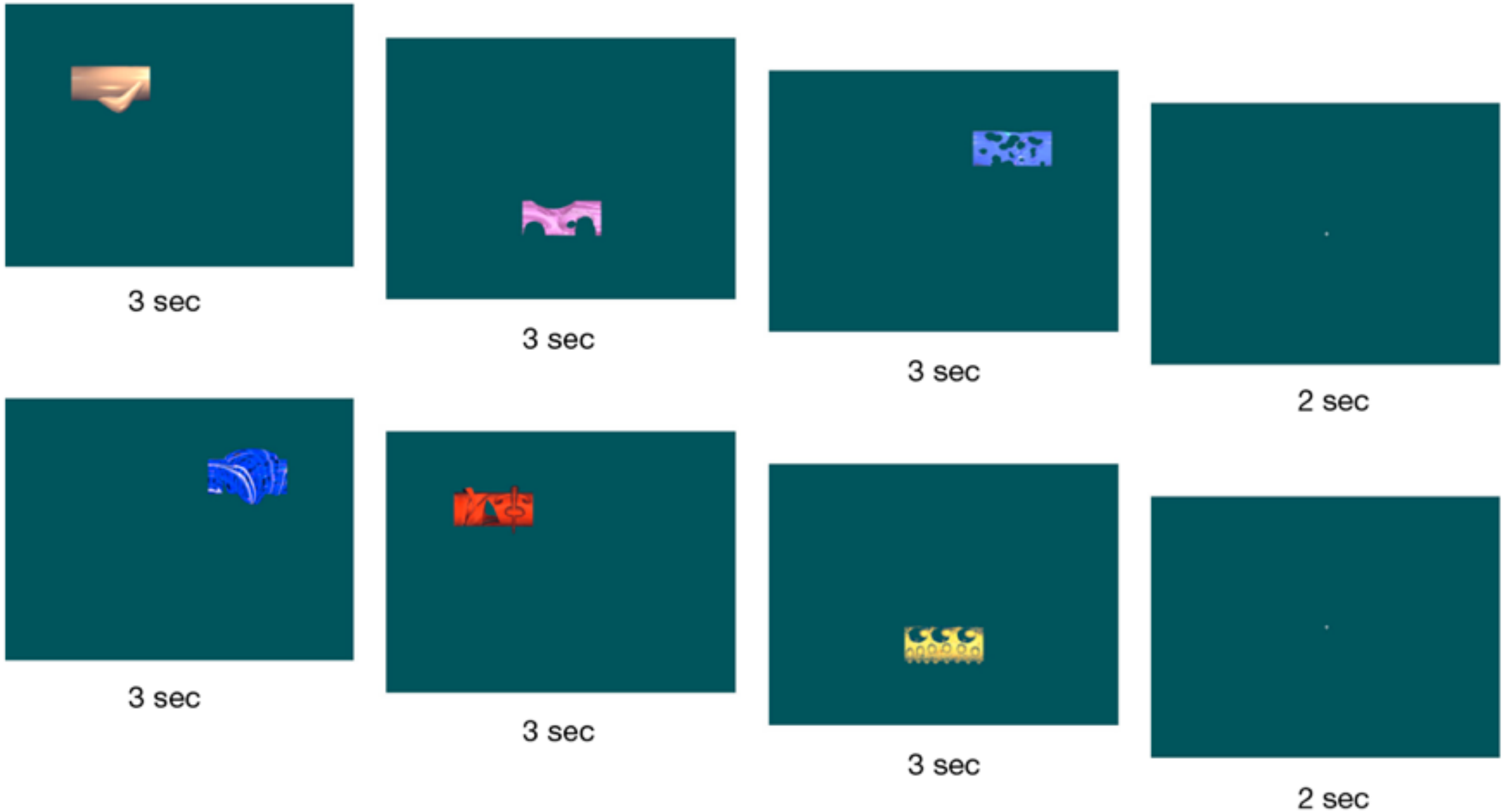


Figure 2 | **Two sample study trials.** Each individual stimulus was displayed for 3 s before being replaced by the next. Each set of three stimuli was separated by a 2-s fixation screen.

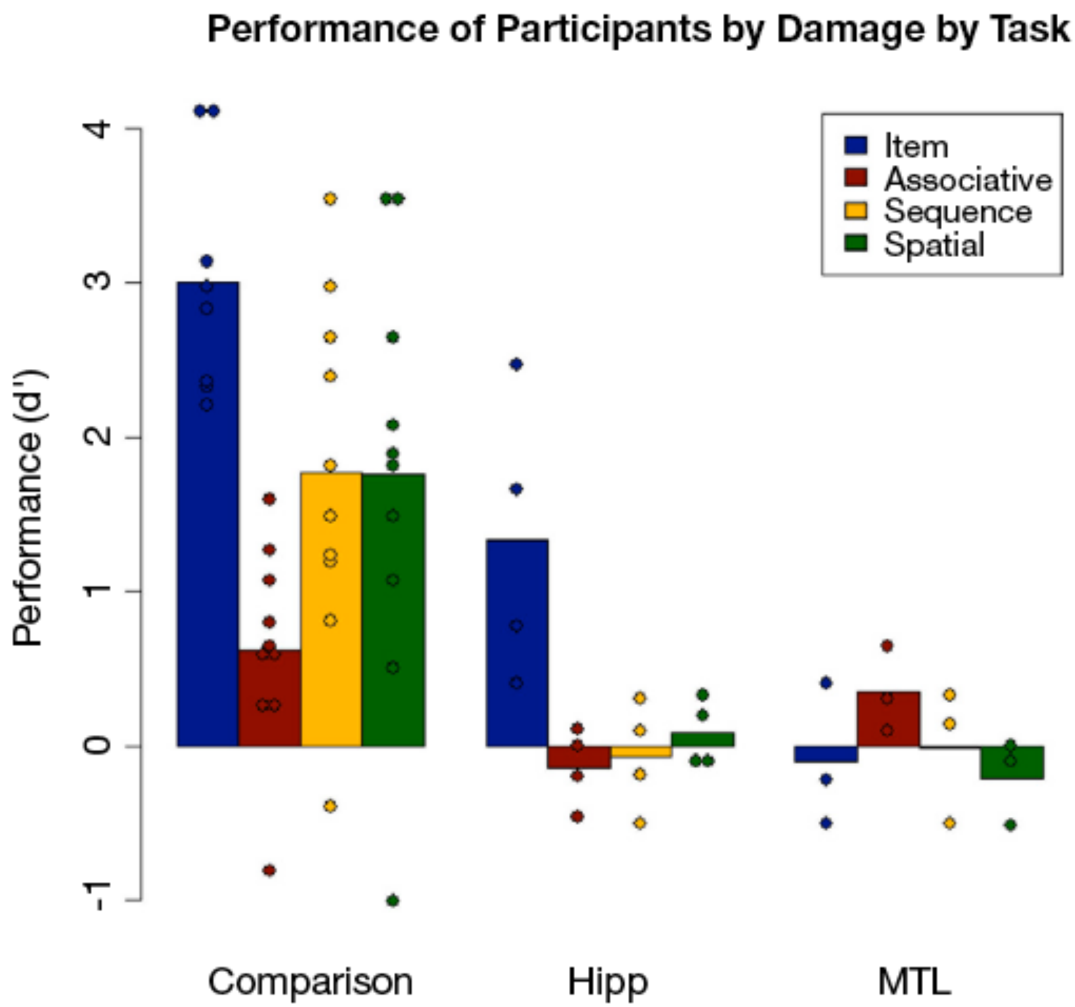


Figure 7 | **Performance on each task by group; dots represent individual participants' scores.**

Subsequent Memory for Item and Relational Information

FMRI studies have used the **subsequent memory** paradigm to explore how different brain regions contribute to encoding of item and relational information

- Scan people during study
- Later memory test
- Sort brain data from the study phase based on memory performance on the test phase (i.e. subsequently remembered, or subsequently forgotten)
- Compare the “subsequently remembered” brain scans to the “subsequently forgotten” brain scans

Subsequent Memory for Item and Relational Information

Davachi et al, (2003) used this paradigm to ask two questions:

- Activity in which brain region(s) predict subsequent **item** memory?
- Activity in which brain region(s) predict subsequent **relational** memory?

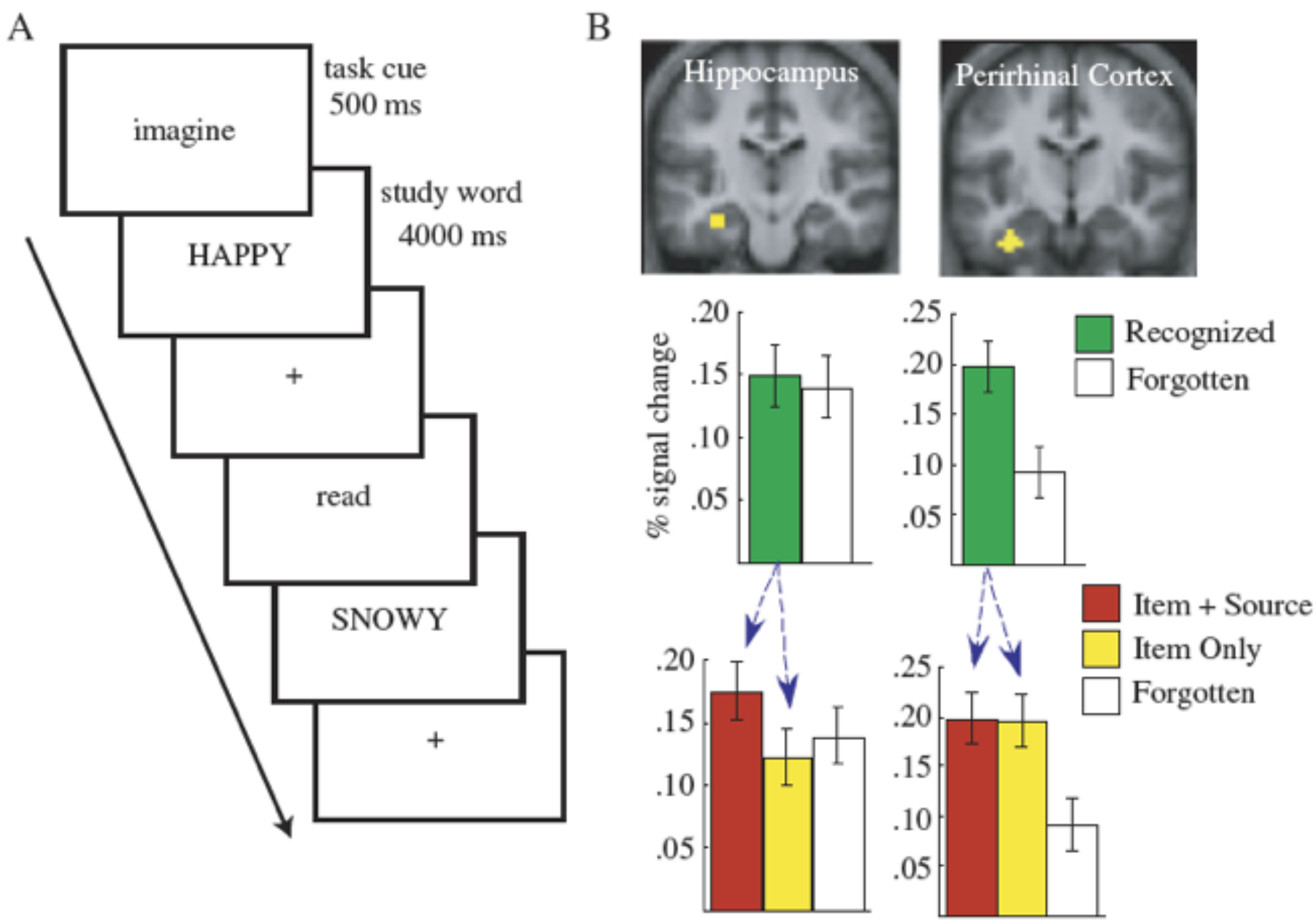


Fig. 1. Encoding paradigm (A) and activation in medial temporal lobe areas as related to subsequent memory (B). In the encoding trial, subjects were scanned while being shown a list of study adjectives (e.g., "happy," "snowy"); before the presentation of each study item, they were instructed to perform one of two encoding tasks (i.e., imagine, read) with it. Encoding activation (as % signal change) in the hippocampus did not predict item recognition when probed later, but it did predict subjects' ability to recollect the task in which they processed the items (i.e., item + source); conversely, encoding activation in the perirhinal cortex predicted later successful item recognition but not recollection of episodic details (Davachi, Mitchell, & Wagner, 2003).

The Medial Temporal Lobe and Memory

- review of relevant anatomy (how hippo connects to cortex)

- how the hippocampus works

- effects of large MTL lesions

- effects of more focal hippocampal lesions

- roles of particular hippocampal substructures

Perirhinal Ctx: item recognition, sharper signal with repetition

Hippocampus: associative memory via pattern separation and completion

Unitization and Associative Memory

Just an aside: there are some circumstances where it appears that cortex (in particular, perirhinal cortex) can rapidly (i.e., based on one study trial) form new associations on its own

Specifically: If subjects can **unitize** the two things being associated, then perirhinal cortex can form the association on its own

The Medial Temporal Lobe and Memory

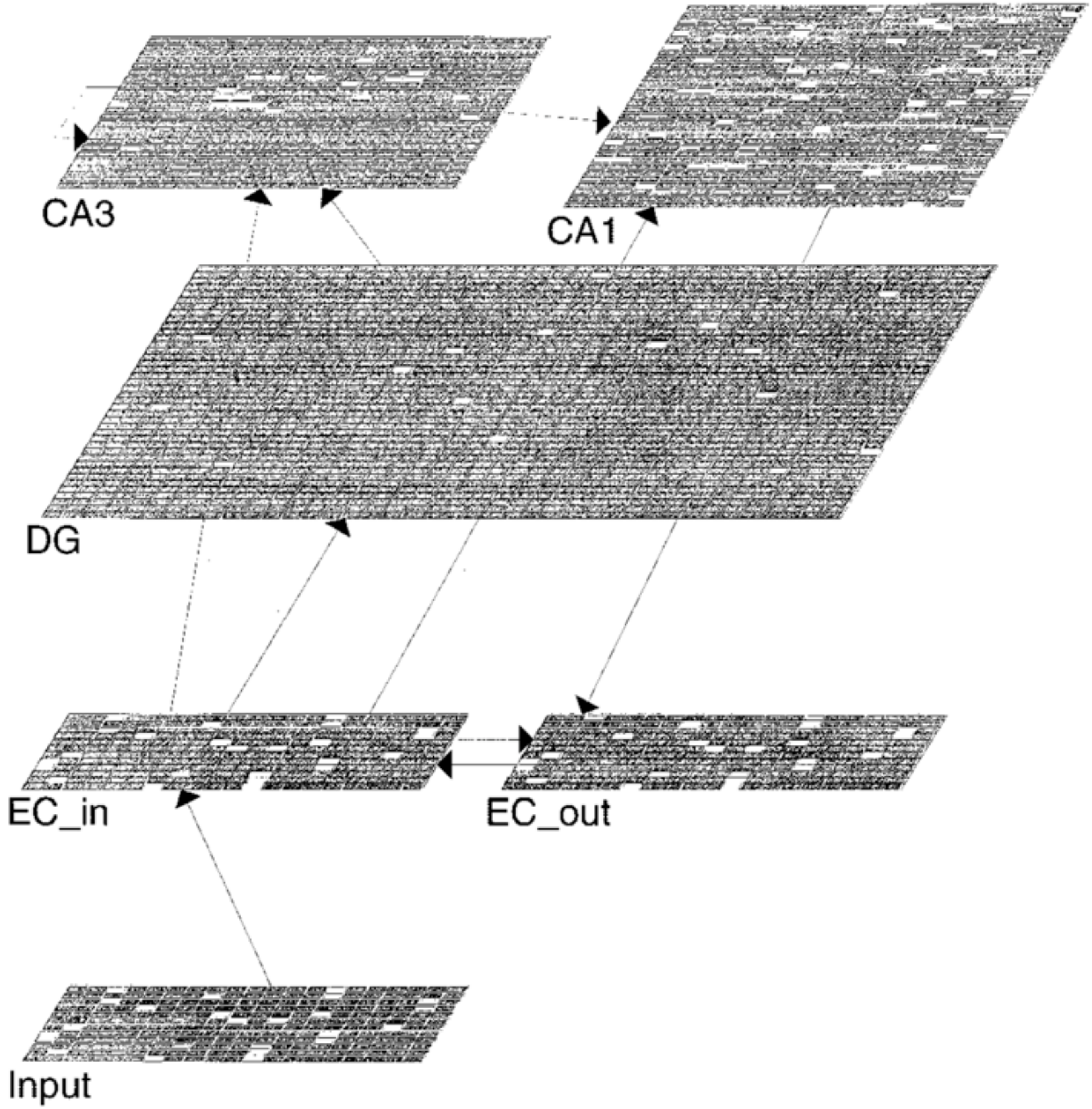
- review of relevant anatomy (how hippo connects to cortex)

- how the hippocampus works

- effects of BIG lesions to the hippocampal region

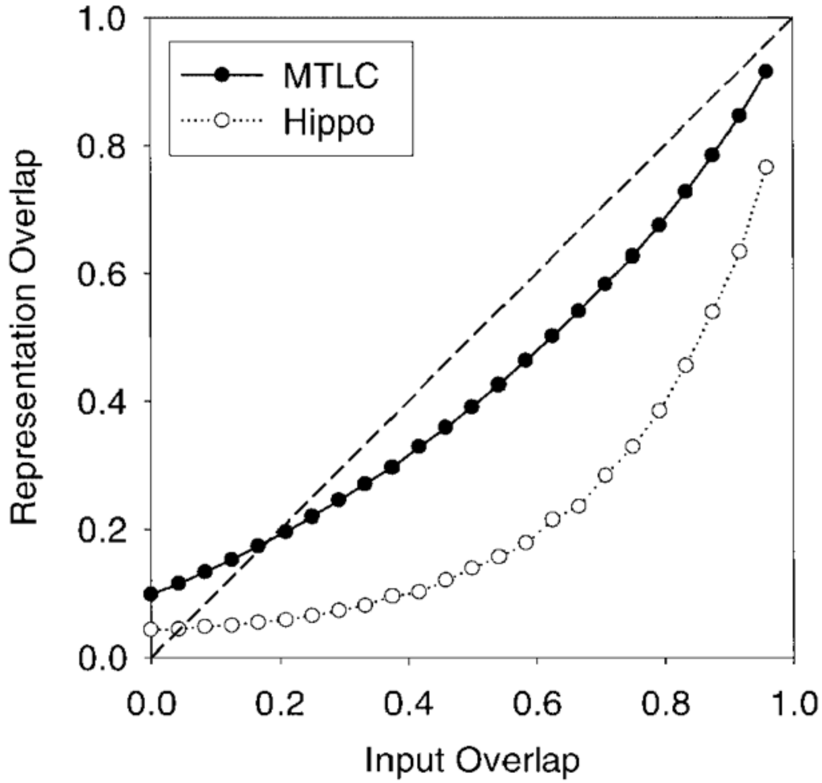
- effects of more focal hippocampal lesions

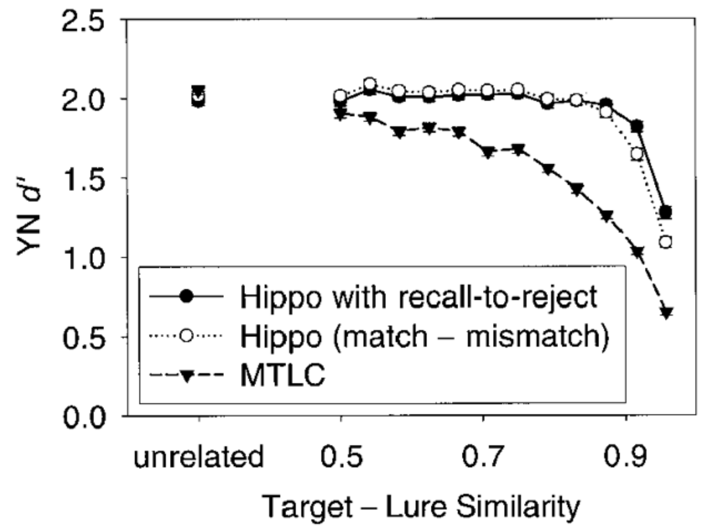
- roles of particular hippocampal substructures

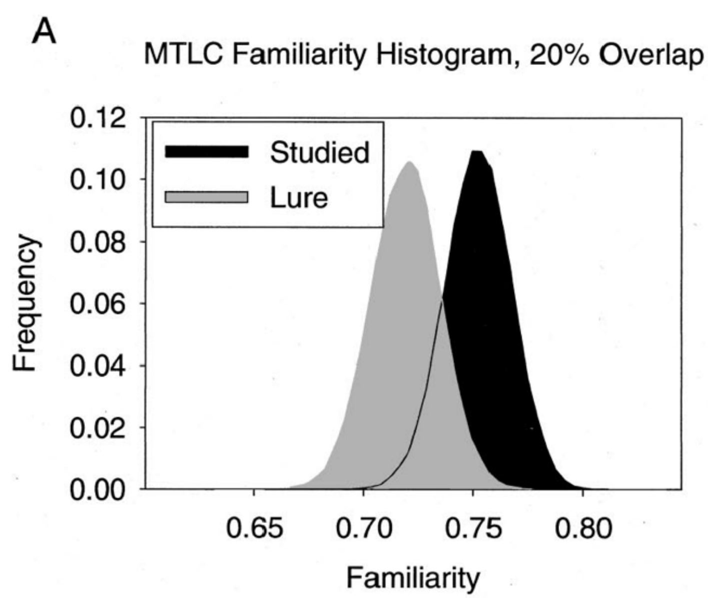

CA3
CA1
DG
EC_in
EC_out
Input

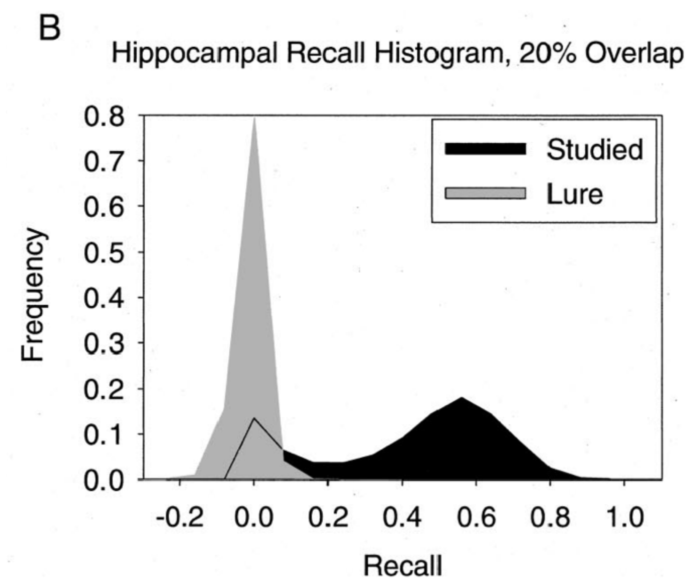
+

Norman and O'Reilly (2003)

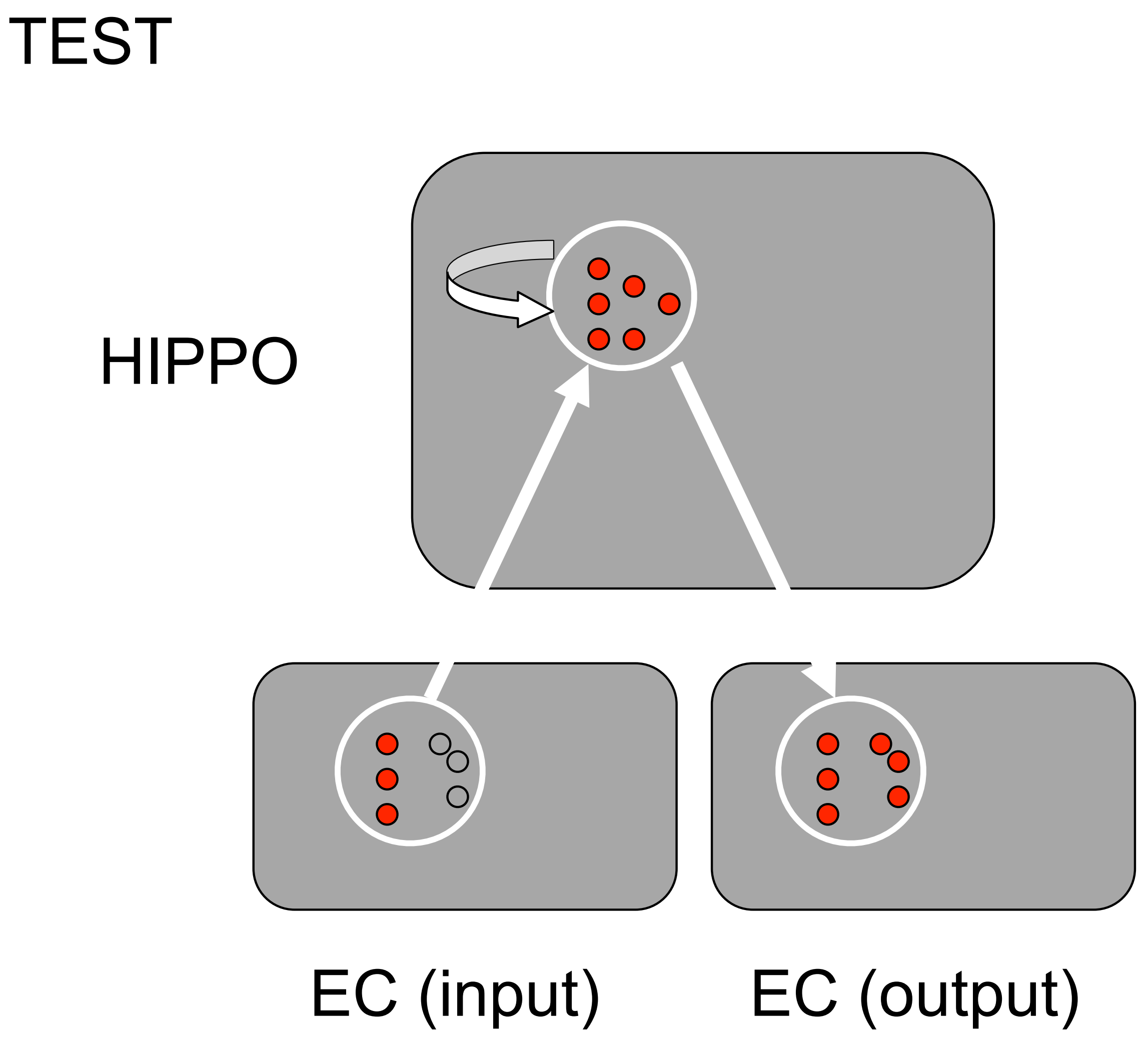




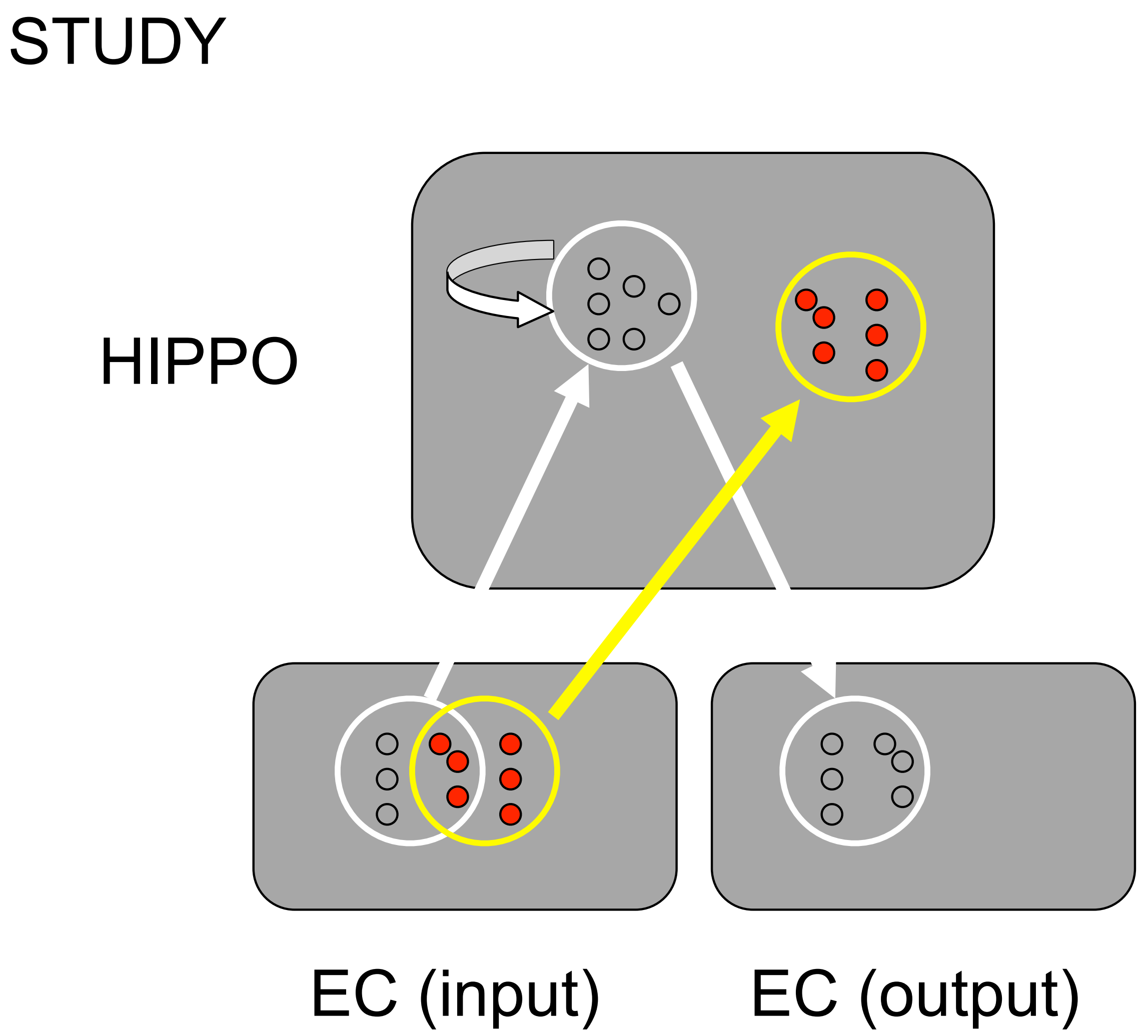




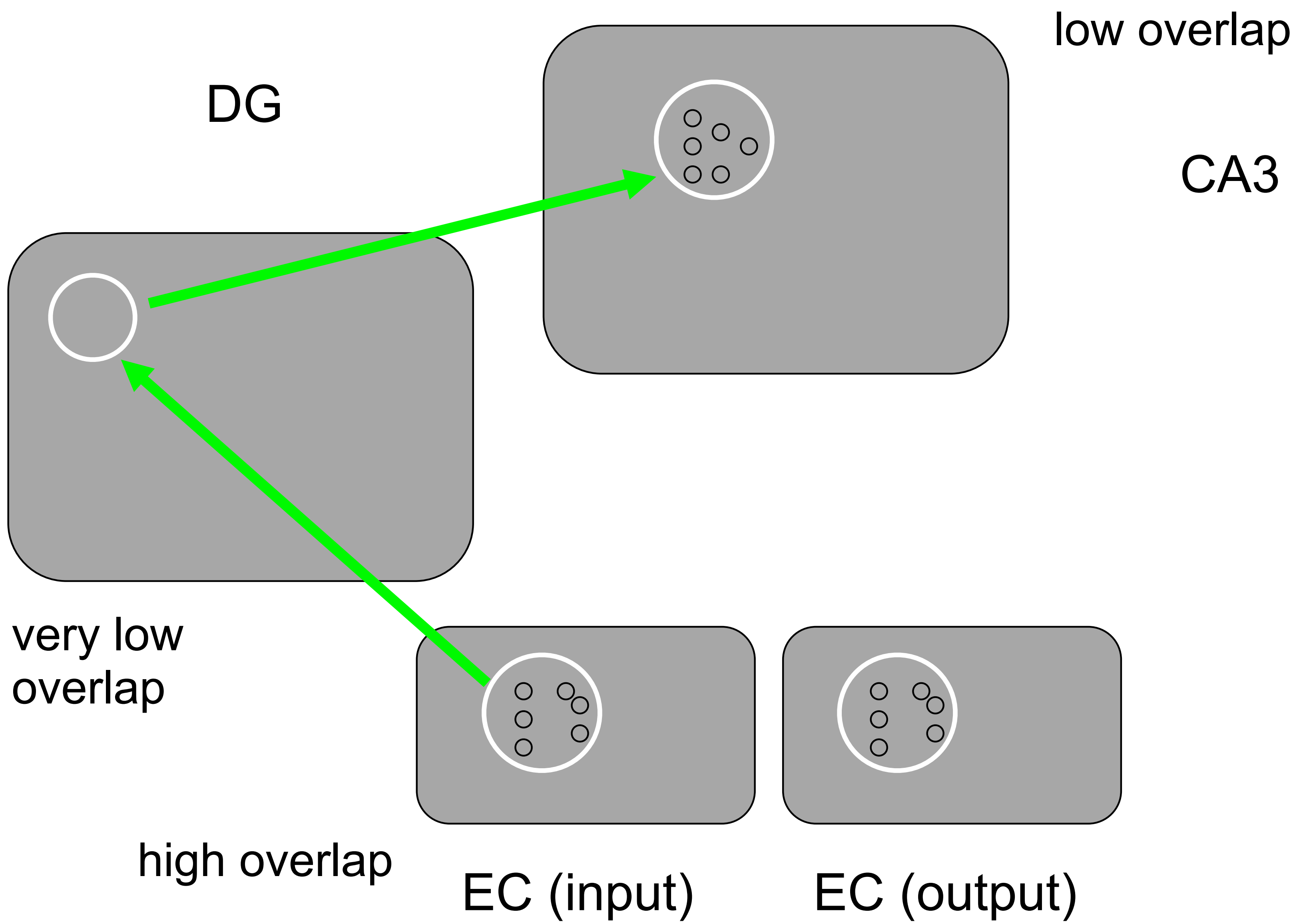
Pattern Completion



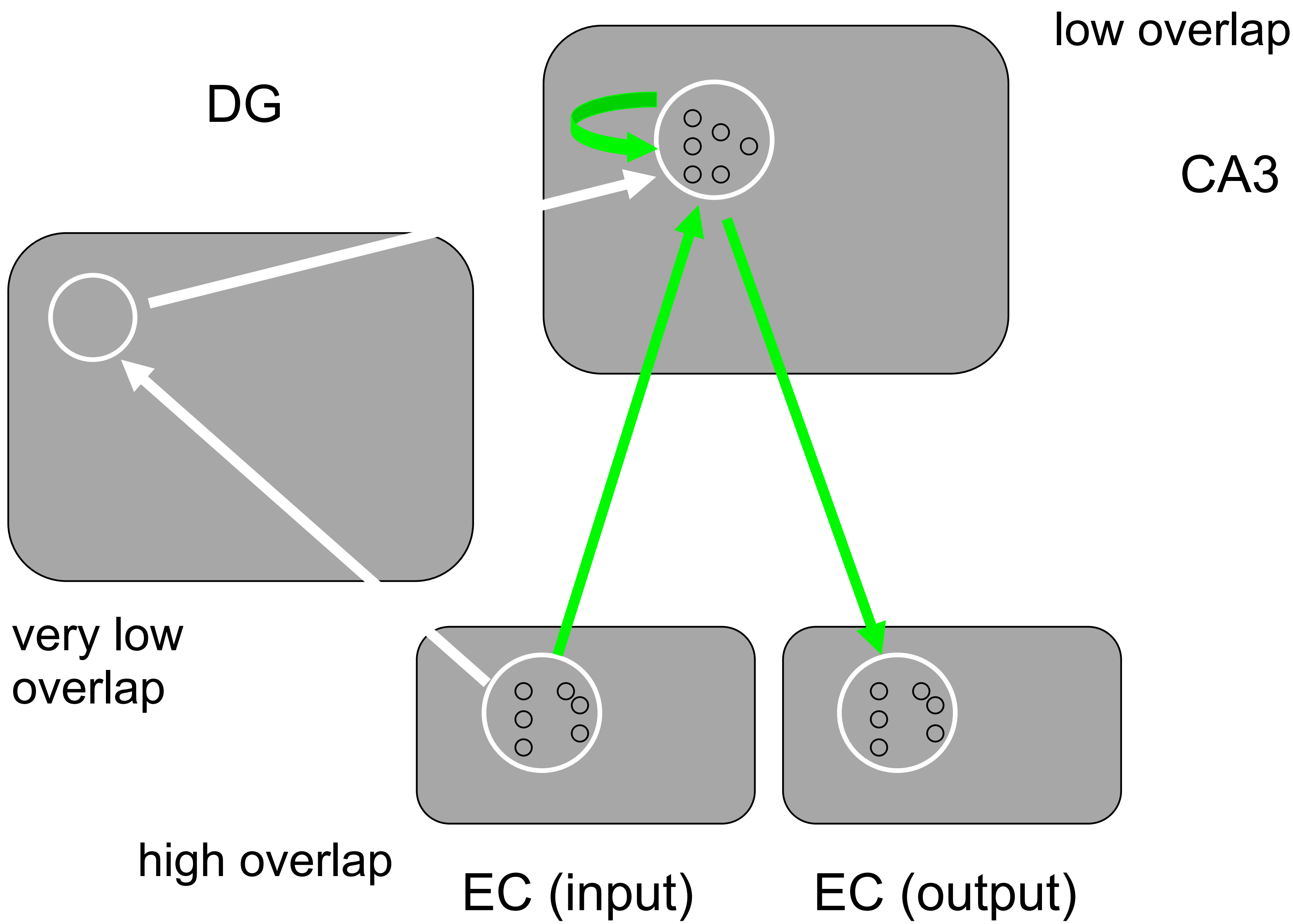
Pattern Separation



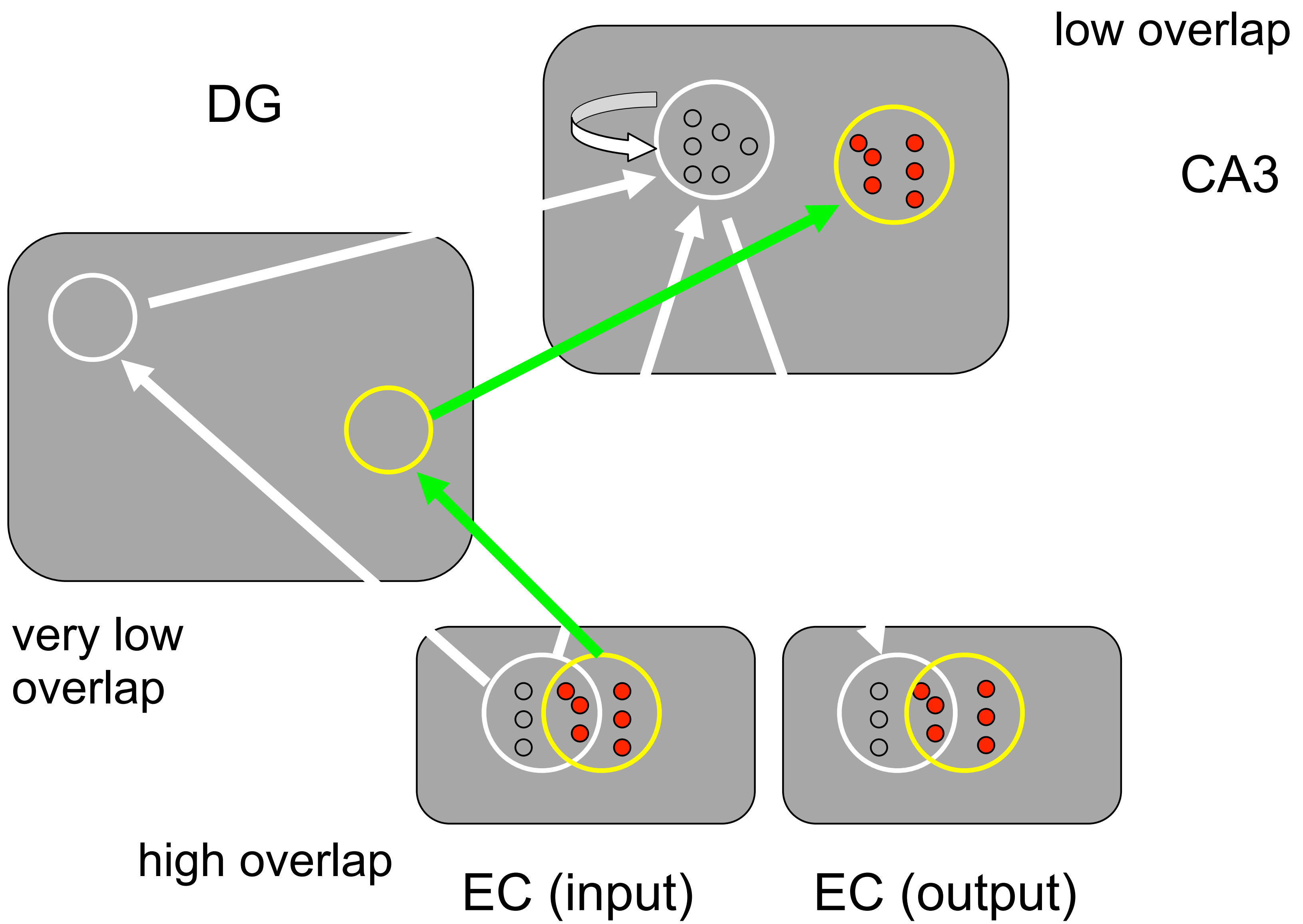
DG = Pattern Separation Turbocharger



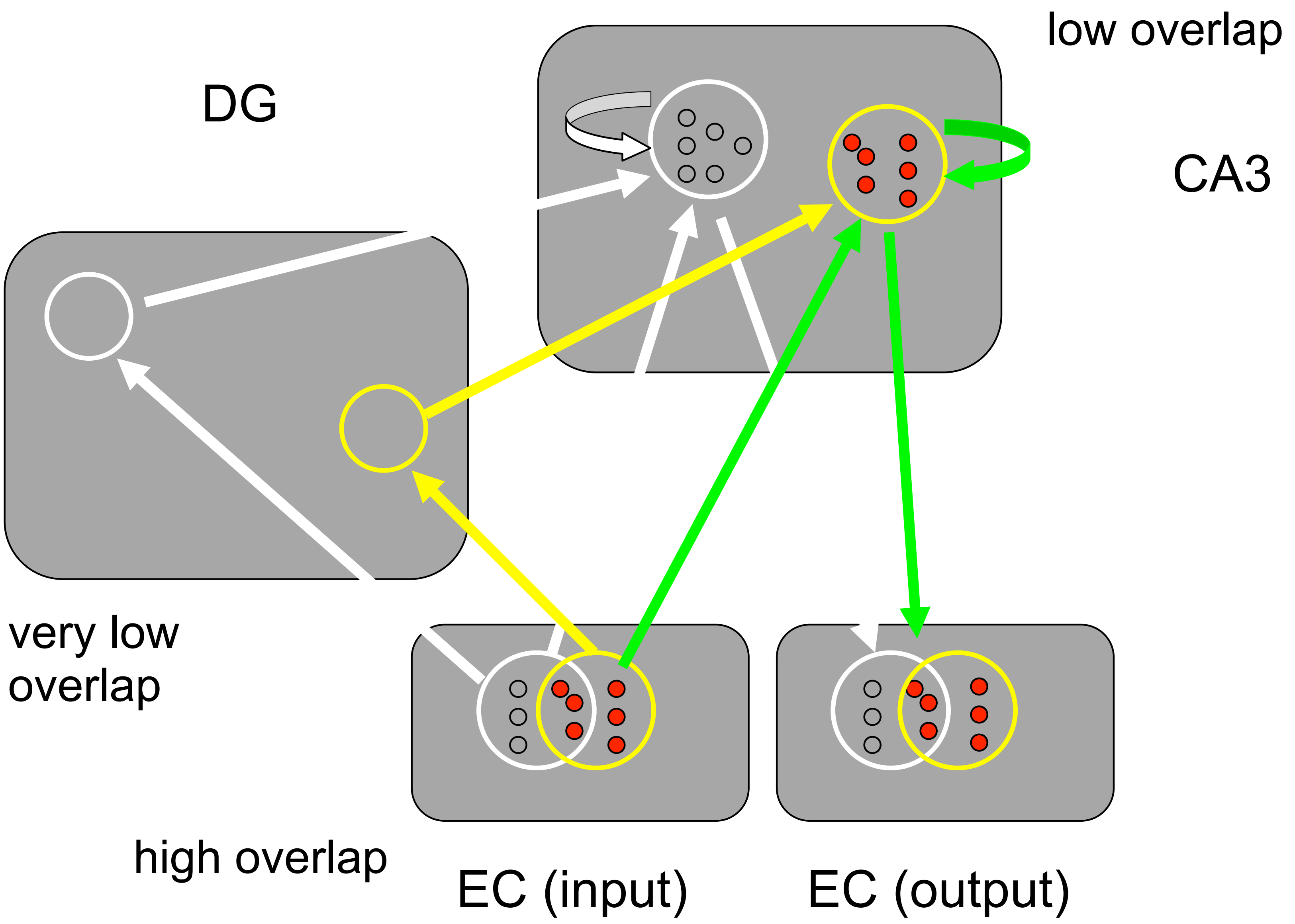
DG = Pattern Separation Turbocharger



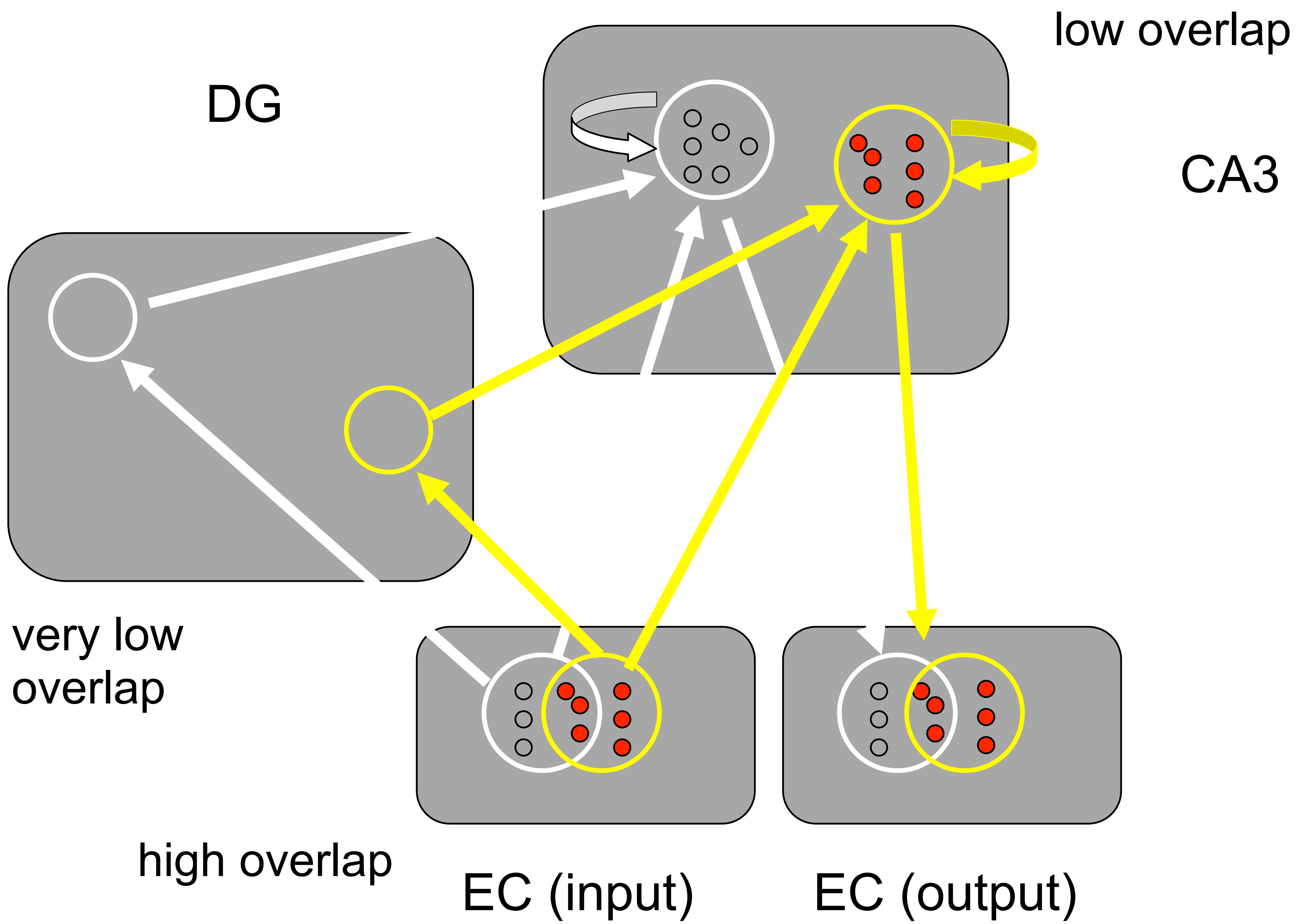
DG = Pattern Separation Turbocharger



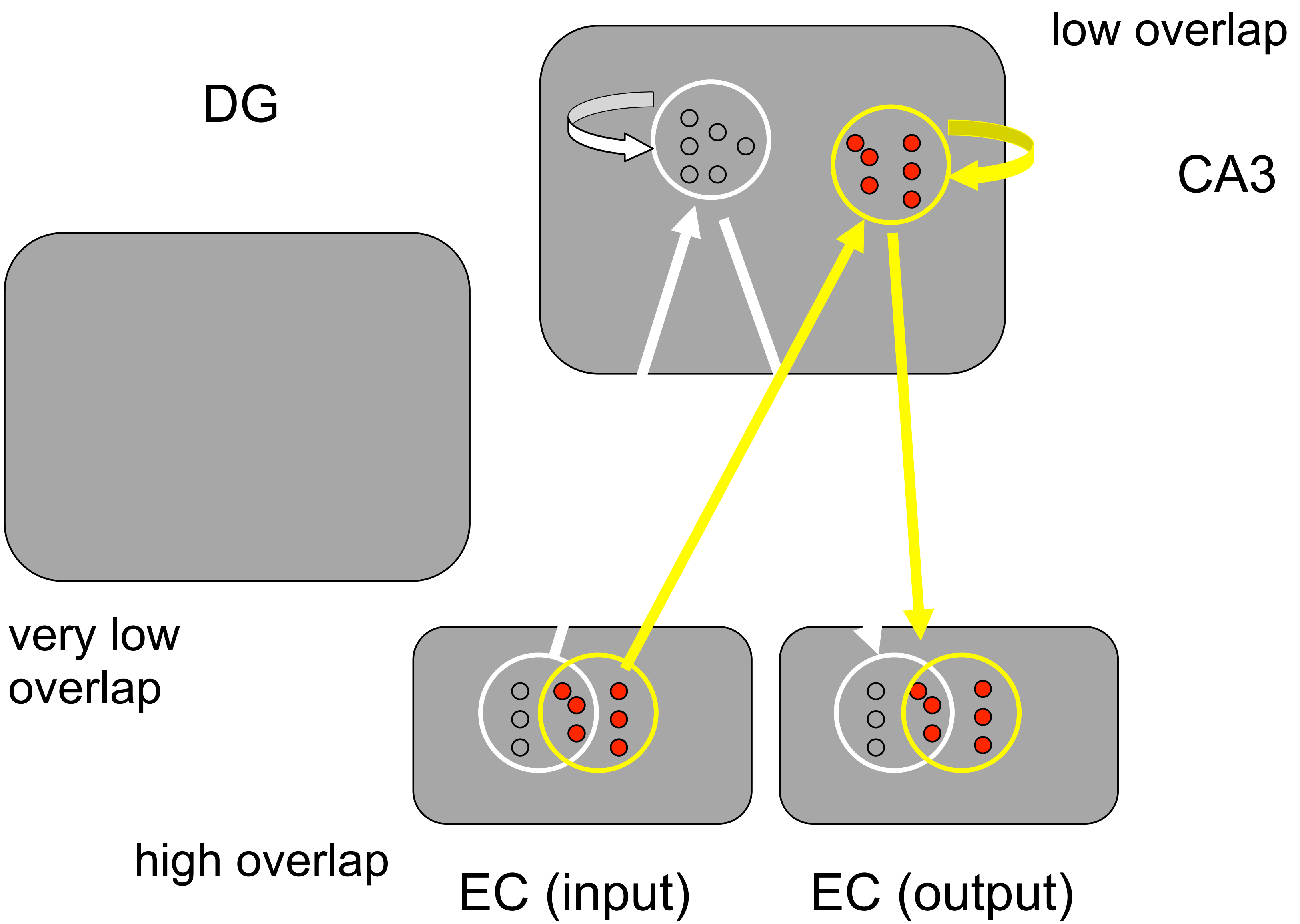
DG = Pattern Separation Turbocharger



DG = Pattern Separation Turbocharger



DG = Pattern Separation Turbocharger



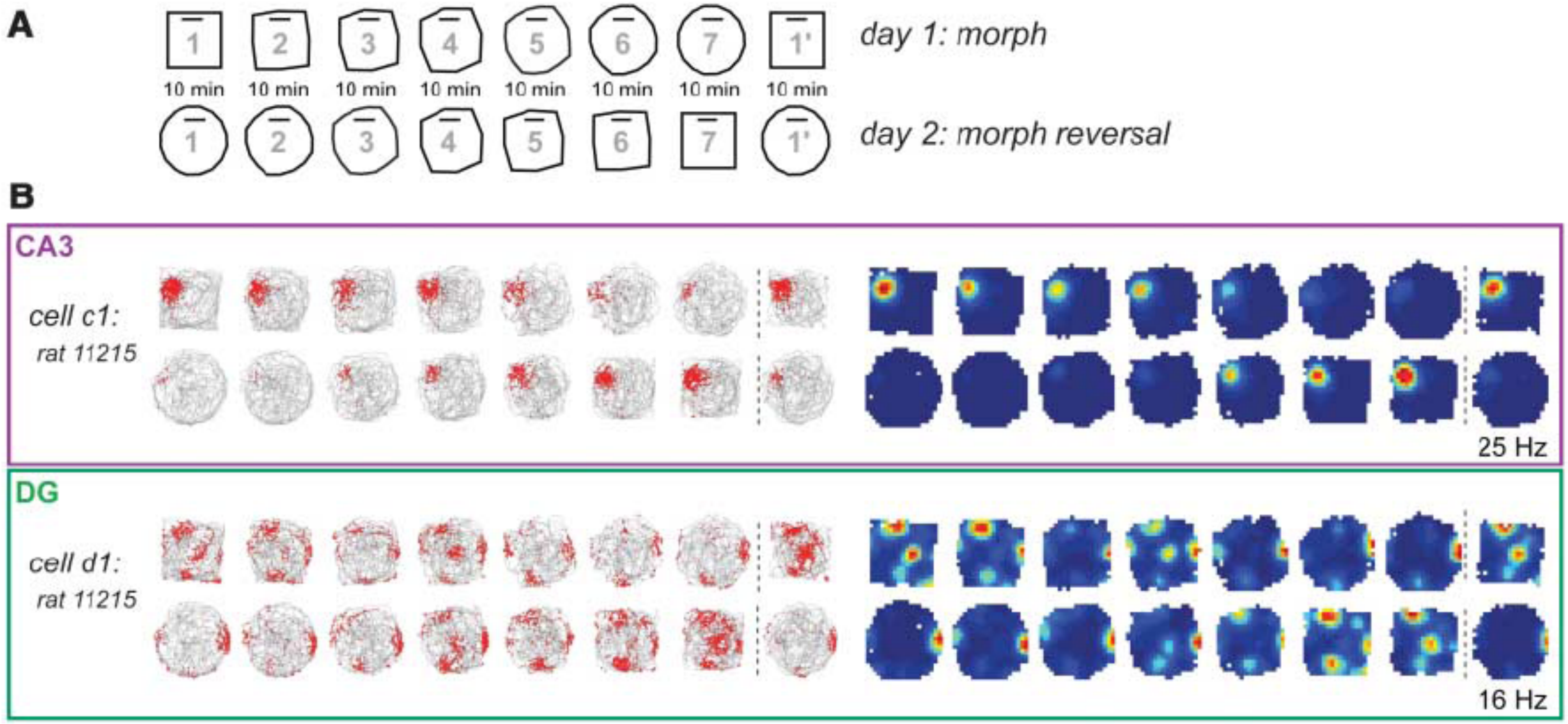
Evidence for DG Pattern Separation

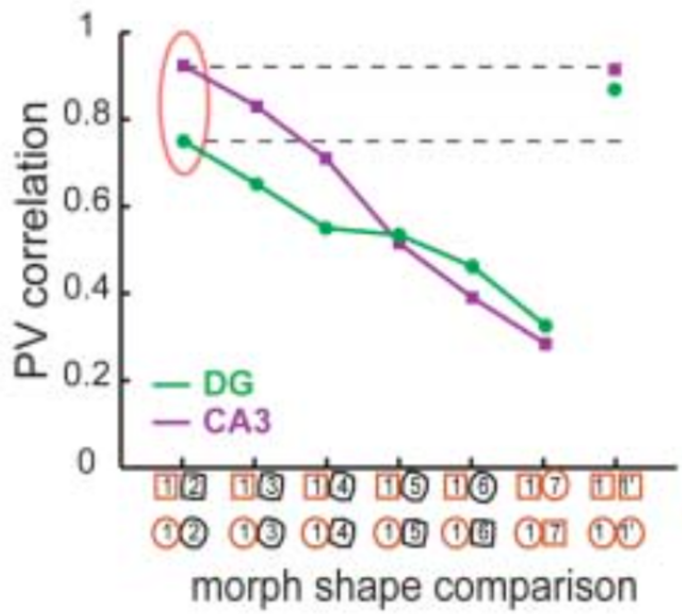
Recently, researchers have started to look for direct evidence that DG is involved in pattern separation

Leutgeb et al. (2007, *Science*)

- Leutgeb et al. (2007) manipulated the shape of a rat cage and recorded hippocampal activity in the dentate gyrus and region CA3

- Prediction: The idea that dentate gyrus is especially prone to pattern separation implies that small changes in the input will have a **larger effect** on dentate gyrus activity than CA3 activity


A
day 1: morph
day 2: morph reversal
10 min
B
CA3
cell c1:
rat 11215
25 Hz
DG
cell d1:
rat 11215
16 Hz


PV correlation
DG
CA3
morph shape comparison

Bakker et al. (2008, *Science*)

- Bakker et al. used high-resolution fMRI to study pattern separation in different hippocampal subregions
- Logic of the study is based on **repetition suppression** effects in fMRI data

Bakker et al. (2008, *Science*)

- If the same stimulus is presented twice, the fMRI response is smaller on the second presentation

- If **very different** stimuli are presented, the fMRI response stays constant

- This implies that we can use the **size of the decrease in the fMRI response** as an indicator of whether a particular brain area thinks that two stimuli are the **same** or **different**

- If the two stimuli are coded in a **similar** fashion, we should see a decrease in the fMRI response

- If the two stimuli are coded **very differently**, the fMRI response should stay relatively constant

Bakker et al. (2008, *Science*)

- In the Bakker et al. study, subjects performed a continuous recognition memory task
- Sometimes they presented lures that were very similar (but not identical) to studied items

Fig. 1. Sample stimuli sets showing versions A and B of the same object.

Bakker et al. (2008, *Science*)

- Predictions:

- Most brain regions should assign **very similar** representations to studied items and similar lures

- For these brain regions, we should see repetition attenuation (decreased responding to similar lures)

- DG should assign **relatively dissimilar** representations to studied items and similar lures

- For these brain regions, we should see **less** repetition attenuation (**less of a decrease** in responding to similar lures)

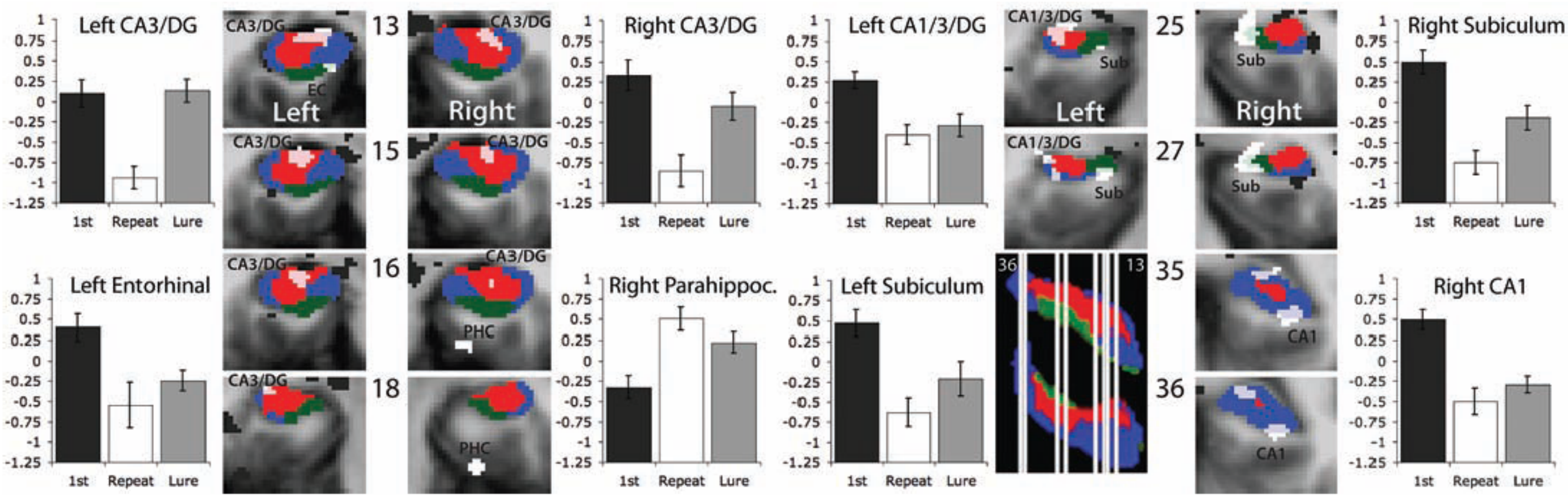
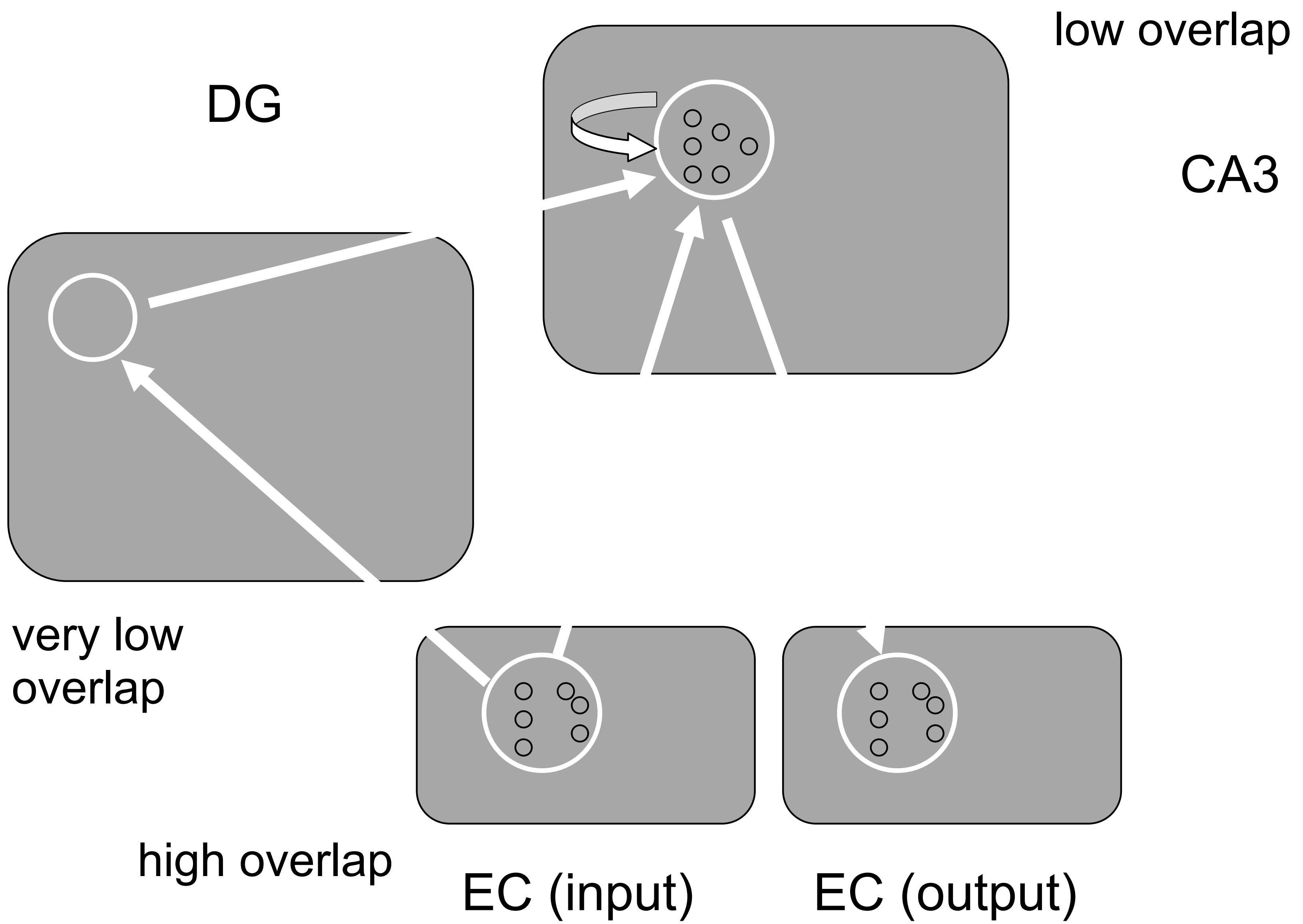
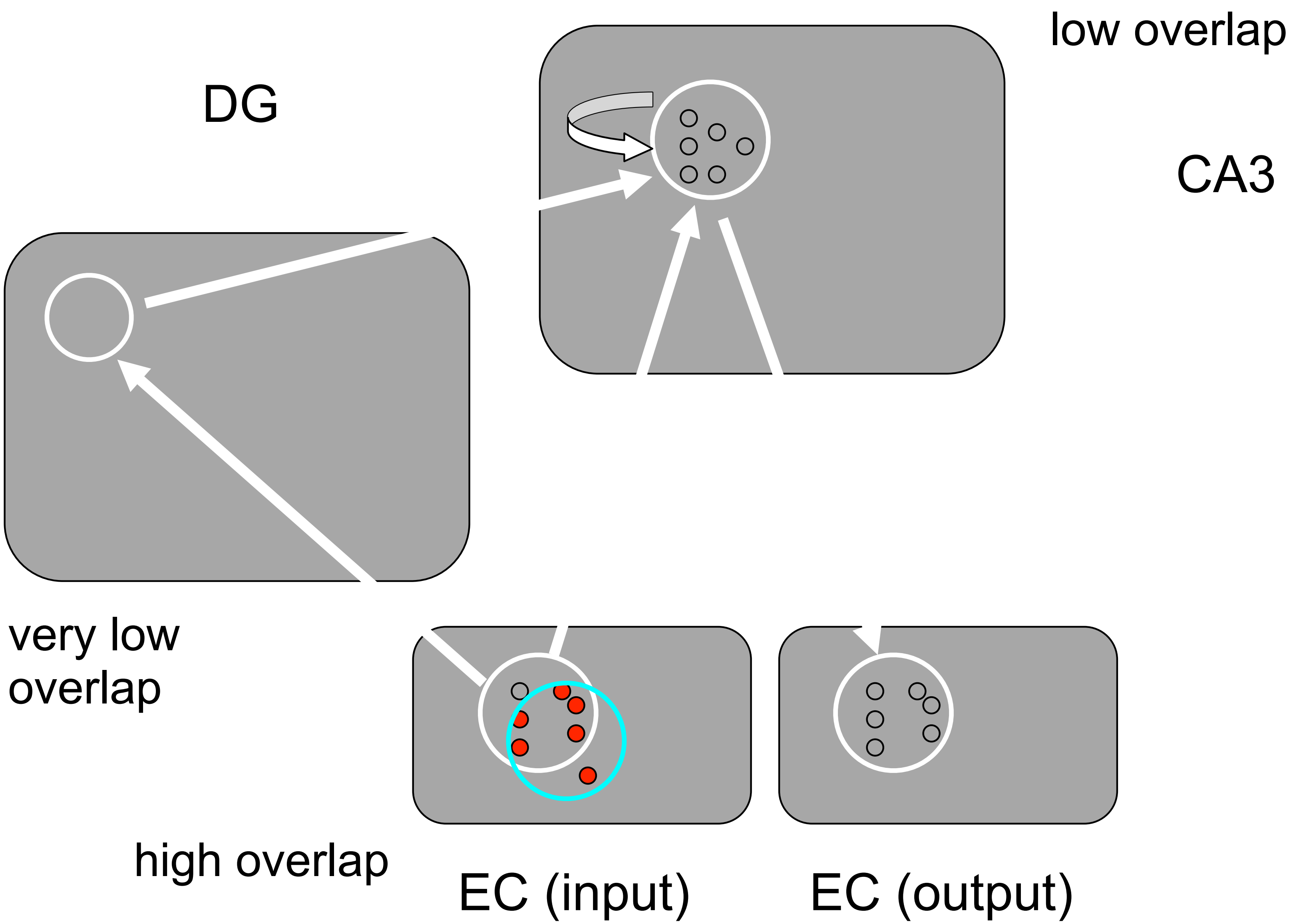


Fig. 2. Anatomical location and mean activity in the three task conditions for each of the eight MTL regions of interest. A model segmentation of hippocampal subfields is overlaid on each brain slice to indicate the location of the subiculum (green), CA1 (blue), and CA3/DG (red). Regions of activity within the MTL are shown in white and labeled within each slice (PHC, parahippocampal cortex; EC, entorhinal cortex). Regions of activity outside the MTL (not part of the analysis) are shown in black. Three-dimensional rendering (lateral, superior view) shows the location of each slice (white lines). The distance of each slice from the anterior commissure ($y = 0$ in Talairach coordinates) is indicated for each slice as well ($y = 13$ to 36 mm). The thicker white lines represent two adjoining slices. Bar graphs show mean activity (summed beta coefficients) in each ROI for each trial condition.

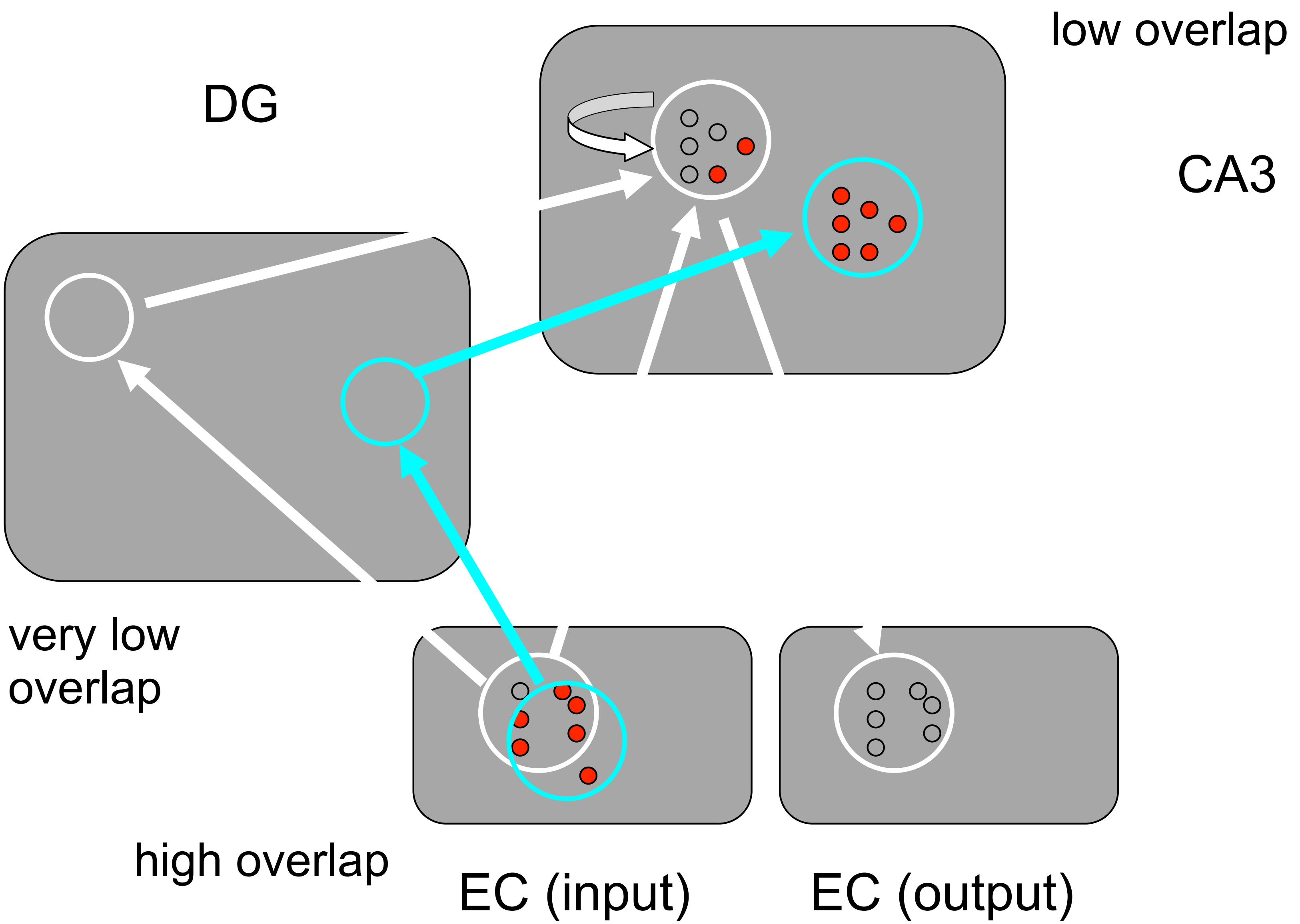
DG is good for pattern separation, but it's bad for recall



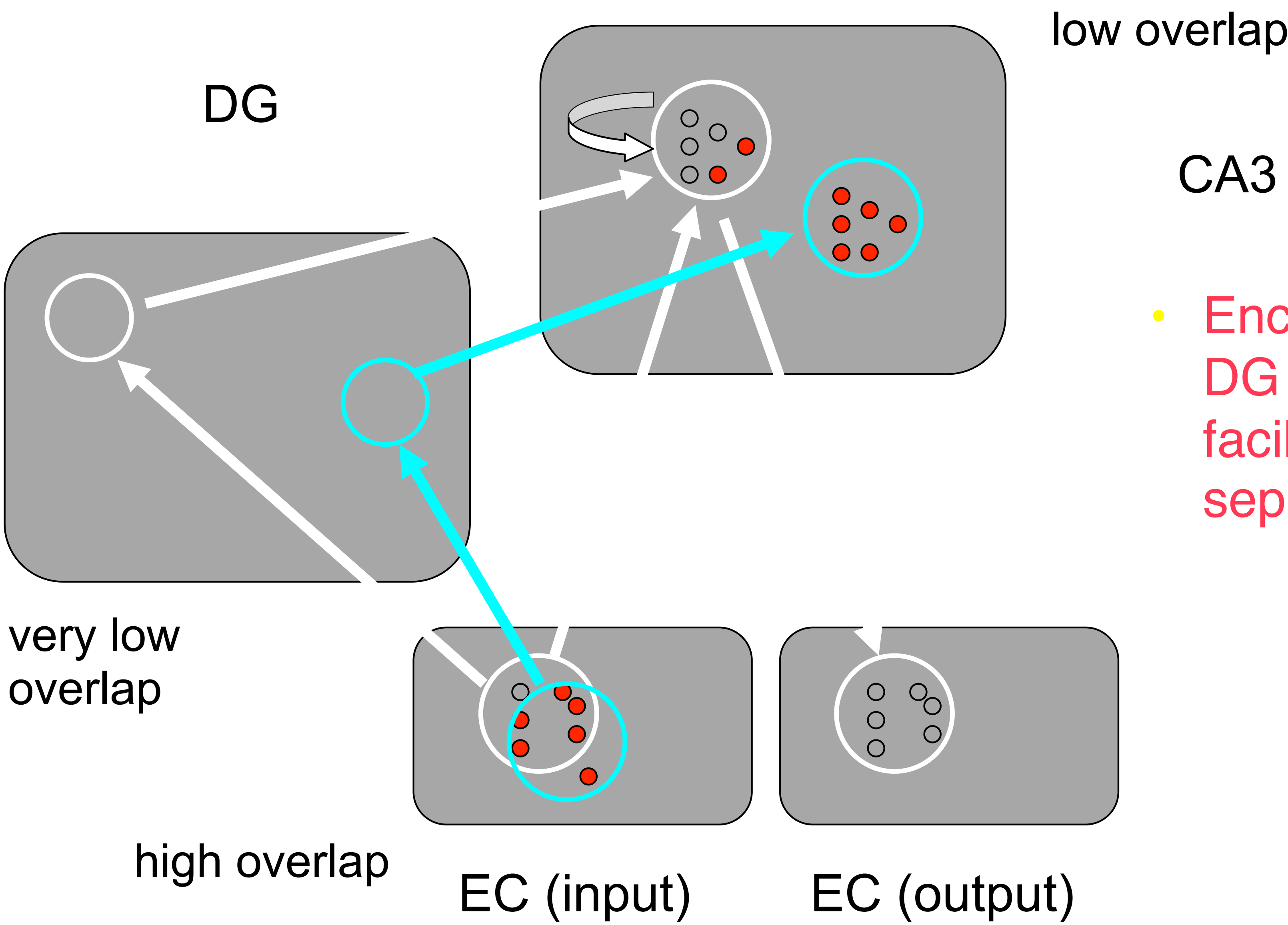
+ DG is good for pattern separation, but it's bad for recall



DG is good for pattern separation, but it's bad for recall

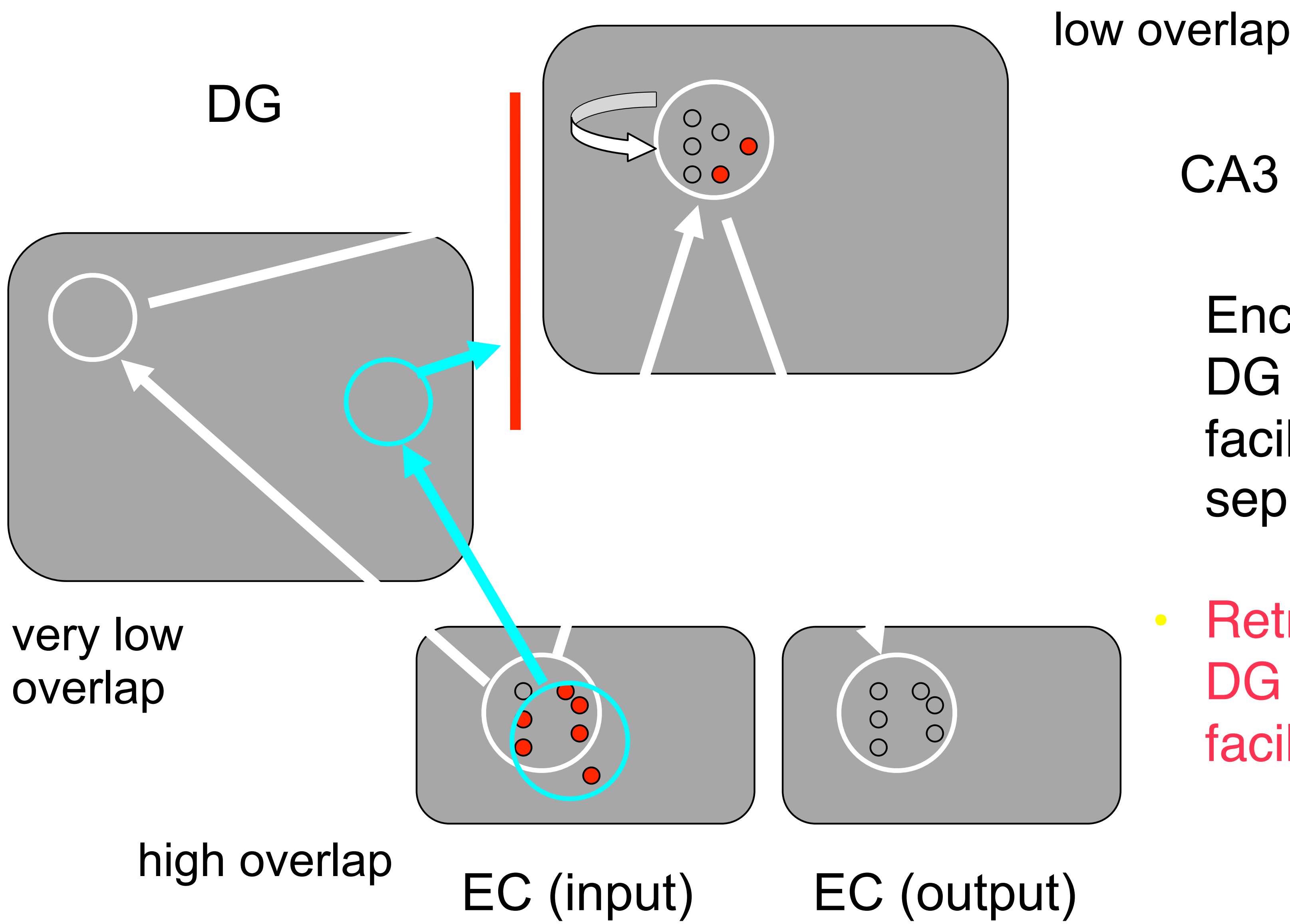


+ Possible Solution: Two Modes



- Encoding mode: DG strong, facilitates pattern separation

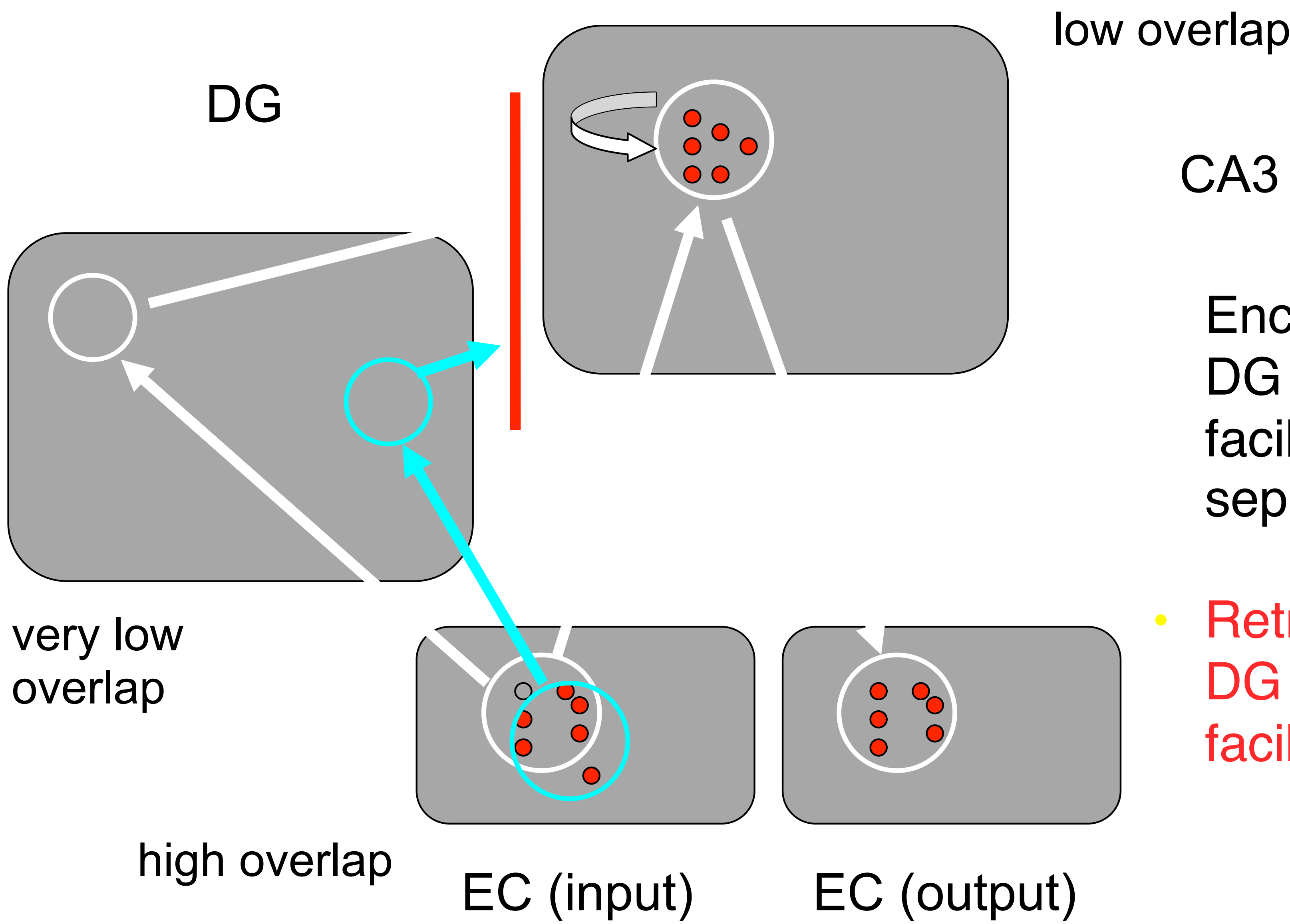
Possible Solution: Two Modes



Encoding mode: DG strong, facilitates pattern separation

- Retrieval mode: DG weak, facilitates recall

Possible Solution: Two Modes



Encoding mode: DG strong, facilitates pattern separation

- Retrieval mode: DG weak, facilitates recall

Evidence for Mode-Setting

There is evidence that acetylcholine might adjust the strengths of hippocampal pathways so as to facilitate encoding (as opposed to retrieval)

For example, ACh blocks transmission along CA3 recurrents while still allowing learning in these synapses: Good for encoding

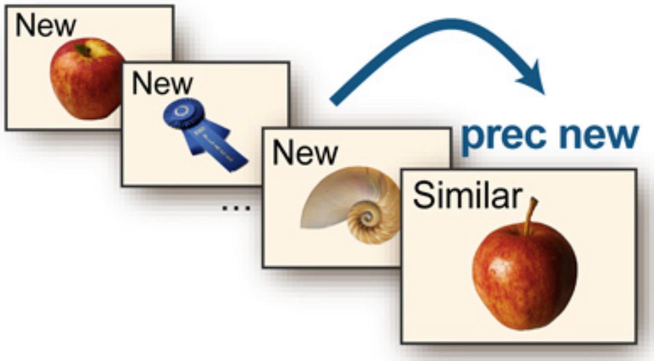
Low ACh facilitates transmission along CA3 recurrents: Good for retrieval

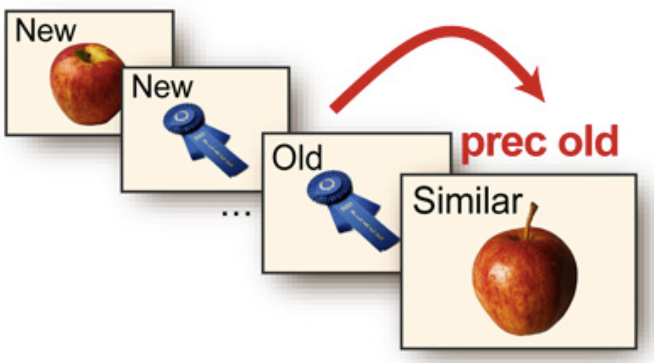
Evidence for Mode-Setting

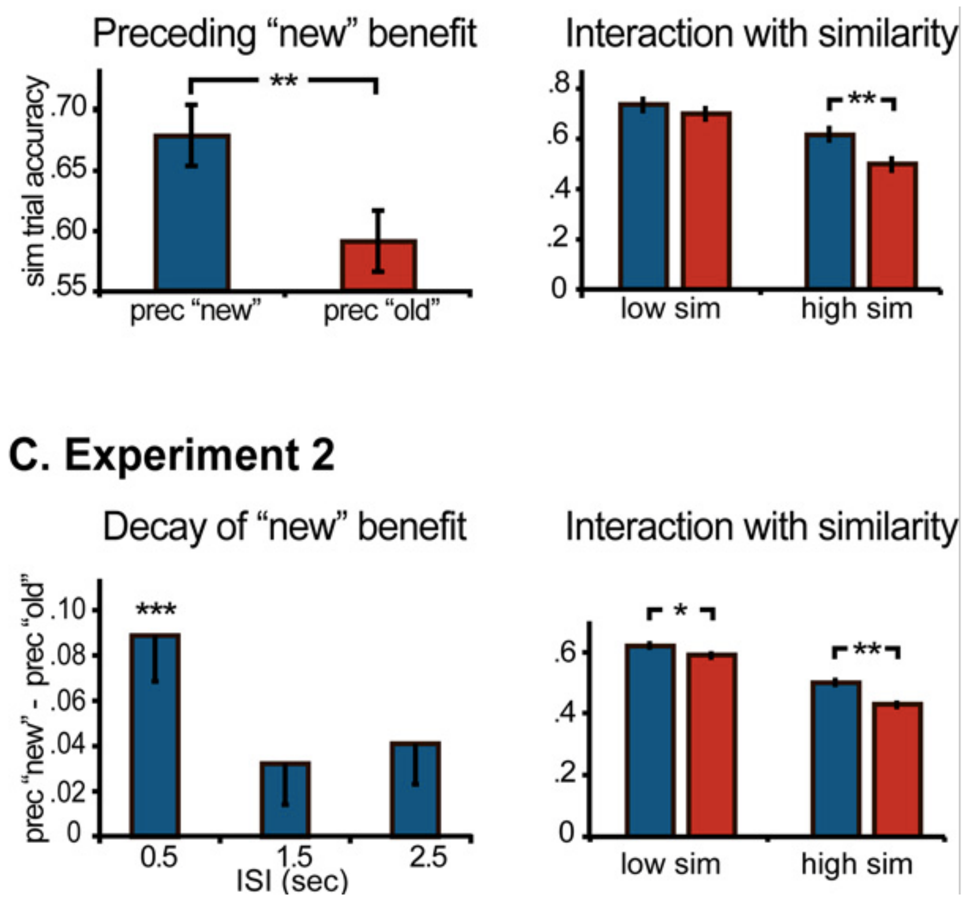
Hasselmo & Wyble (1997) set forth a detailed story about how novel stimuli could trigger ACh and flip the hippocampus into encoding mode

Note that the time course of ACh release is relatively slow...

Evidence for mnemonic ‘modes’?

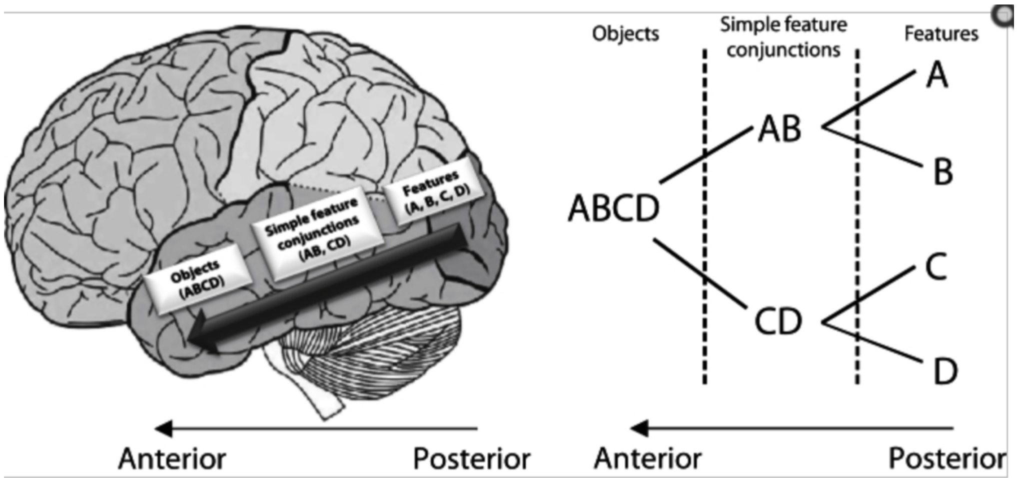


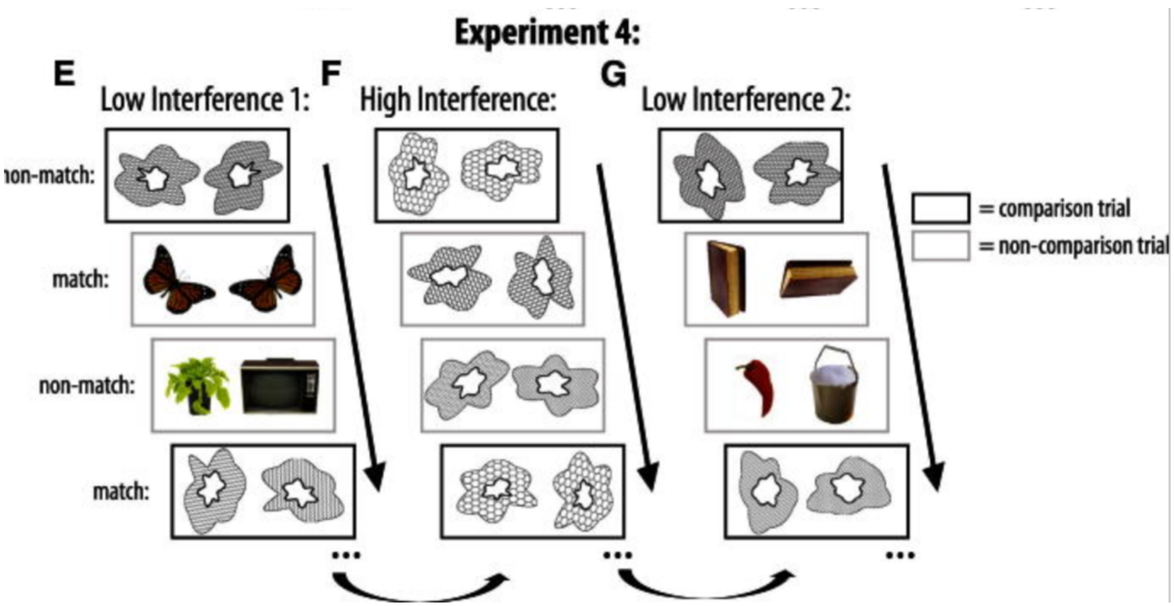


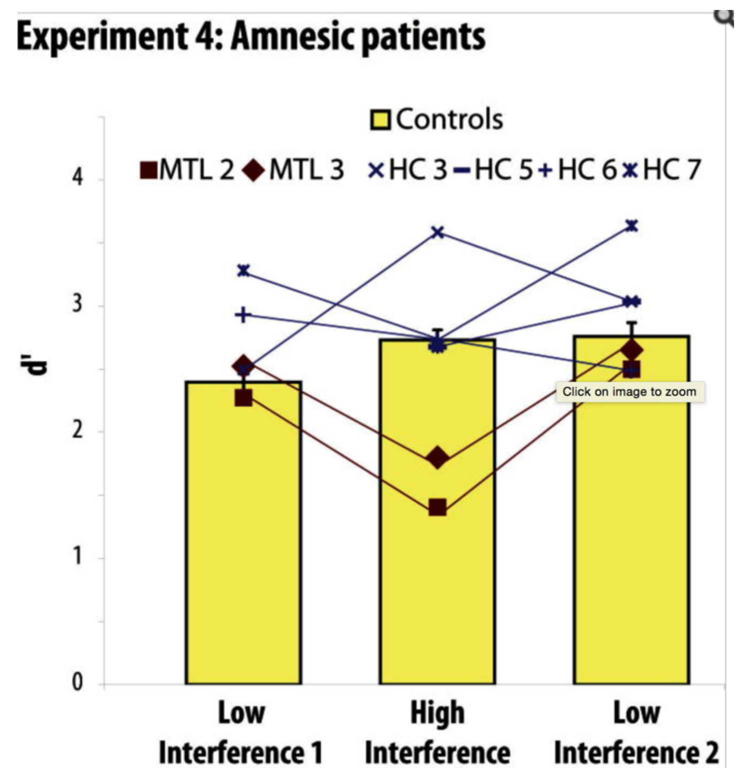


(Duncan, Sadanand and Davachi, 2012 Science)

+ Barense et al 2012







The take-home

- Computations supported by the hippocampus and cortical structures appear fundamentally distinct
- Recollection and familiarity are likely independent processes that support memory function
- Interplay between empirical findings in animal and human research with modeling has played a big role in memory research
- Has led to a focus on the computations performed in the brain - not so much on identifying a ‘memory system’ or a ‘perceptual system’

Neurogenesis in the Dentate Gyrus

- The dentate gyrus site of **neurogenesis**
-

- Hypothesis: New neurons help to “randomize” the dentate gyrus response, thereby ensuring that similar stimuli activate different sets of dentate gyrus neurons

- The idea that neurogenesis in the dentate gyrus **helps** pattern separation implies that blocking neurogenesis will **hurt** pattern separation

Clelland Et Al. (2009, *Science*)

- Clelland et al. (2009) set out to explore how disrupting neurogenesis in the dentate gyrus affects learning

- They used a task where animals visited a place, and then they had to discriminate between the previously visited place and another place

- In some conditions, the animals had to discriminate between **nearby** places, and in other conditions the animals had to discriminate between **far away** places

Clelland Et Al. (2009, *Science*)

- Key idea: Hippocampal pattern separation is important for discriminating between **nearby** places but it is not needed to discriminate between **far away** places

- Prediction:

- IF disrupting neurogenesis impairs pattern separation
- AND pattern separation is important for discriminating between **nearby** places
- THEN disrupting neurogenesis should impair discrimination between **nearby** places but it should not disrupt discrimination between **far away** places

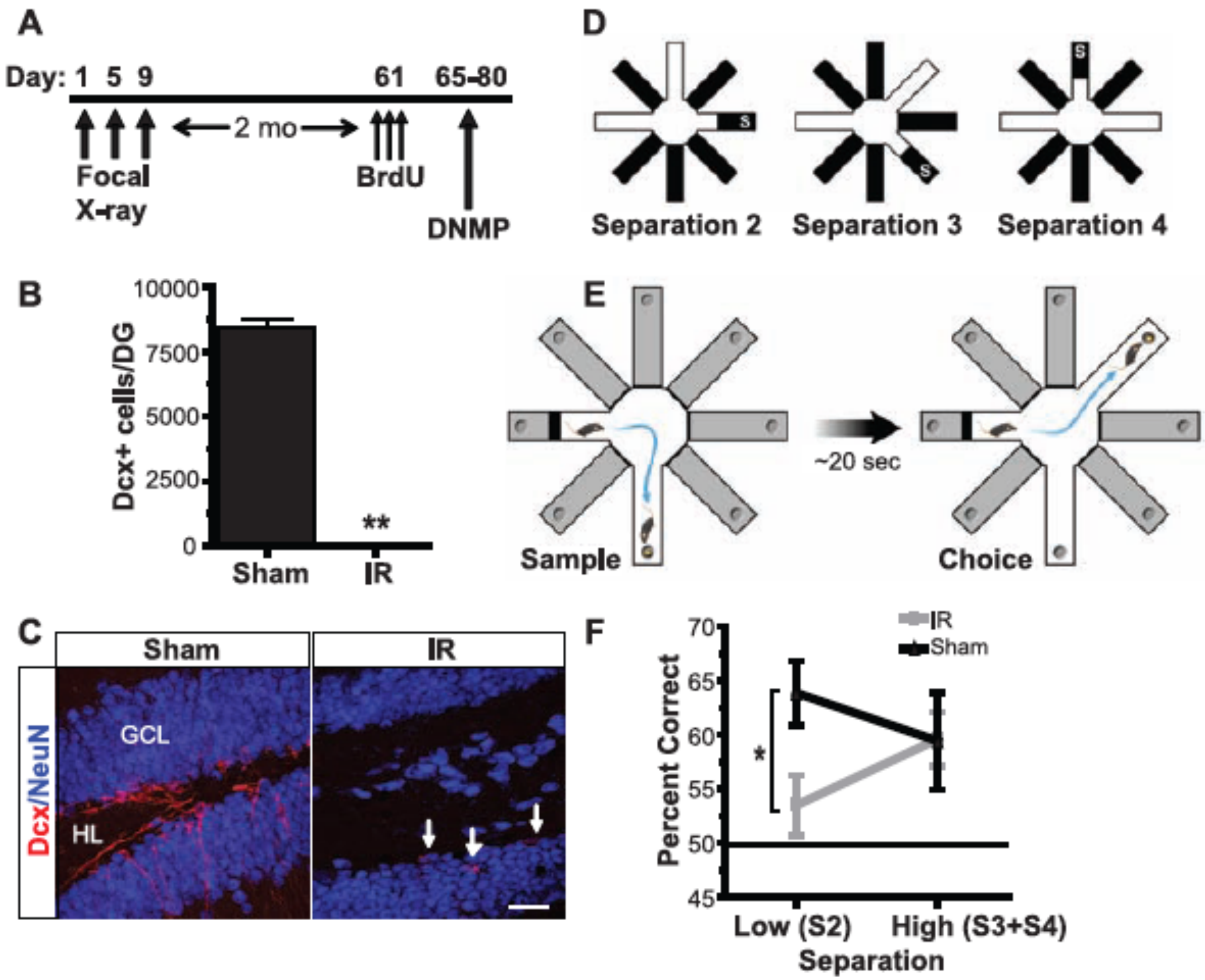


Fig. 1. Mice with ablated neurogenesis due to focal x-irradiation show impaired spatial memory for similar, but not distinct, spatial locations in the radial arm maze. (**A**) Mice were irradiated 2 months before behavioral testing. (**B** and **C**) Irradiation significantly reduced the total numbers of immature Dcx+ cells in IR mice [white arrows at right in (C)] compared with sham controls [(C), left] [independent samples *t* test, $t(17) = 29.82$, $P < 0.001$]. GCL, granule cell layer; HL, hilus. (**D**) Pattern separation was tested using a DNMP protocol in the RAM by varying the distance between sample and correct arms: S2, low; S3 and S4, high (S, start arm). (**E**) Each trial consisted of a sample phase (left) and a choice phase (right). The mouse had to nonmatch to the new location. (**F**) IR mice were impaired at low (S2) but not high (S3 and S4) separations in the DNMP task. The horizontal black line represents chance. Error bars indicate SEM. Scale bars, 25 μm. $^{**}P < 0.01$; $^{*}P < 0.05$.

Summary

Key Principals for the Semester

- **Learning and memory are closely related and intertwined states of information processing**
- **Major insights about learning and memory have come from studies of the brain**
- **The concept of multiple memory systems unifies the study of learning and memory**
- **The underlying bases of learning and memory are the same in humans and animals**
- **Our theoretical approaches to studying learning are always closely tied to technological advances that are unfolding in general society (e.g., today - machine learning)**