Computational modelling of synaptic plasticity and learning in brains

Dr. Cian O'Donnell

c.odonnell2@ulster.ac.uk

@cianodonnell.bsky.social

https://odonnellgroup.github.io

Why study learning?

KordingLab 👨‍💻🧠∇🔬📈,🏋️⛷️🏂🛹🕺⛰️☕🦖
@KordingLab

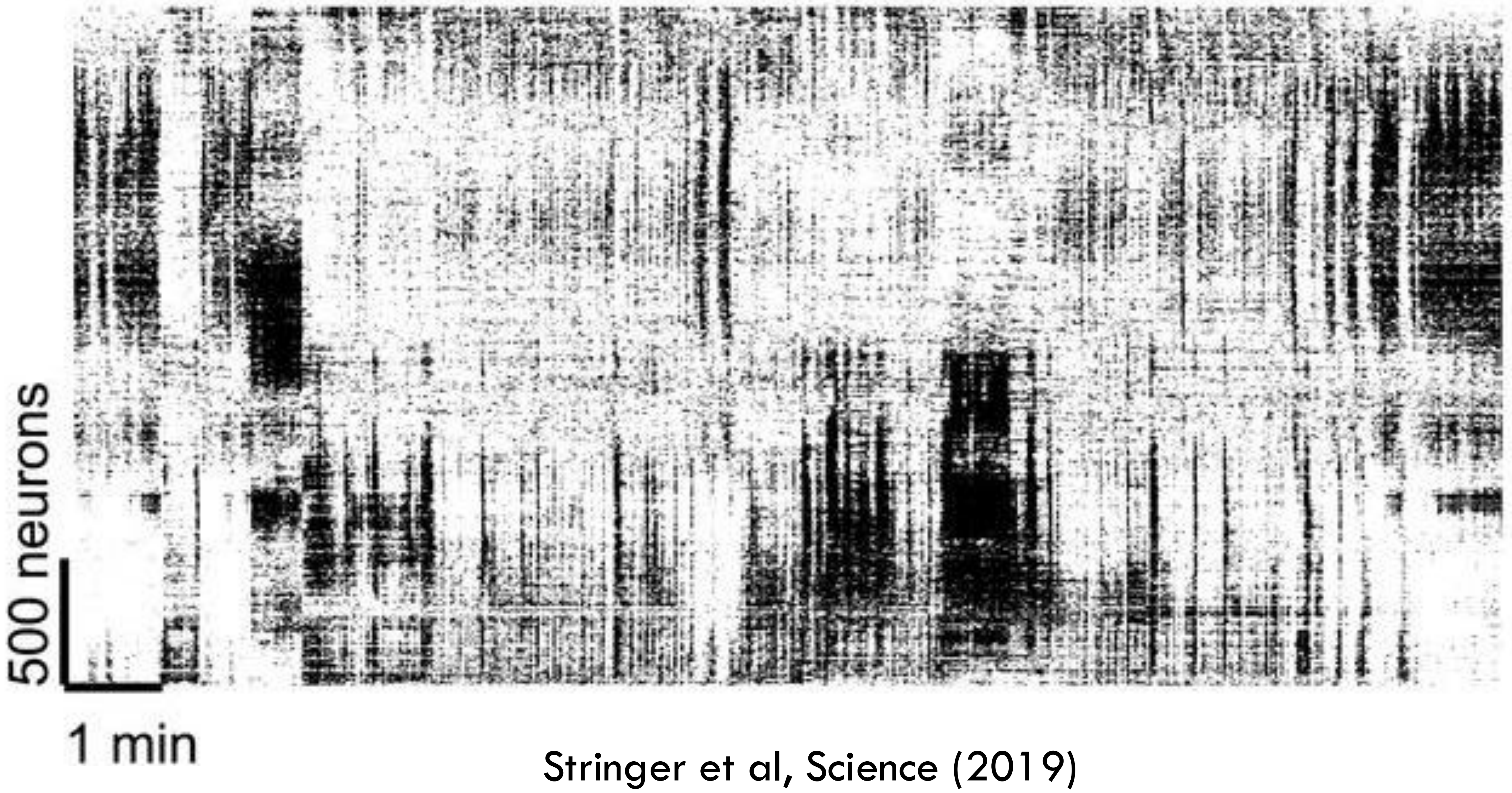
In computational neuroscience can we realistically understand computation (say how we see a hand) or can we at best hope to understand the learning that gives rise to the computation?

4:30 PM · Oct 14, 2021 · Twitter Web App

15 Retweets **5** Quote Tweets **130** Likes

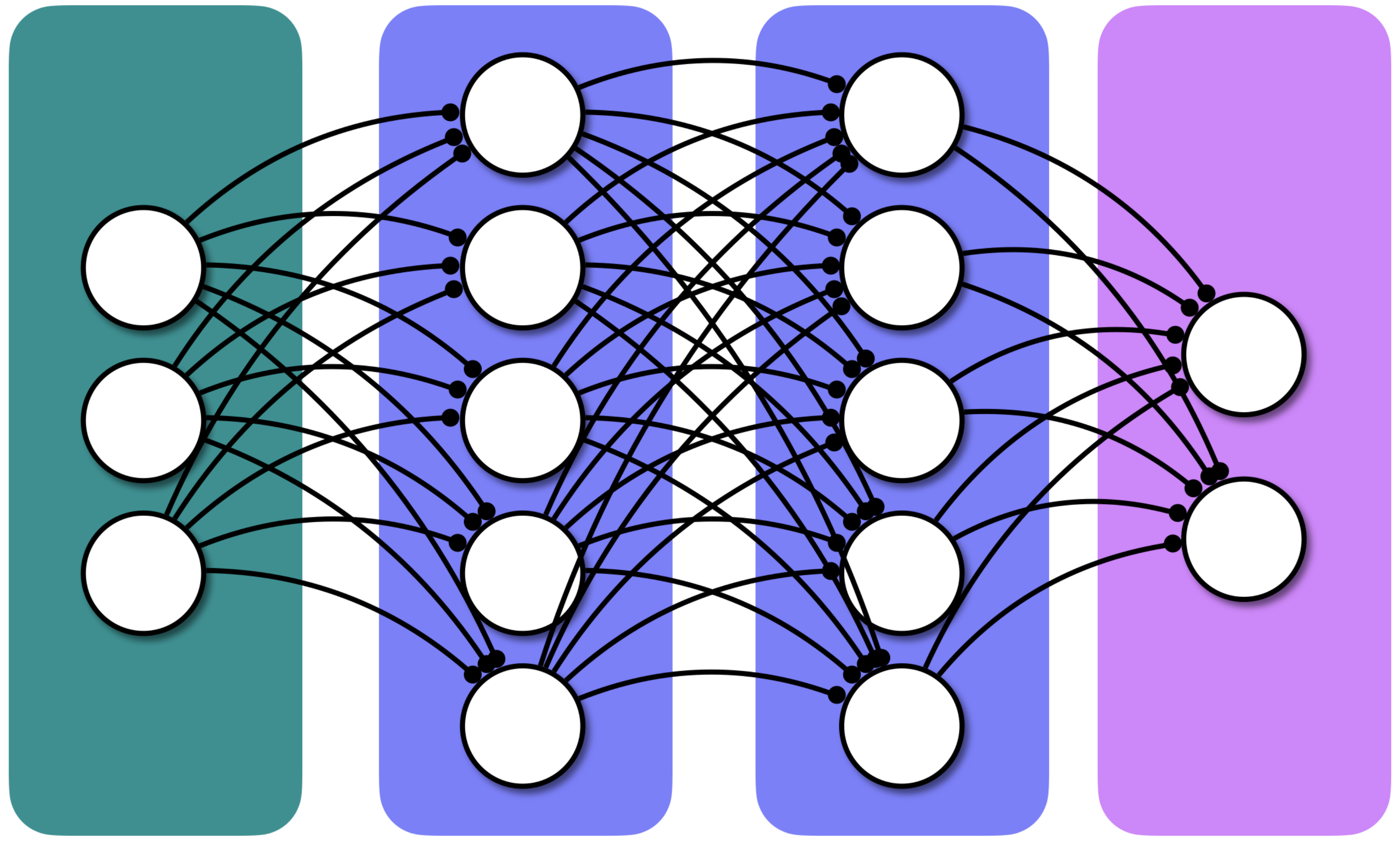
Why study learning?

Population activity in the brain



Stringer et al, Science (2019)

Deep neural network



Why study learning?

One answer:

- Explaining what every single neuron is doing and how they contribute to computations may be impossible.
- But we may hope to someday write down (understand) the rules of brain learning.

What does it mean to understand a neural network?

Lillicrap & Kording, arXiv (2019)

What we will cover

1. Synaptic plasticity

- Classic rate-based models of plasticity, Hebbian learning.
- Spike-timing dependent plasticity.

2. Associative memories and their network models

- The hippocampus
- Feedforward models of hippocampal memory
- Hopfield networks

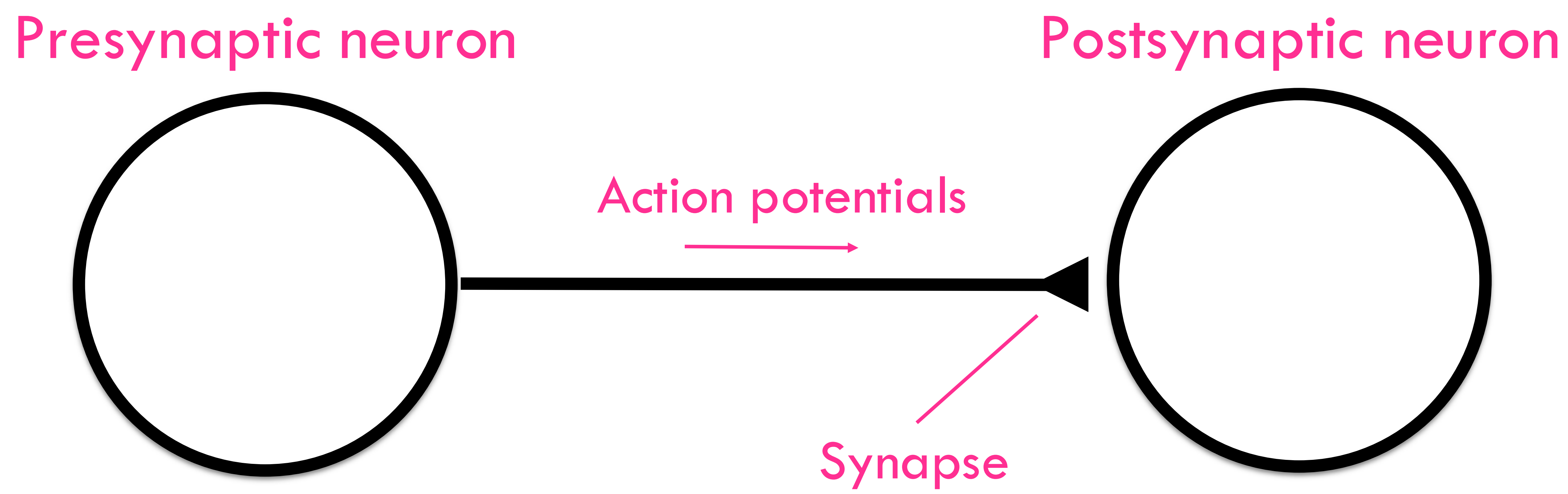
3. Links to artificial neural networks

- The backpropagation rule
- Mapping deep learning to neurobiology

1. Synaptic plasticity

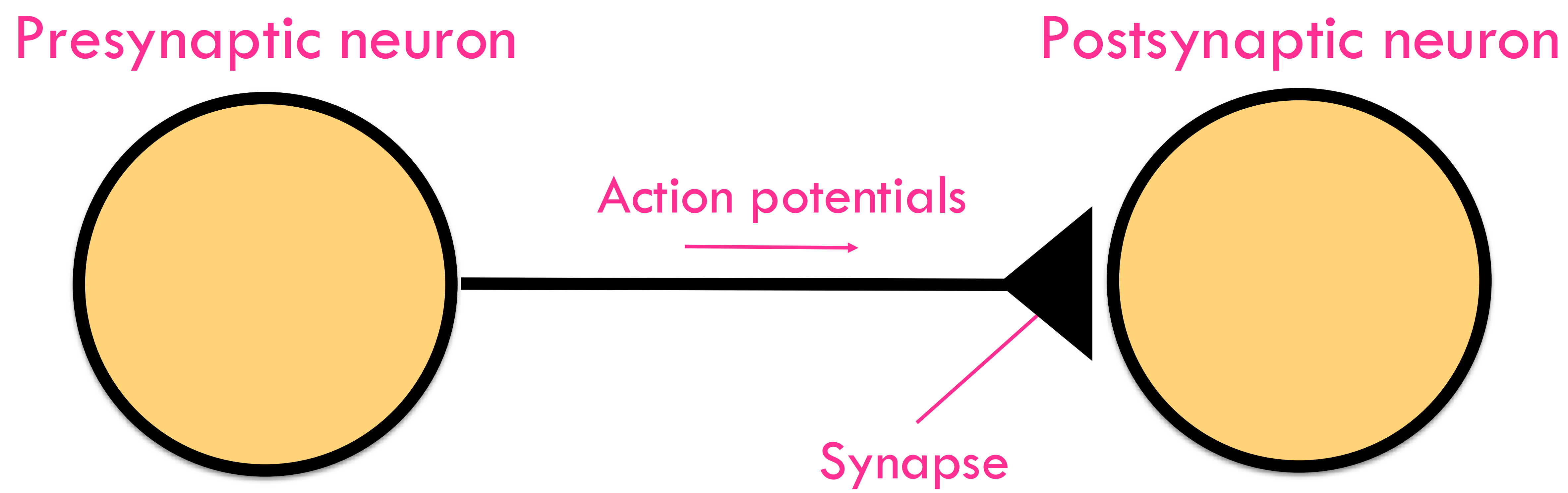
Long-term synaptic plasticity

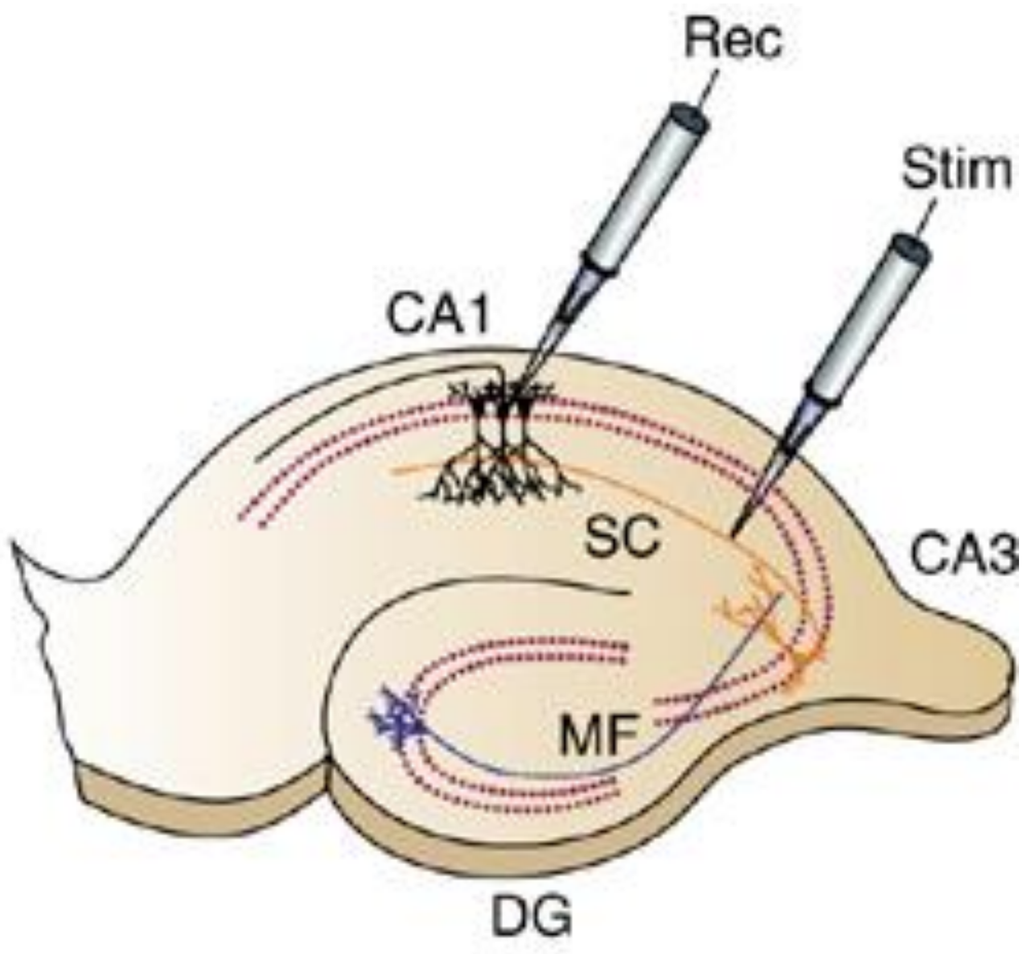
Long-term synaptic plasticity is a (activity-dependent) semi-permanent change in the strength of the connection from one neuron to another.

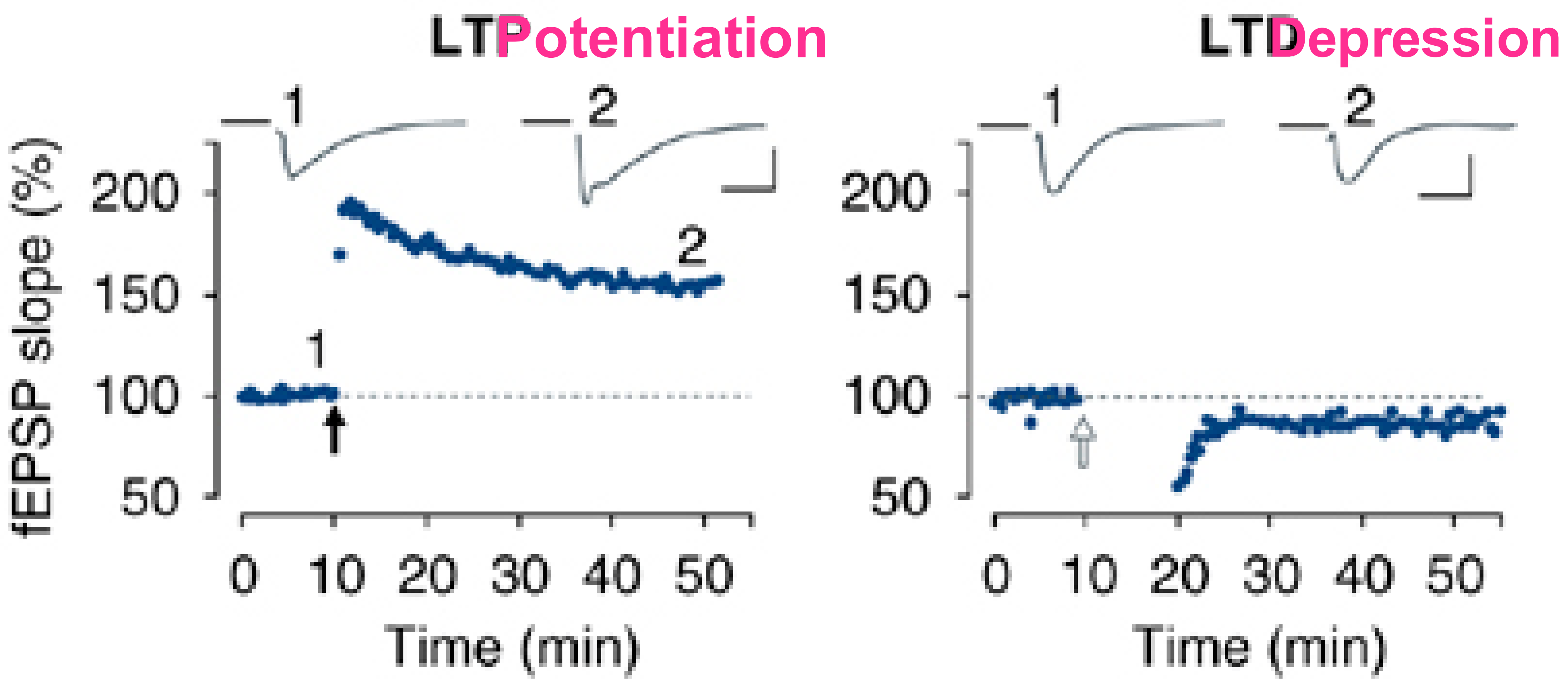


Long-term synaptic plasticity

Long-term synaptic plasticity is a (activity-dependent) semi-permanent change in the strength of the connection from one neuron to another.







Learning and memory via synaptic plasticity

- Synaptic plasticity is generally believed to underlie long-term memory.
- Synapses change their strength according to certain 'rules of plasticity'.
- Synaptic plasticity is linked to learning and memory as follows:
  - Neural activity during learning triggers synaptic strength changes.
  - Synaptic strength changes alters the propensity for neurons to fire.
  - Next time the same neural circuit receives an input, it responds in a different fashion than it otherwise would have.

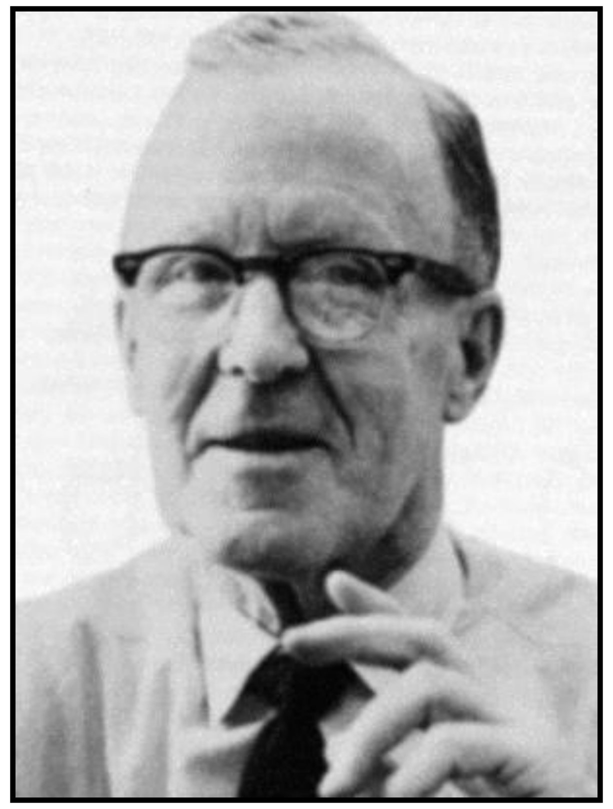
That's memory.

What are these 'rules of synaptic plasticity'?

Some mathematical function of the pre- and post-synaptic neurons' activities… and maybe other stuff?

$$f(pre, post, weight, error, \ldots) = ?$$

Hebbian plasticity

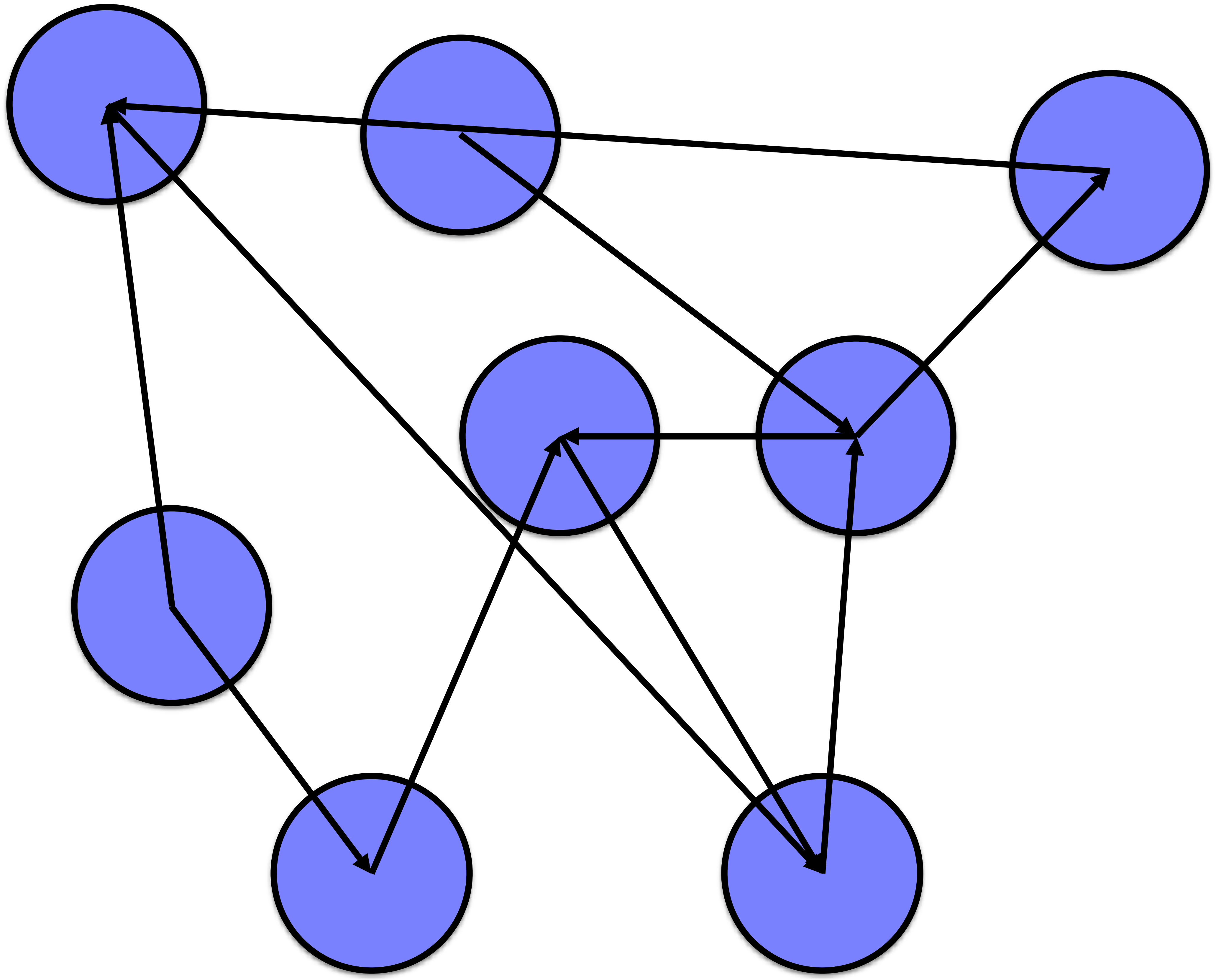


Donald Hebb

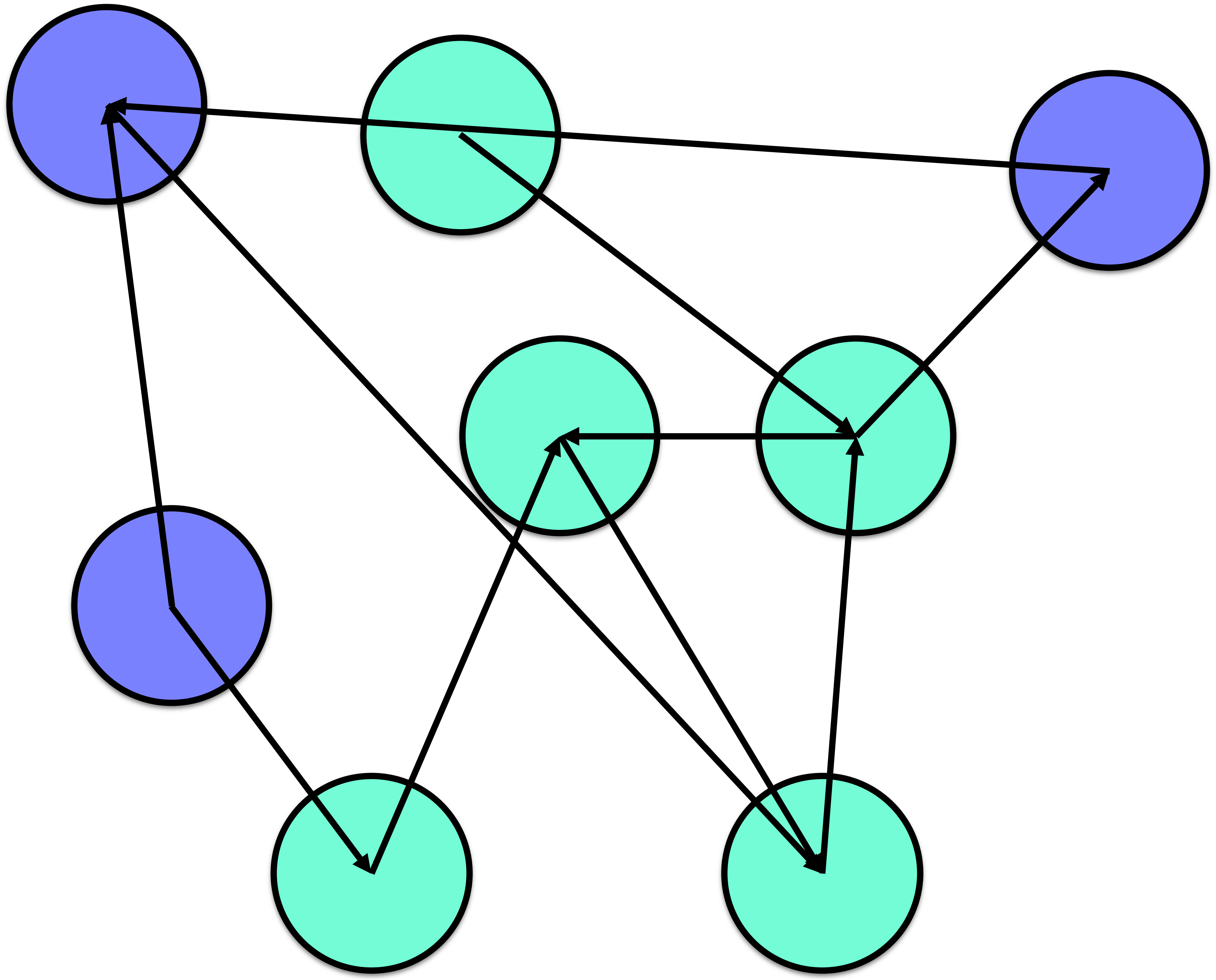
"When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased."
— Donald Hebb (1949)

a.k.a. "neurons that fire together wire together." (Carla Shatz)

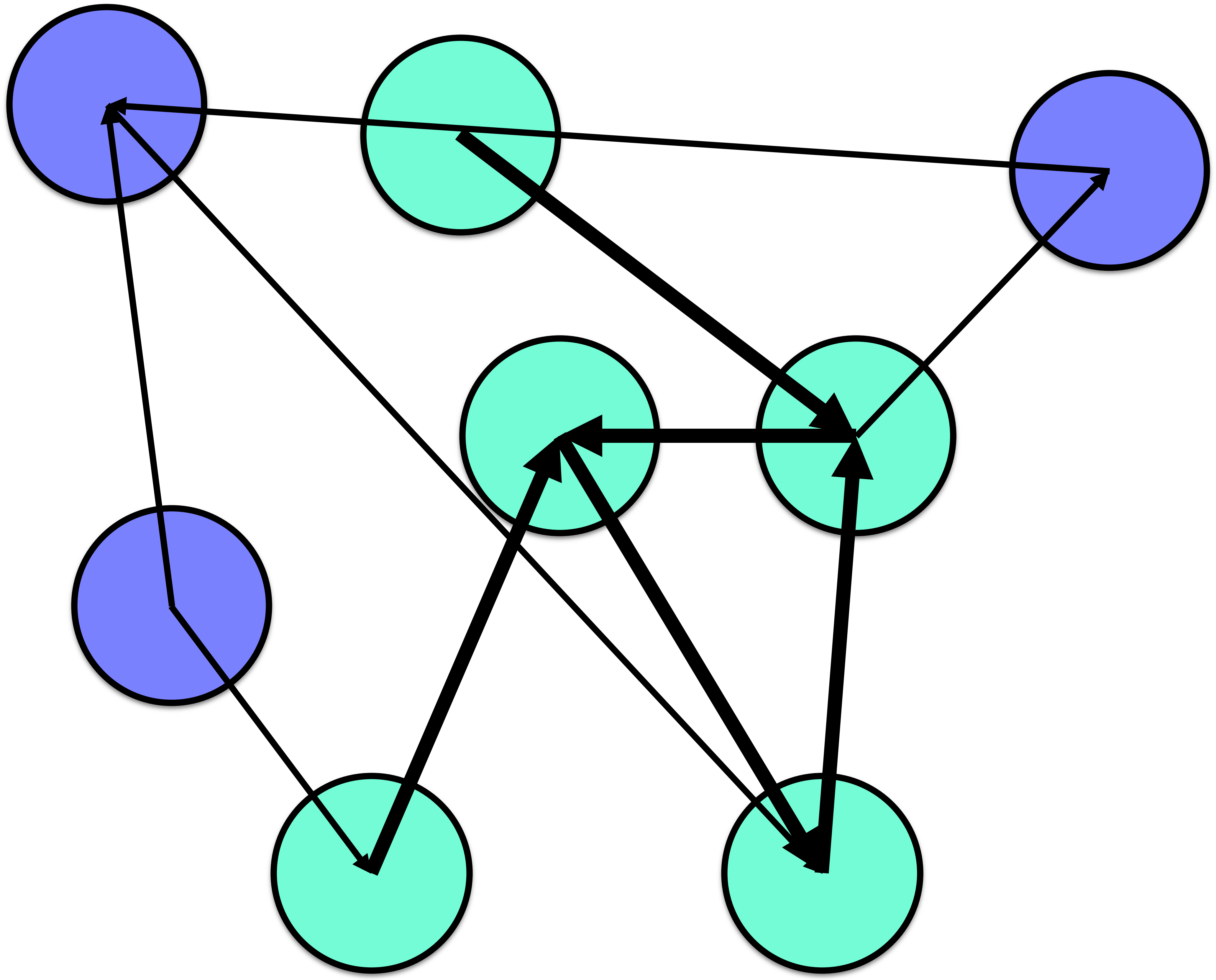
Hebbian plasticity



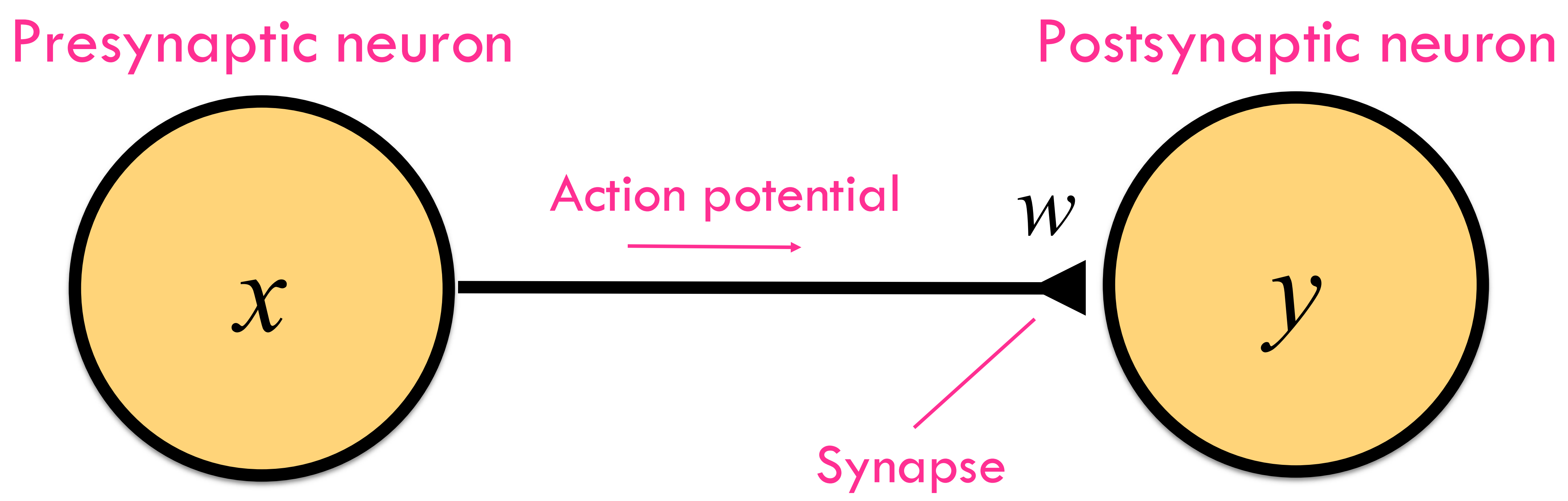
Hebbian plasticity



Hebbian plasticity



Modelling Hebbian plasticity

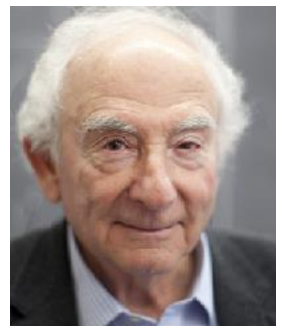


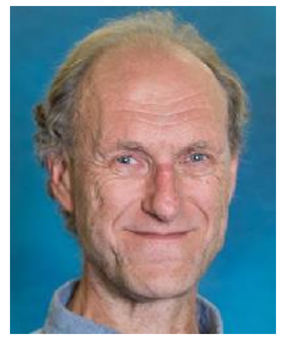
$$\Delta w = f(x, y) \quad = ???$$

A Hebbian rule: $$\Delta w = \eta x y$$

Note that dynamics are unstable: w and therefore y grow without bound.

How to stabilise plasticity?

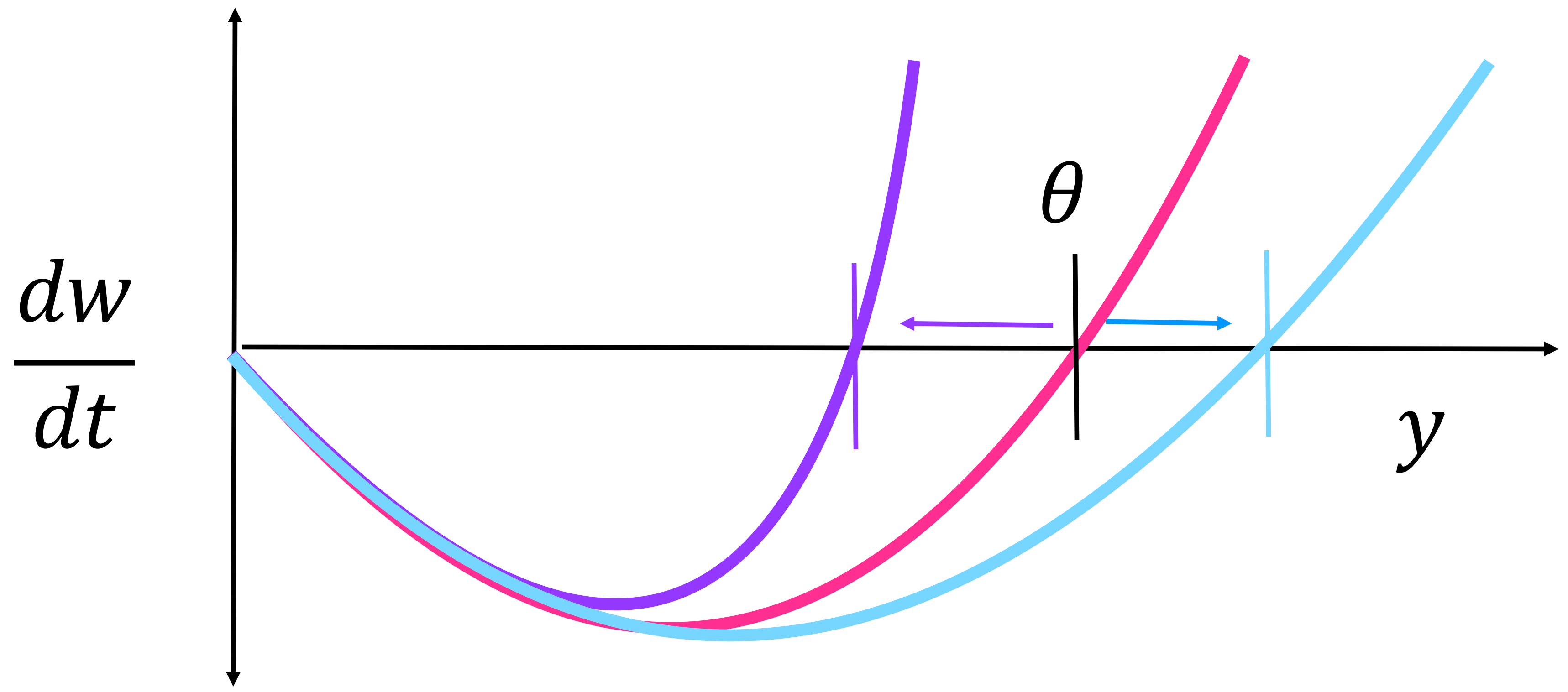




Idea 1: BCM (Bienenstock, Cooper & Munro, 1982)

Modify the basic Hebbian rule by including a postsynaptic threshold for plasticity:

$$\frac{dw}{dt} = \eta_w xy(y - \theta)$$



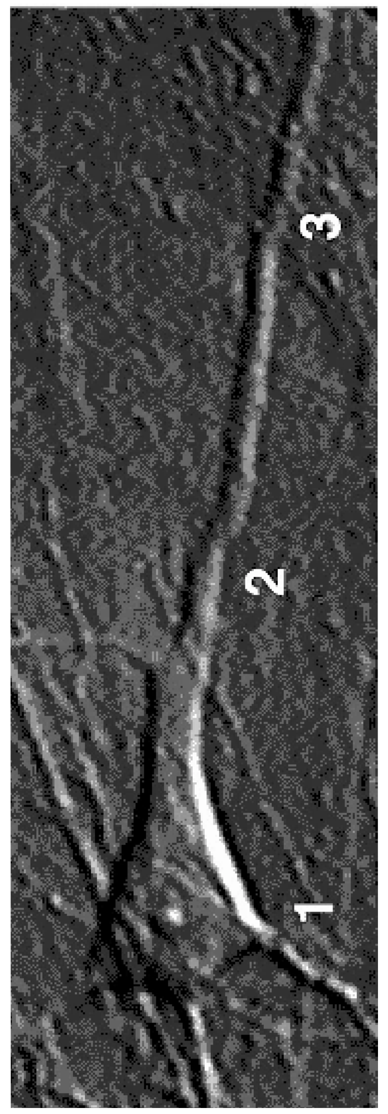
The key idea for stability is that the threshold is also plastic:

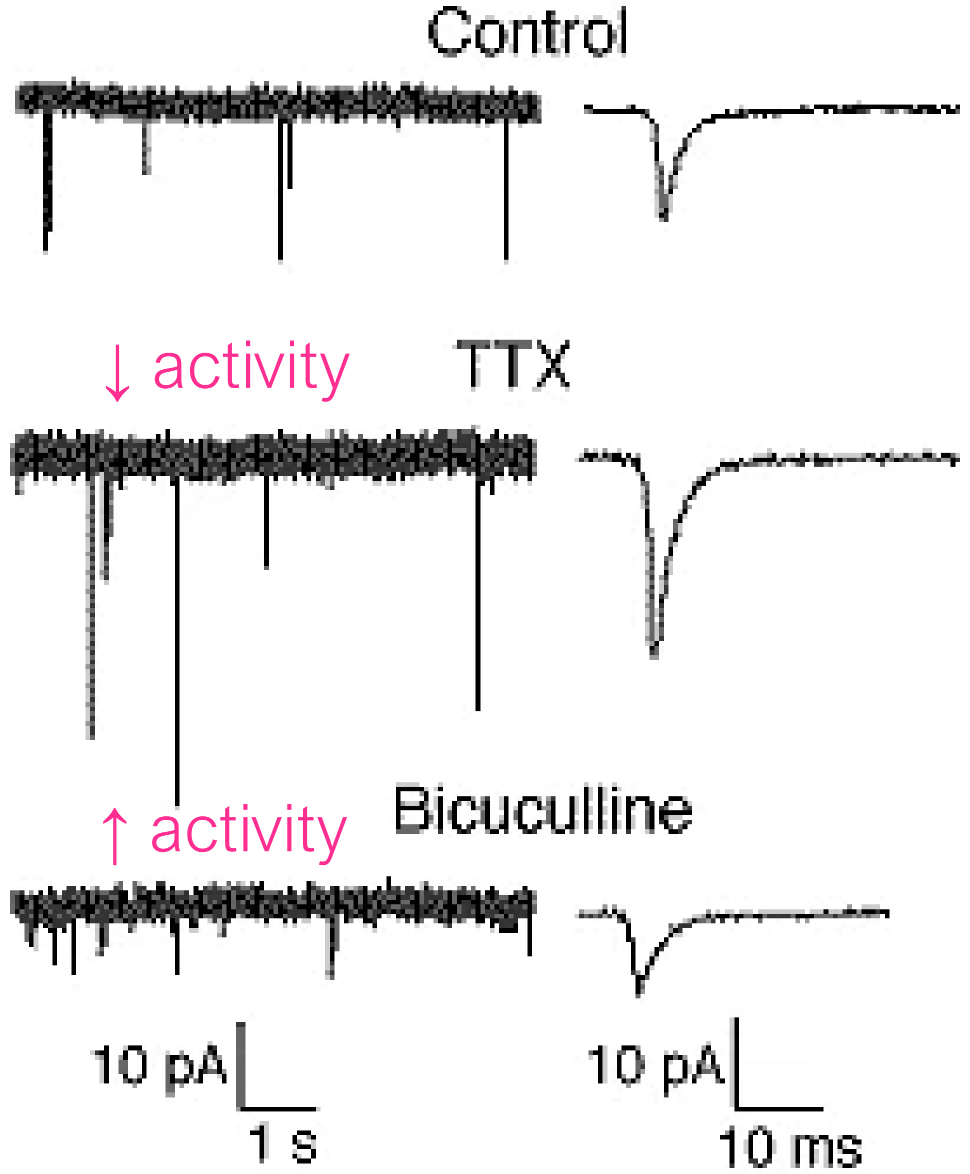
$$\frac{d\theta}{dt} = \eta_\theta (y^2 - \theta_\infty)$$

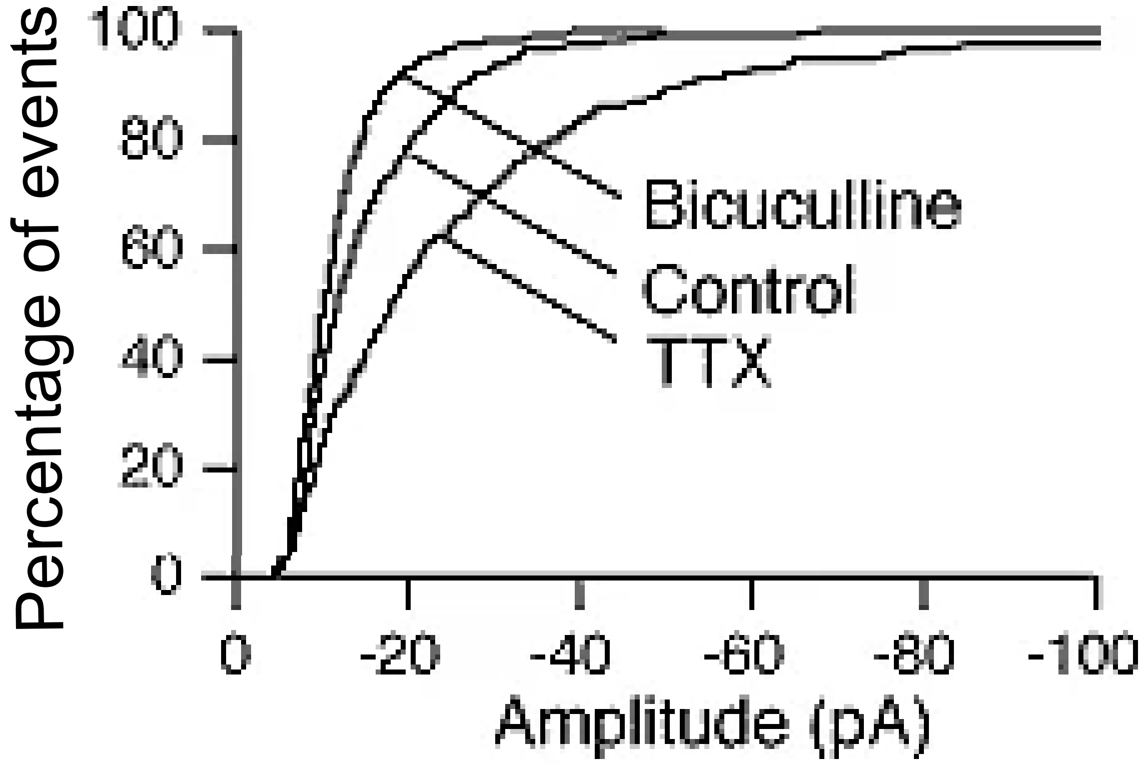
How to stabilise plasticity?

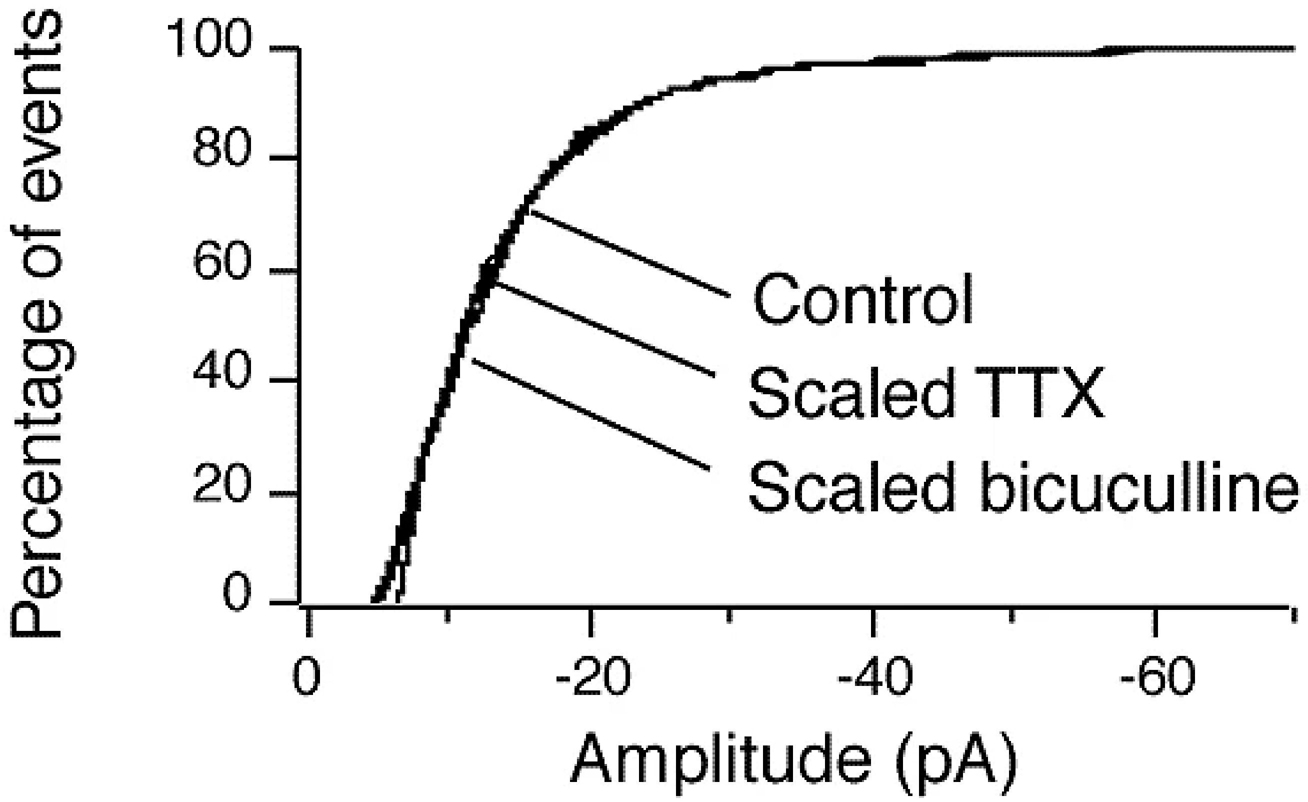
Idea 2: Homeostatic plasticity/synaptic scaling (Turrigiano et al, 1998)

Gina Turrigiano







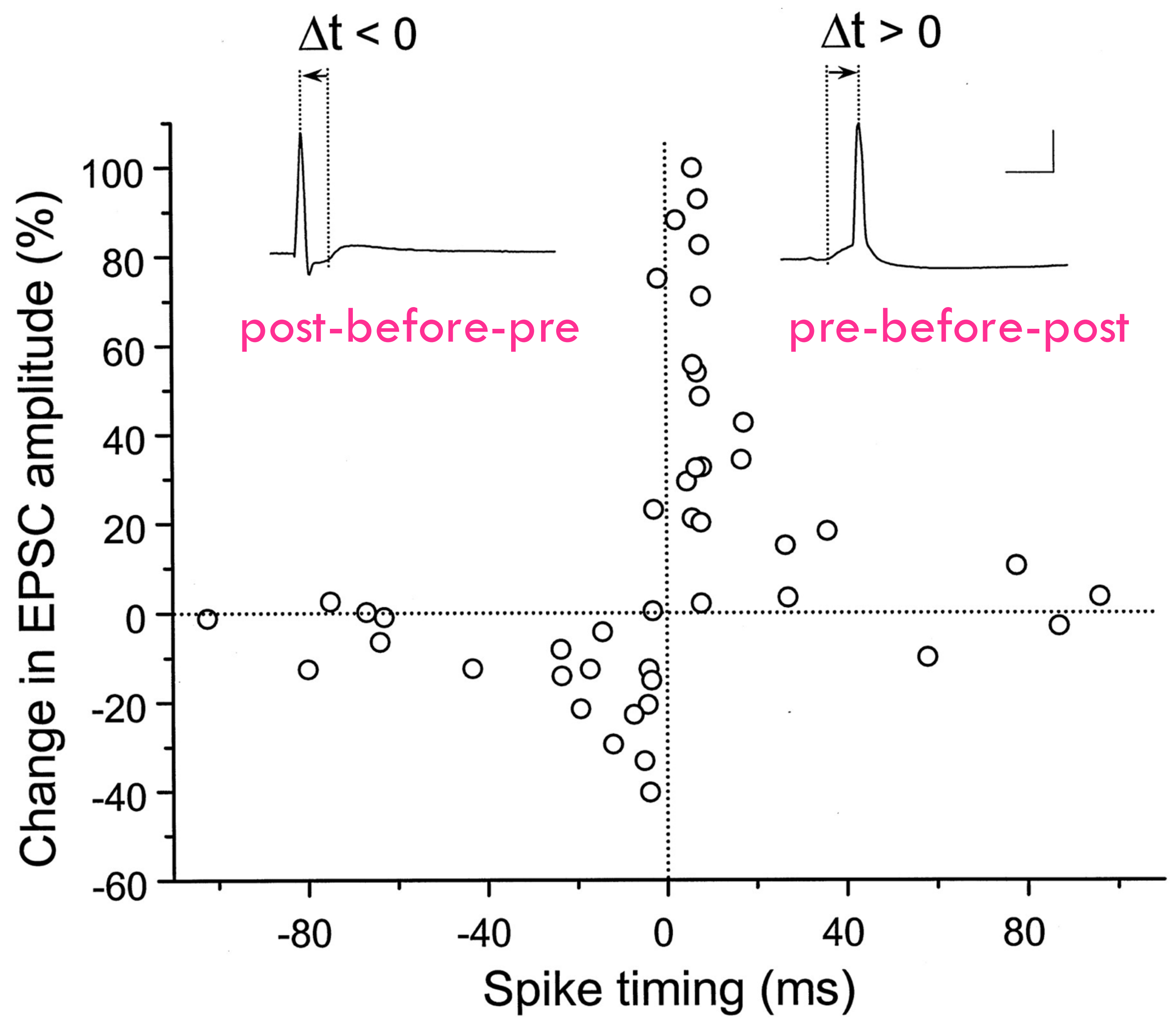


Add a term that makes the weights scale multiplicatively on a slow timescale, to keep postsynaptic activity at some target level:

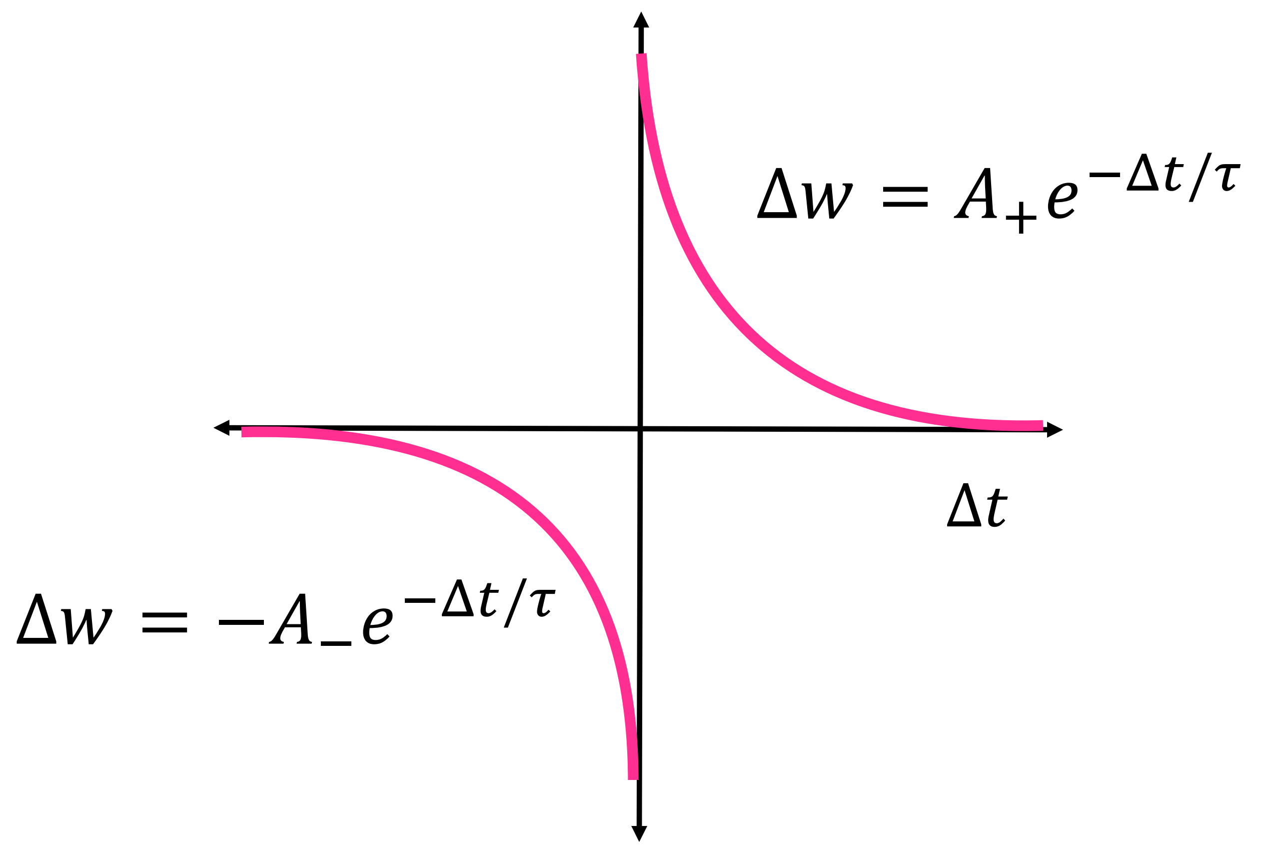
$$\frac{dw}{dt} = \eta_{Hebb} xy + \eta_{ss} w(y_{target} - y)$$

Spike-timing-dependent plasticity

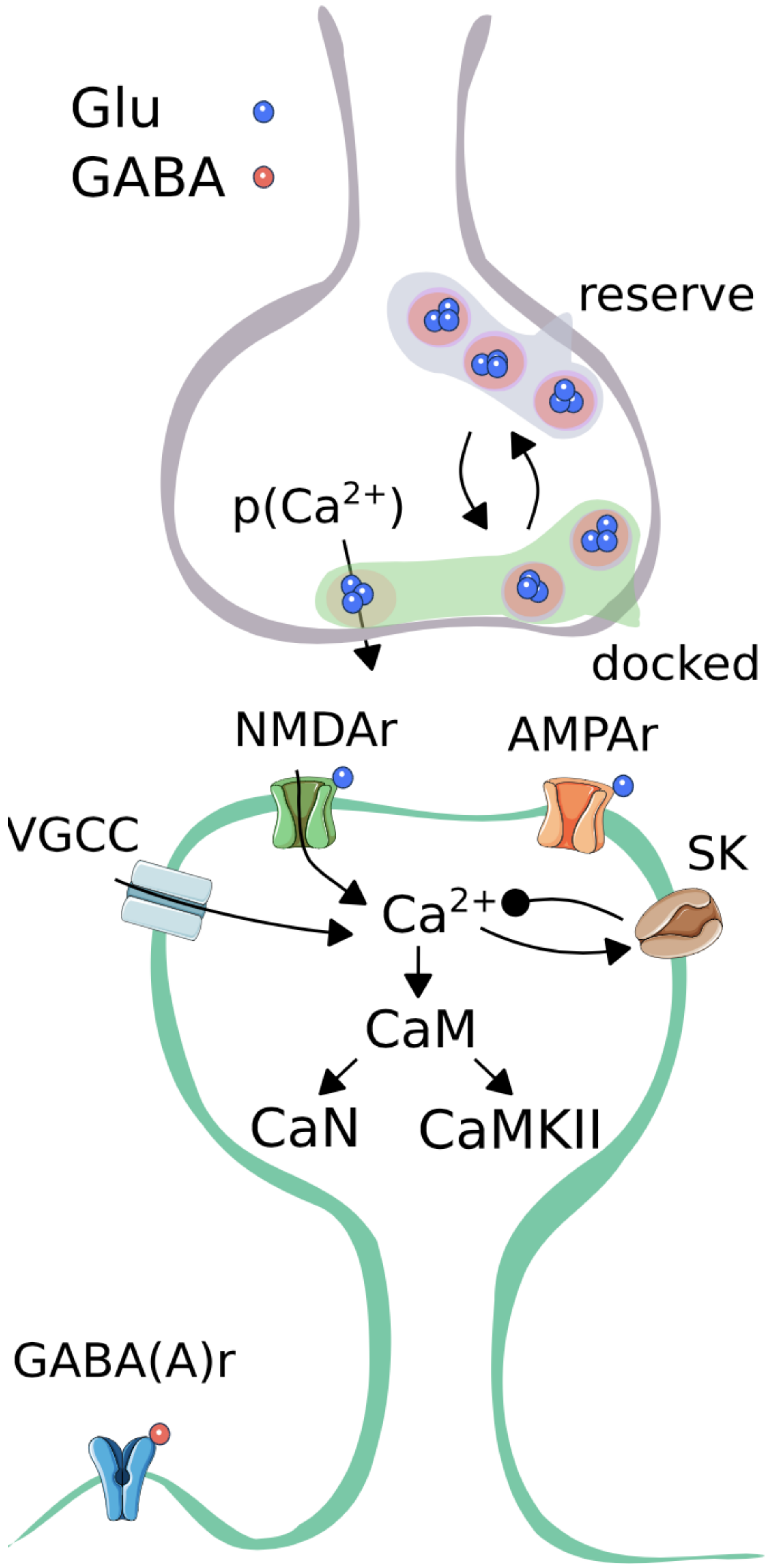
- STDP (discovered in late 1990s) encapsulates the idea of causality implied by Hebb:
  - if presynaptic spike A happened just before postsynaptic spike B, A could have caused B.
  - on the other hand, if presynaptic spike A happened just after postsynaptic spike B,
    A could not have caused B.
- Classic STDP: Pre-before-post causes LTP, post-before-pre causes LTD.
- STDP's existence implies that synapses can detect millisecond-level differences in spike timing when deciding whether to strengthen or weaken.
- When first discovered it was seen as the possible "atom of plasticity".
- "Things turned out to be just as simple as we first thought"
  — No biologist, ever

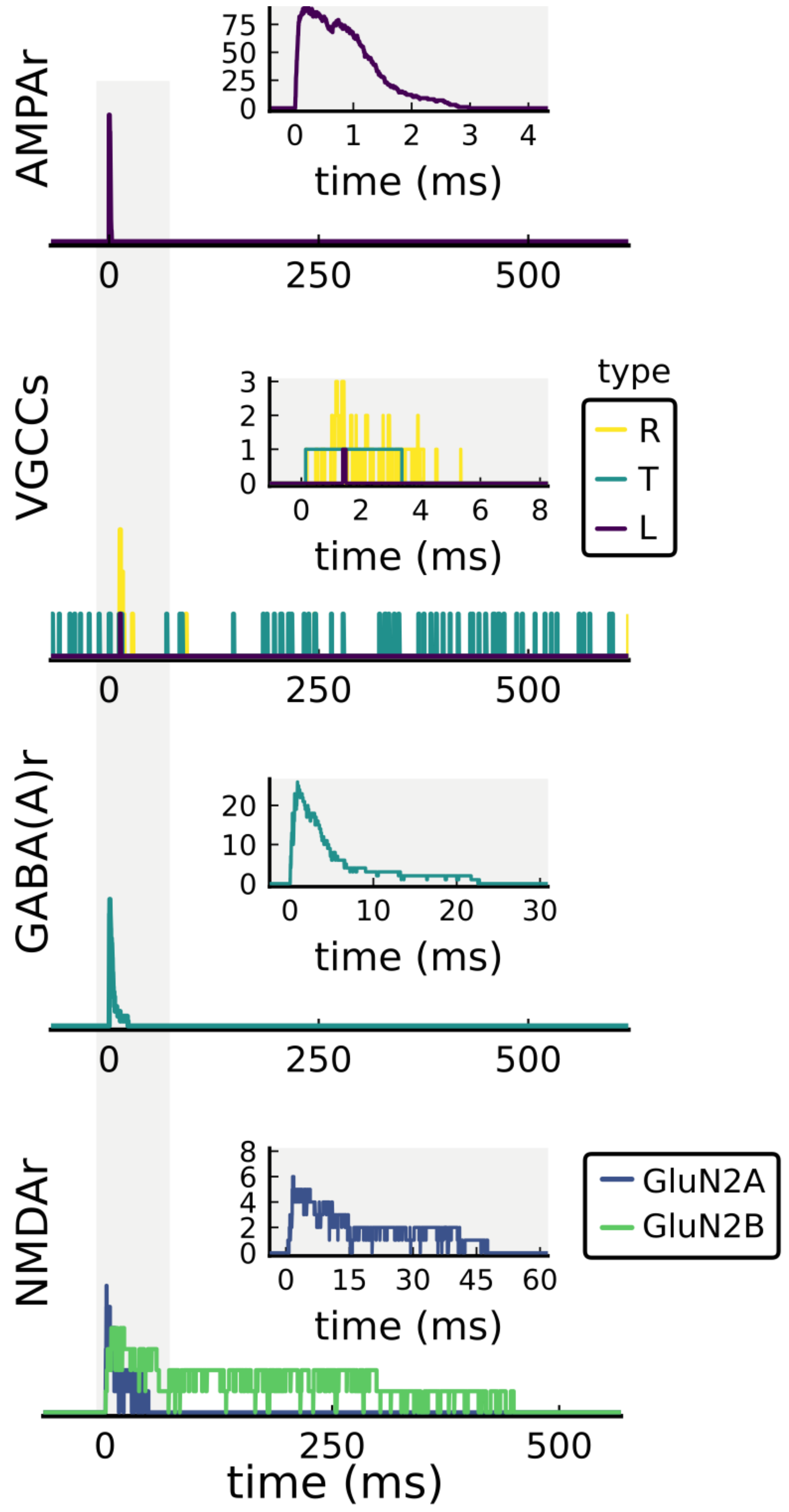


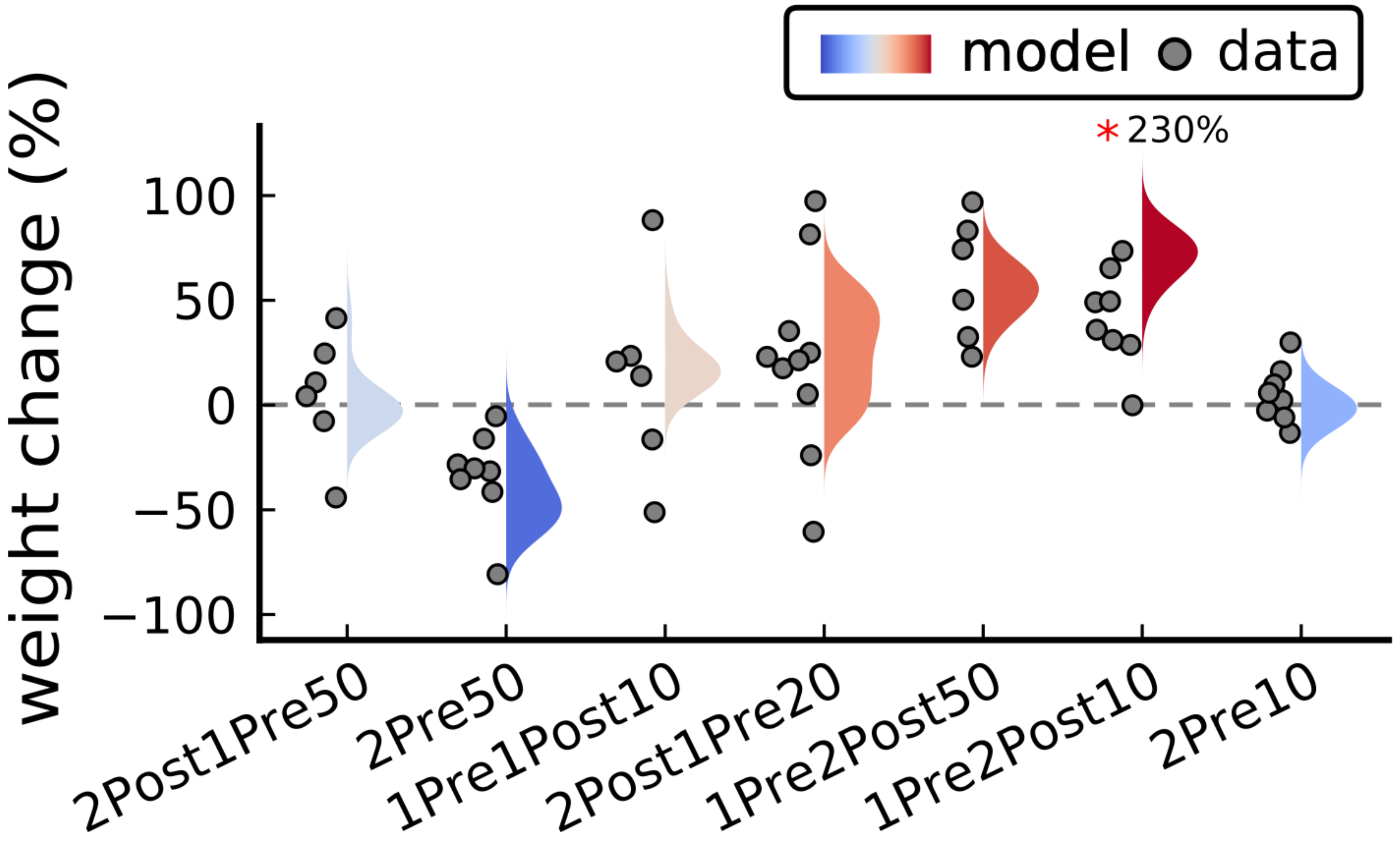
Bi & Poo, J Neurosci (1998)



Adding more biology

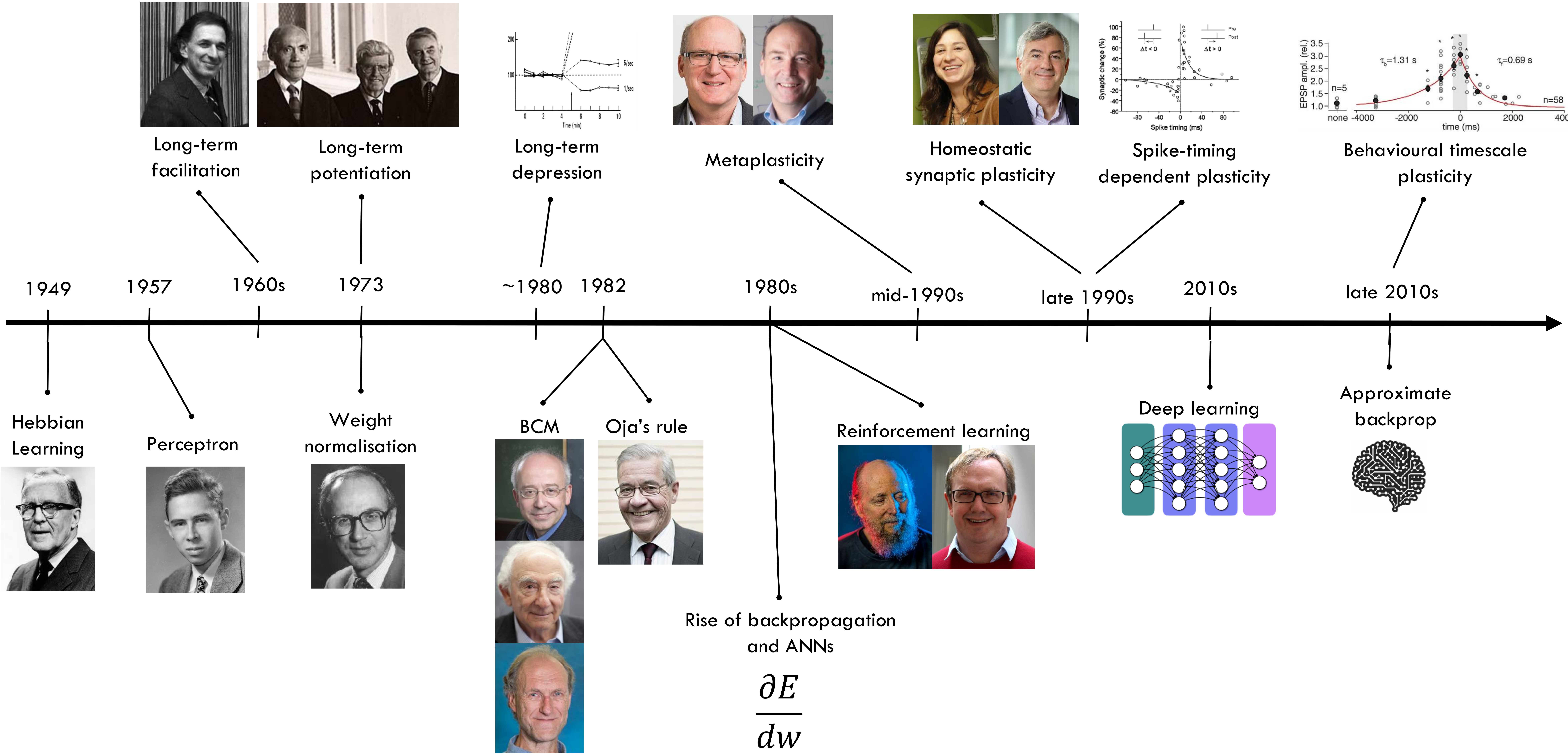






Rodrigues, Tigaret, Marie, O'Donnell, Veltz, eLife (2023)

A brief history of learning rules



Where we are

Learning type	Unsupervised	Supervised	Reinforcement
Example rules	Hebbian learning, STDP, BTSP	Backpropagation	Temporal Difference, Q-learning
Can achieve human-level performance on complex tasks?	Not really ☐	Yes, many and improving 🔥	Yes in specific cases 😎
Known mapping to neurobiology?	Pretty good	Several deep incompatibilities	Less clear, mostly high-level

Further reading on synaptic plasticity

- Simple rate-based plasticity models:
  Dayan and Abbott book (2001), chapter 8.

- STDP:
  Feldman, D.E. (2012). The spike-timing dependence of plasticity. Neuron 75, 556–571.

- Problems with STDP:
  Lisman, J., and Spruston, N. (2005). Postsynaptic depolarization requirements for LTP and LTD: a critique of spike timing-dependent plasticity. Nat Neurosci 8, 839–841.

2. Network models of memory

2a. The Hippocampus

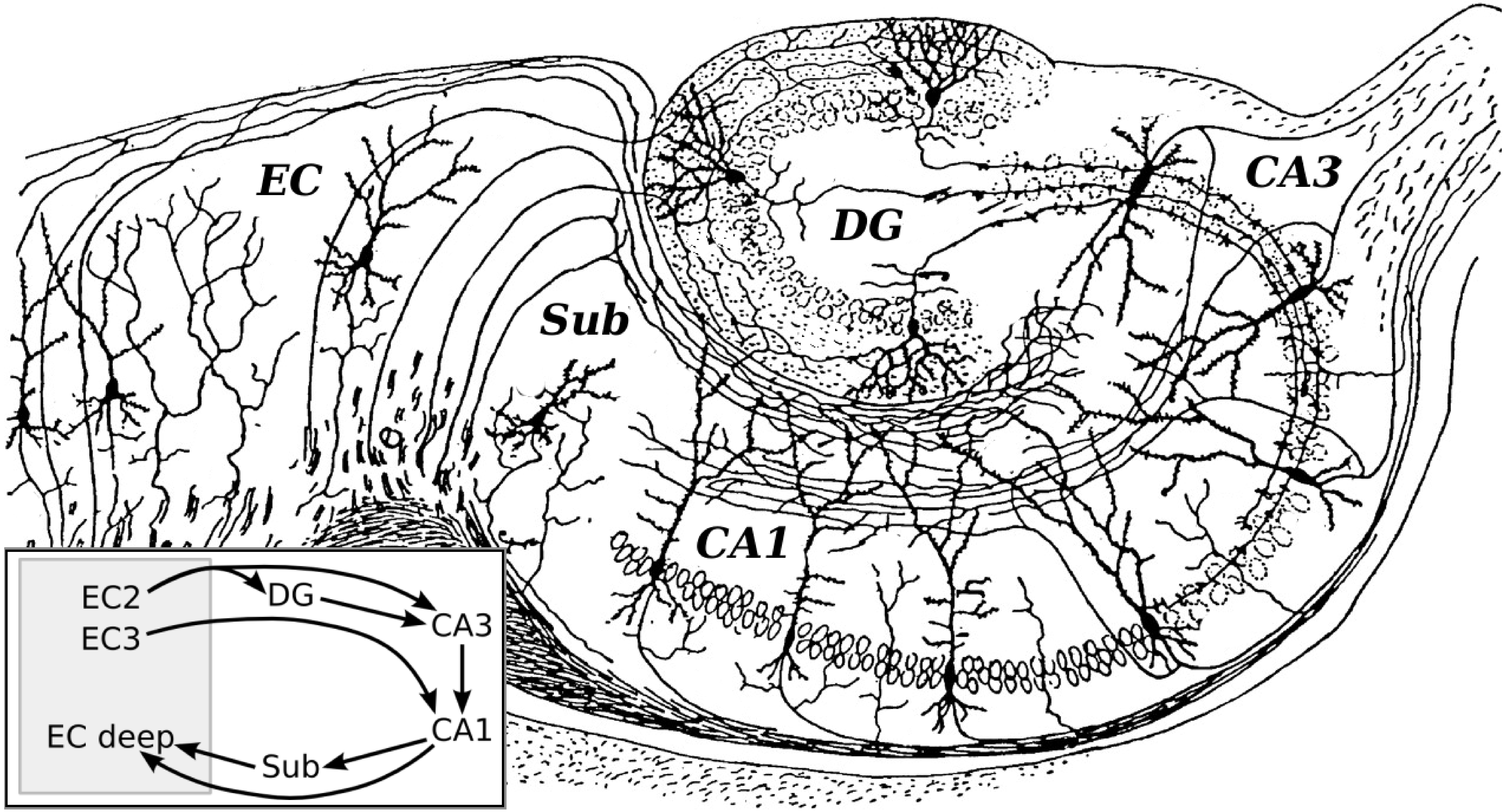
Hippocampus, from the greek words for “horse" and “sea-monster”

https://en.wikipedia.org/wiki/Hippocampus#/media/File:Hippocampus_and_seahorse_cropped.JPG

Anatomy of the hippocampus

https://mouse.brain-map.org

Anatomy of the hippocampus

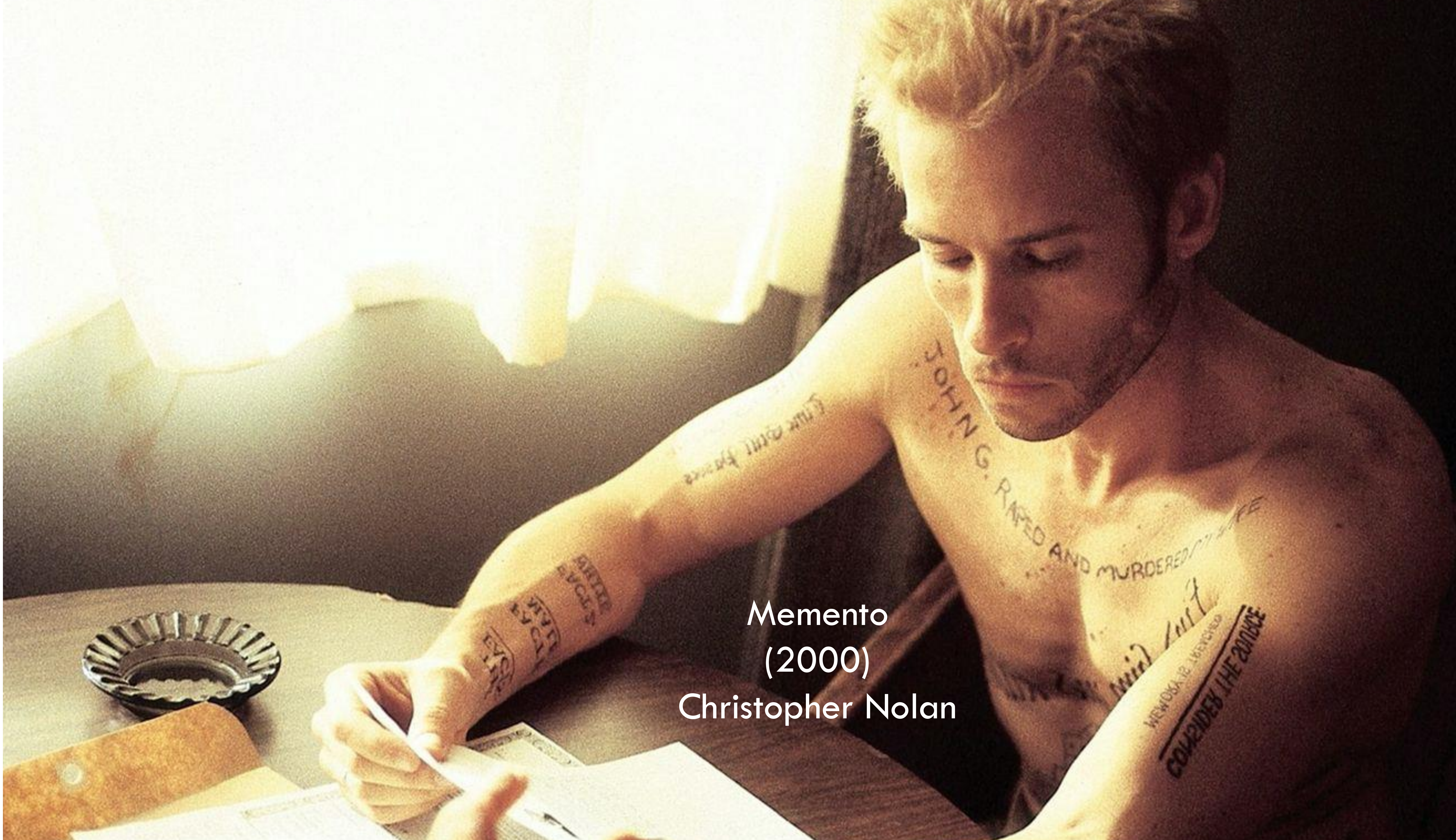


Original drawing by Ramon y Cajal (circa 1900)

https://en.wikipedia.org/wiki/Hippocampus#/media/File:CajalHippocampus_(modified).png

Hippocampus and memory

- Patient HM (who had his hippocampus surgically removed) could not form new long-term memories, and also had time-limited retrograde amnesia.
- The hippocampus is specifically needed for encoding new episodic memories, but is not necessary for other memories (e.g. procedural).
- Memory encoding requires synaptic plasticity in the hippocampus.



Memento
(2000)
Christopher Nolan

Hippocampus is needed for forming episodic memories

Long-term memory types

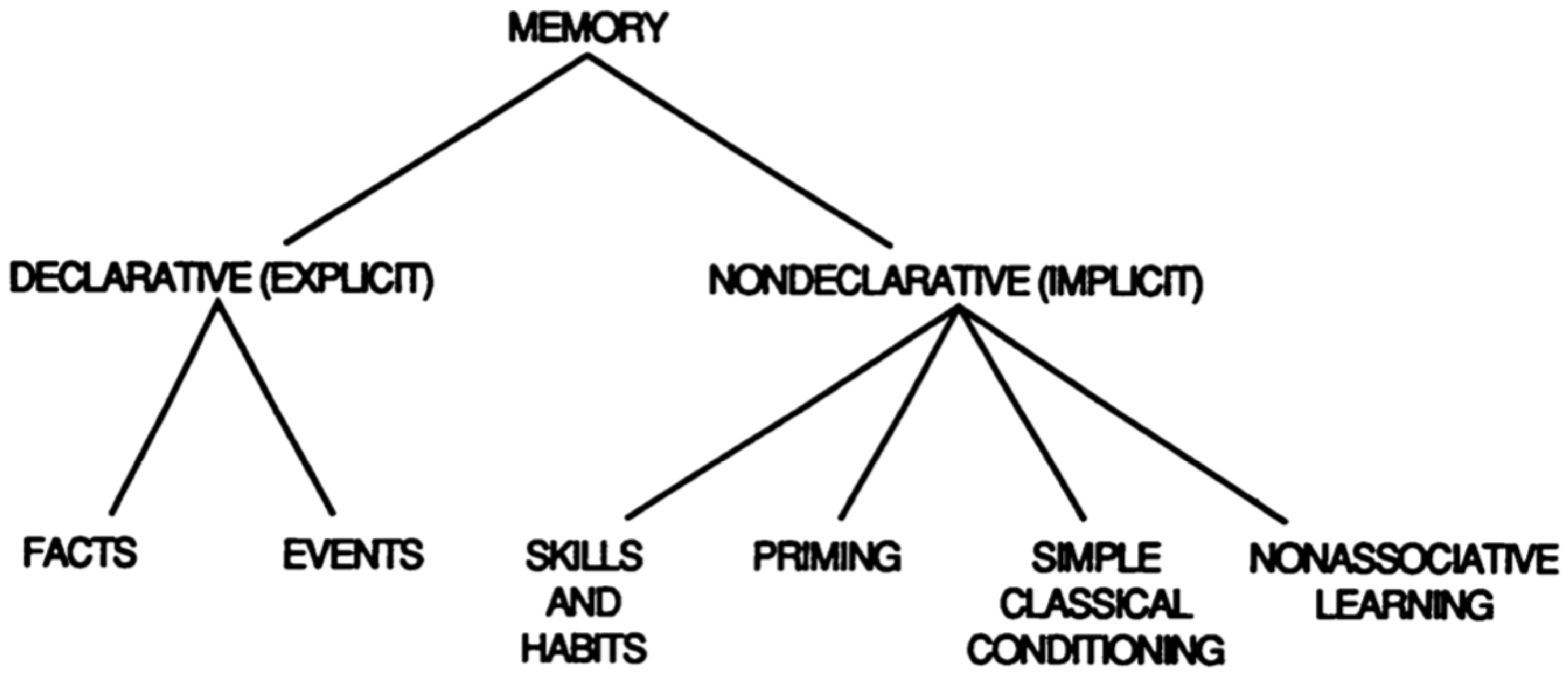
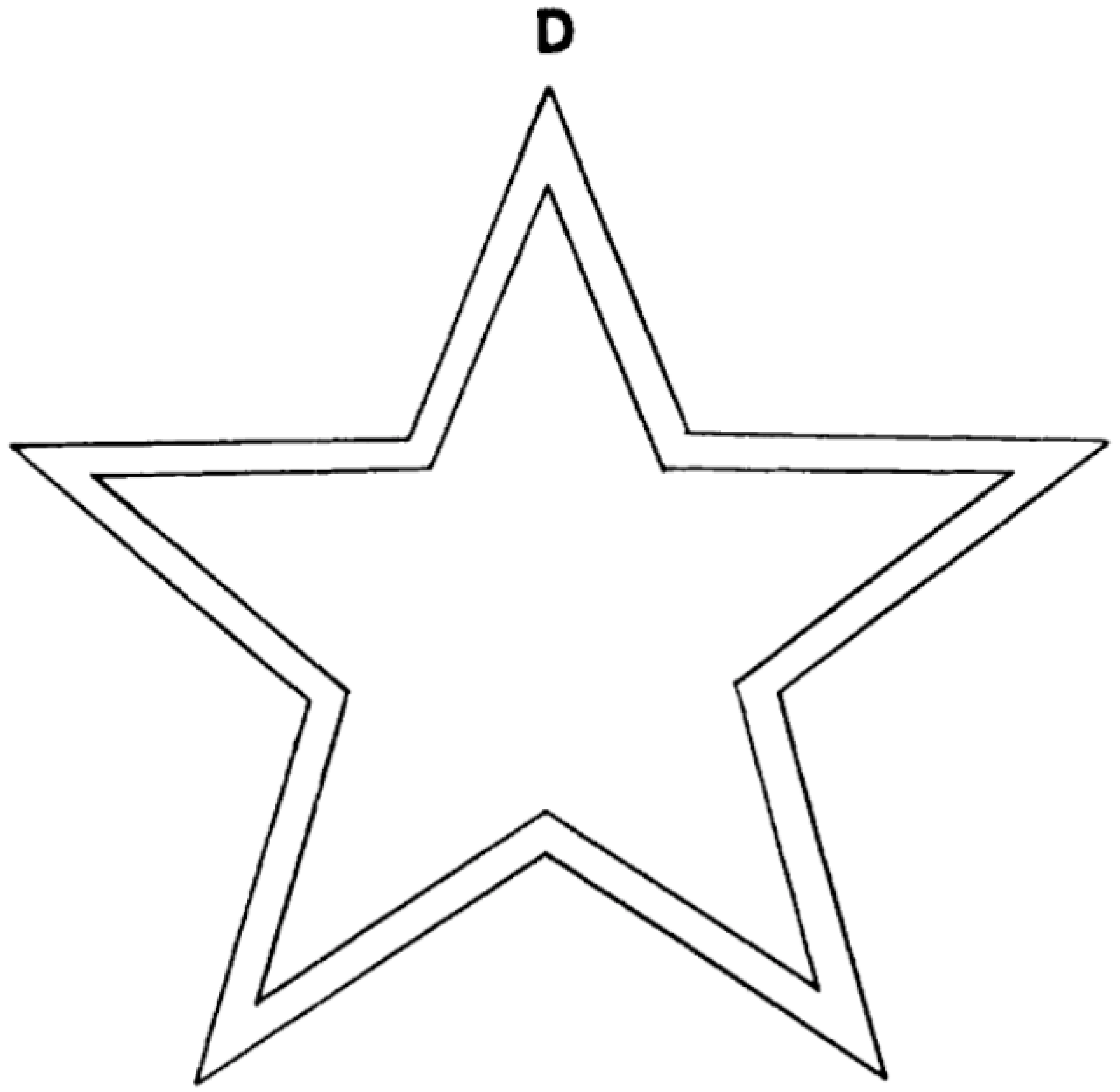


Fig. 3. Classification of memory. Declarative (explicit) memory refers to conscious recollections of facts and events and depends on the integrity of the medial temporal lobe (see text). Nondeclarative (implicit) memory refers to a collection of abilities and is independent of the medial temporal lobe (*60*). Nonassociative learning includes habituation and sensitization. In the case of nondeclarative memory, experience alters behavior nonconsciously without providing access to any memory content (*19*, *20*).

Squire & Zola-Morgan, Science 1991

H.M. could form new motor memories

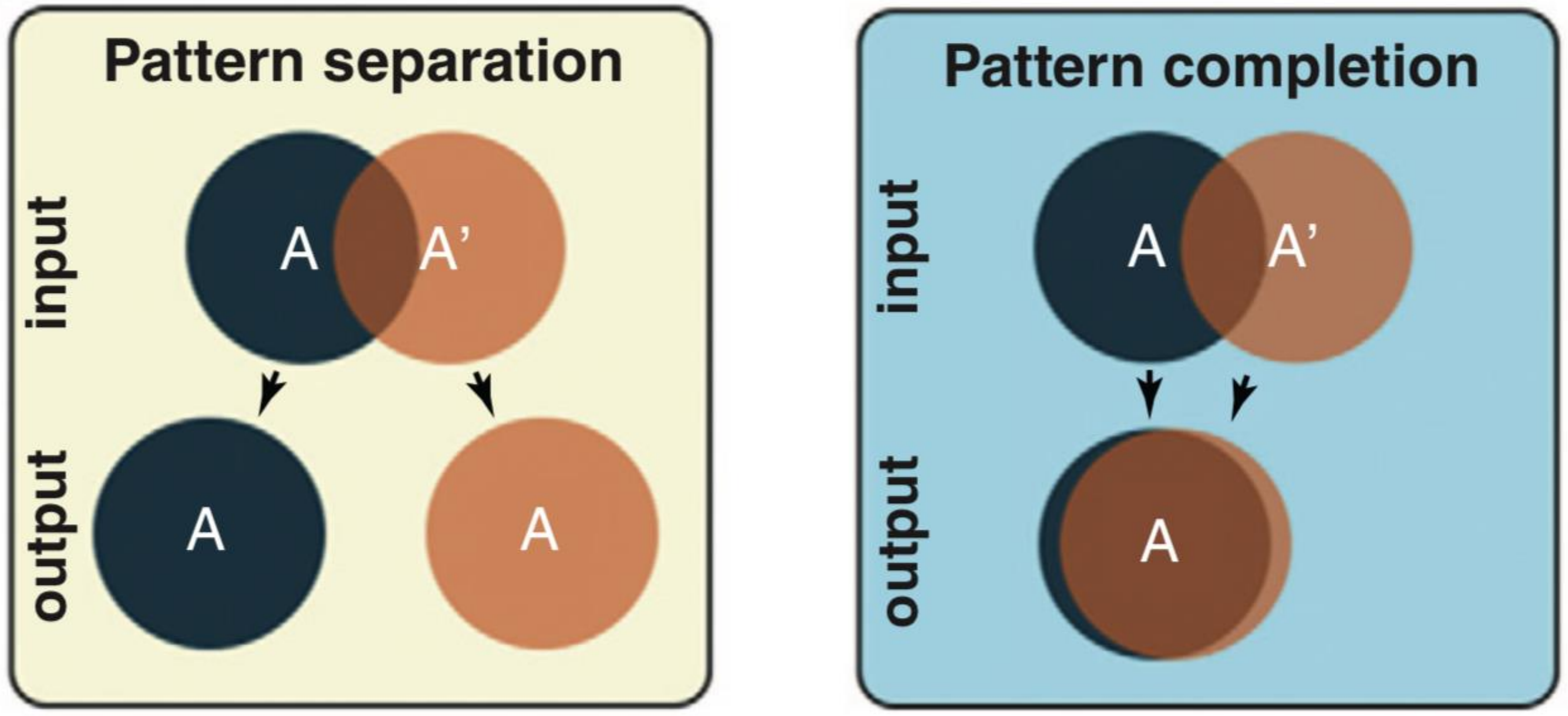


Milner (1962)

2b. Feedforward network models of hippocampus

Pattern separation vs pattern completion

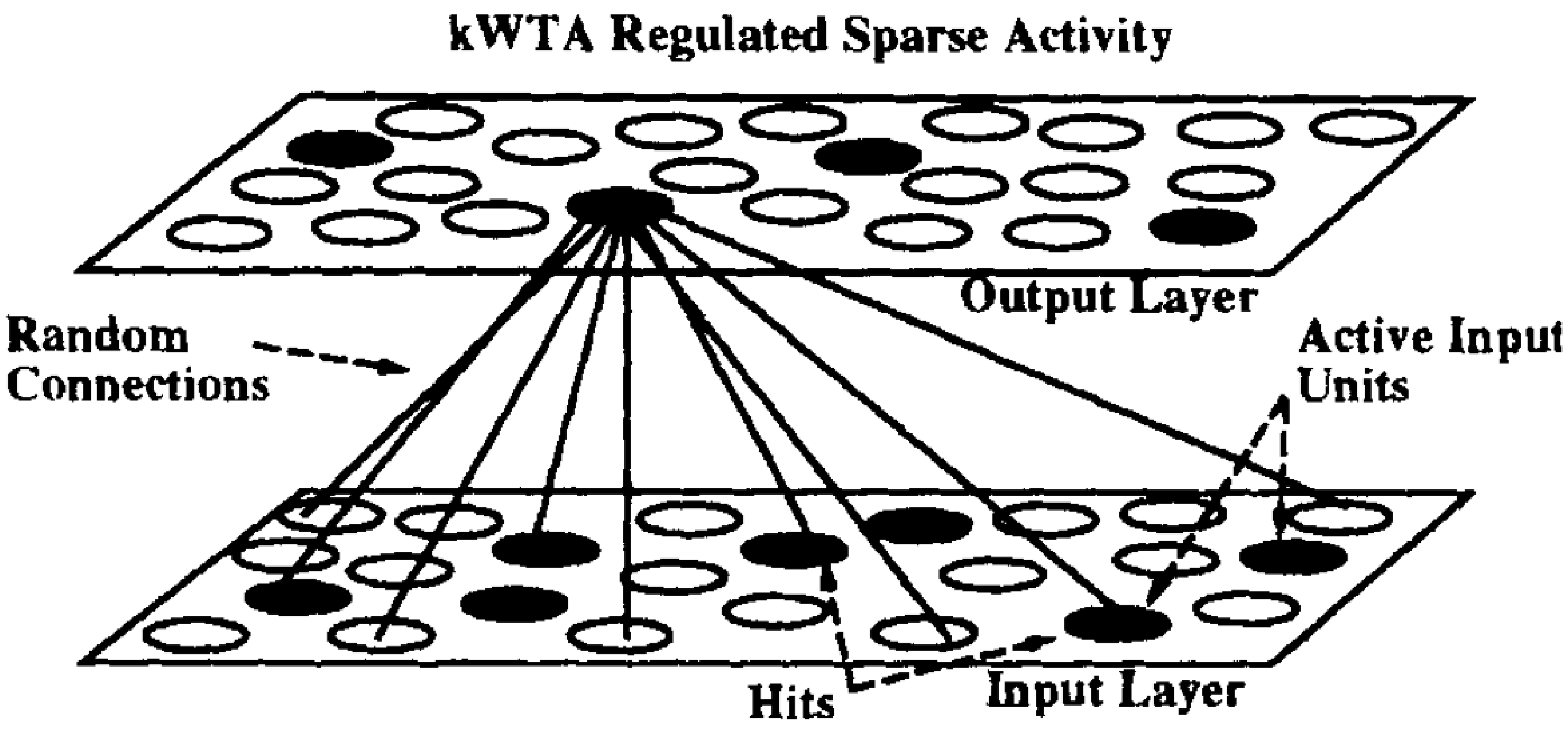
During memory encoding and retrieval, the hippocampus performs pattern separation and pattern completion.

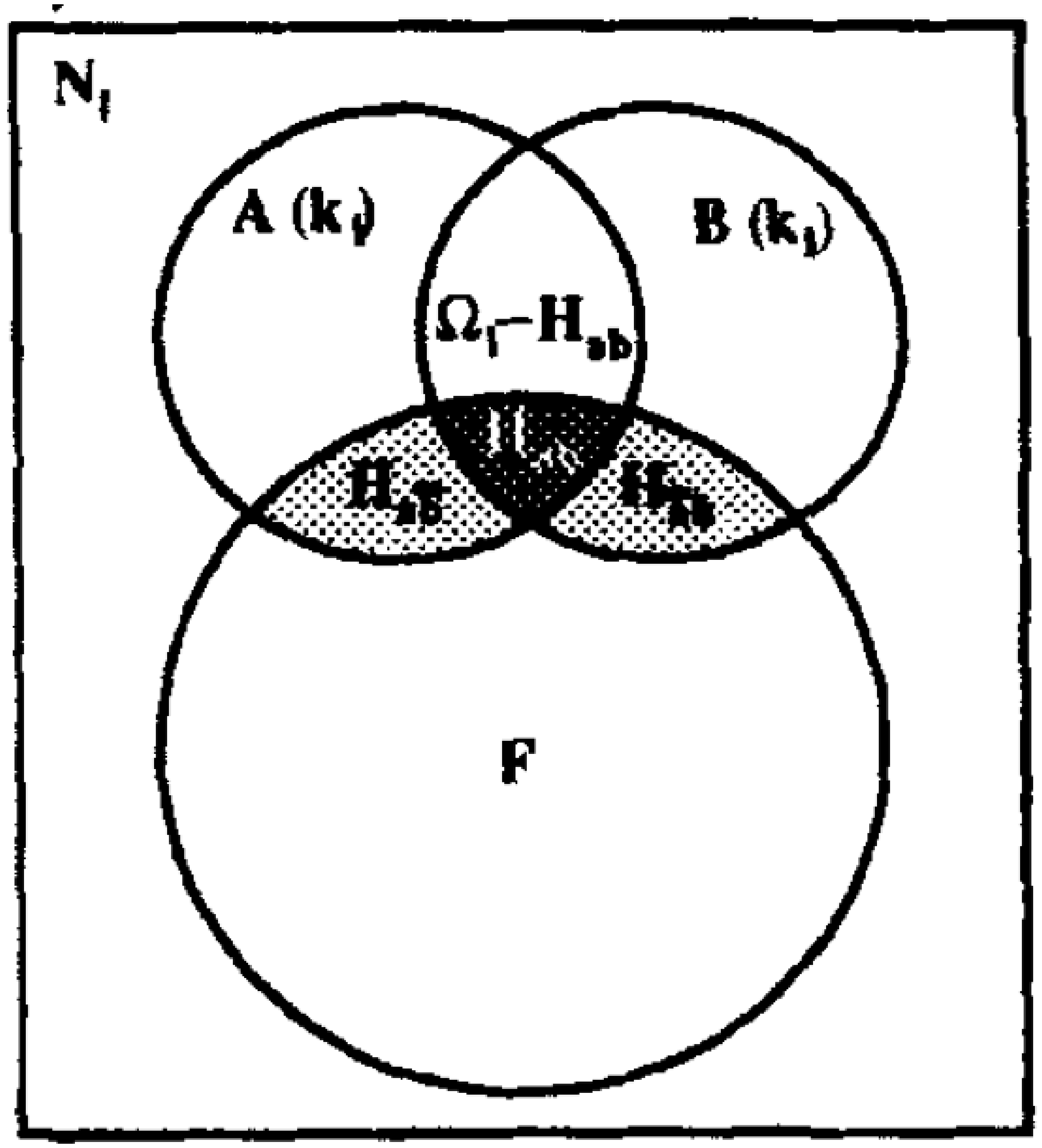


Yassa & Stark, Trends Neurosci (2011)

- Dentate gyrus is thought to do pattern separation.
- CA3 is thought to do pattern completion.

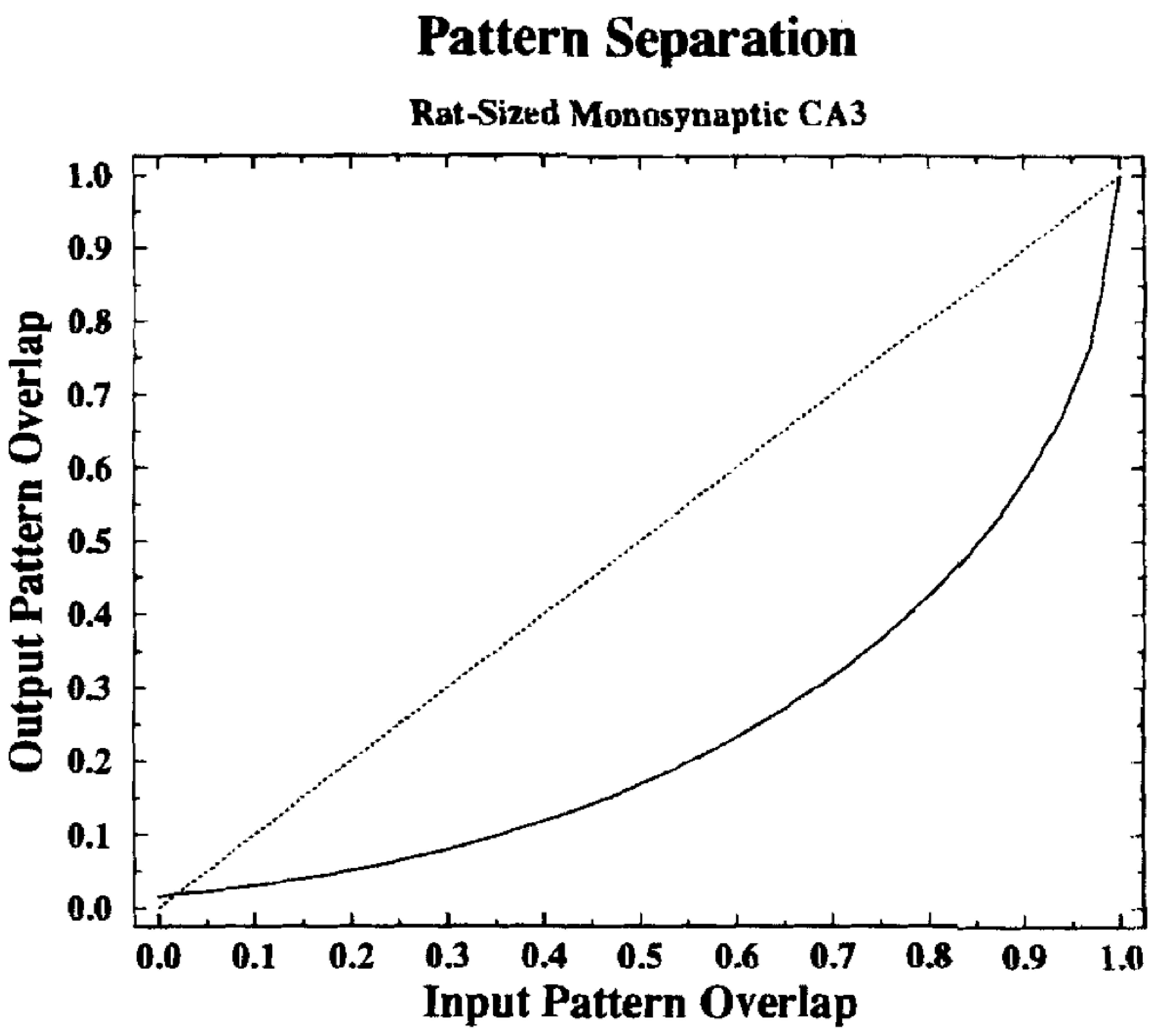
Feedforward model of hippocampus

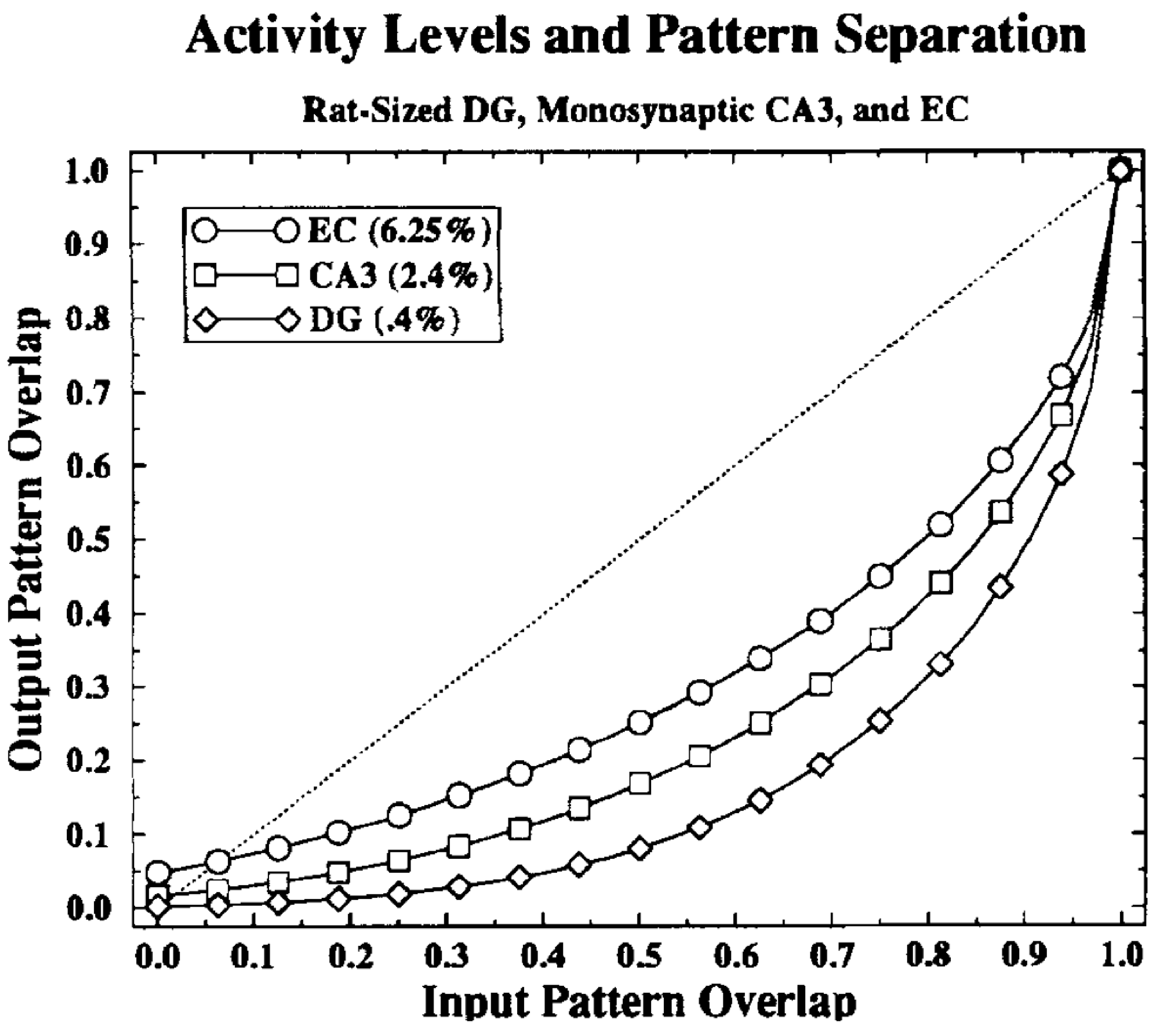




O'Reilly & McClelland, Hippocampus (1994)

Feedforward model of hippocampus



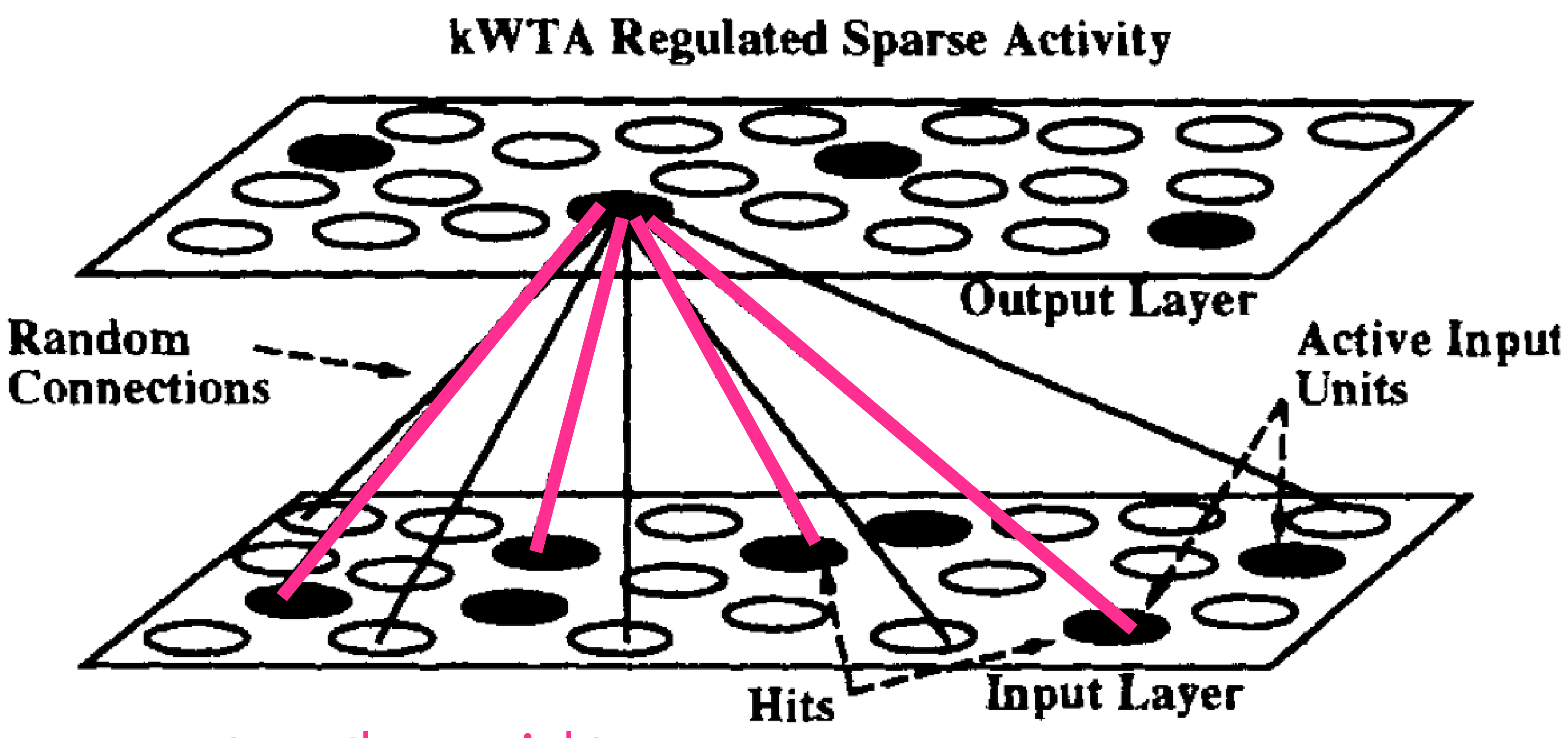


→ sparsity increases pattern separation!

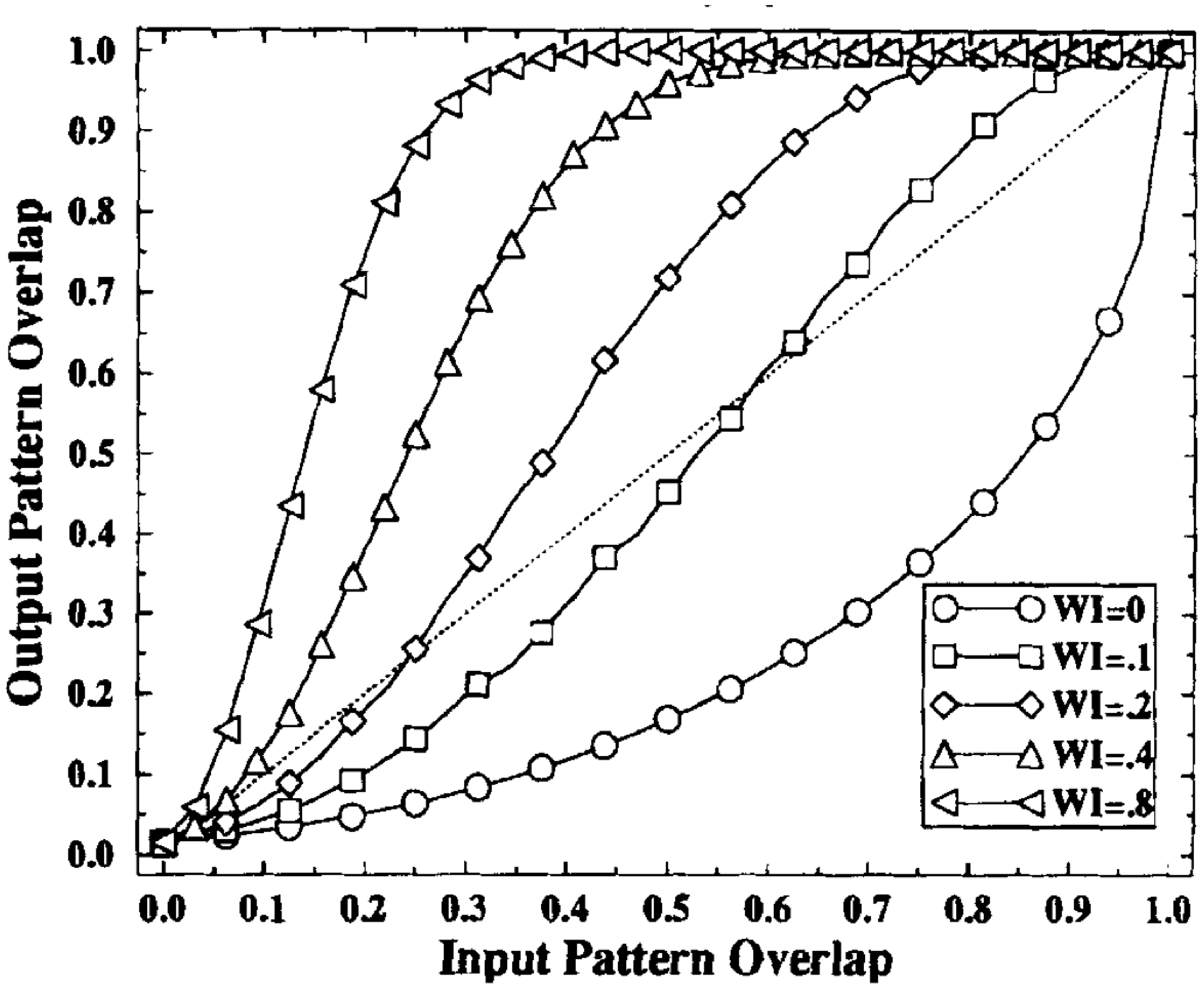
O'Reilly & McClelland, Hippocampus (1994)

Feedforward model of hippocampus

Hebbian strengthening of synaptic weights induces pattern completion.



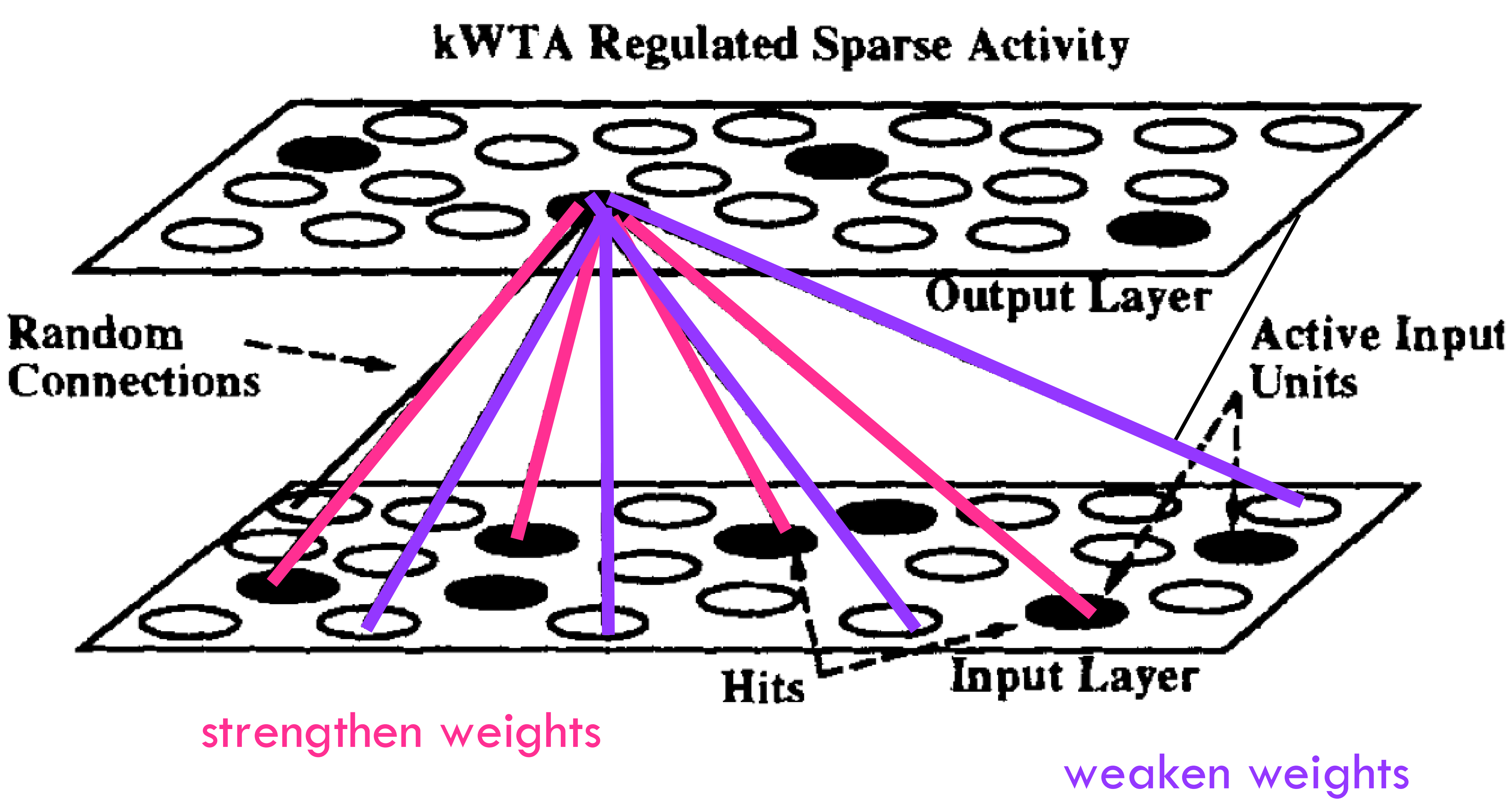
strengthen weights

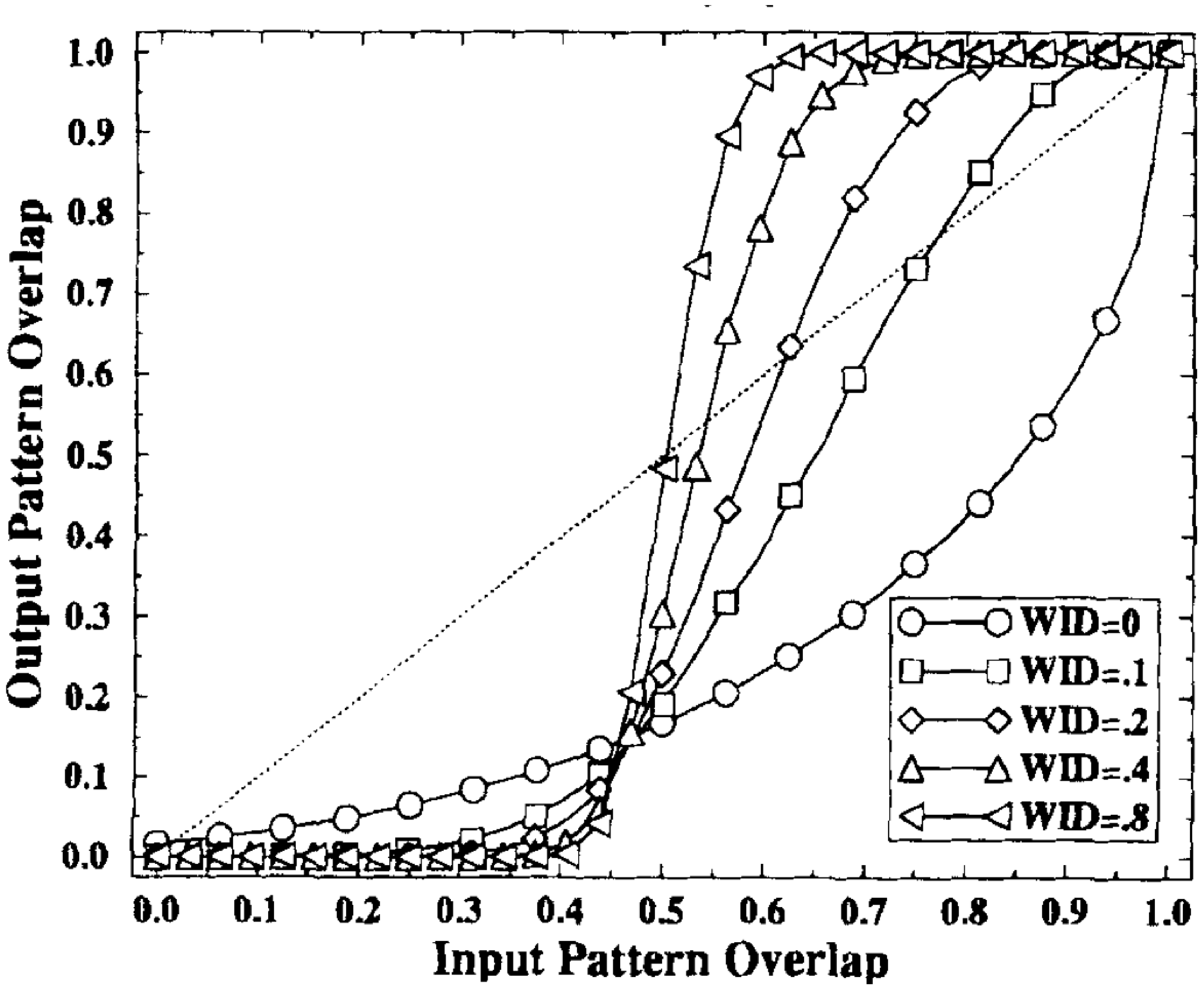


O'Reilly & McClelland, Hippocampus (1994)

Feedforward model of hippocampus

Heterosynaptic weakening of ‘off target’ synapses preserves separation for low-overlap patterns.

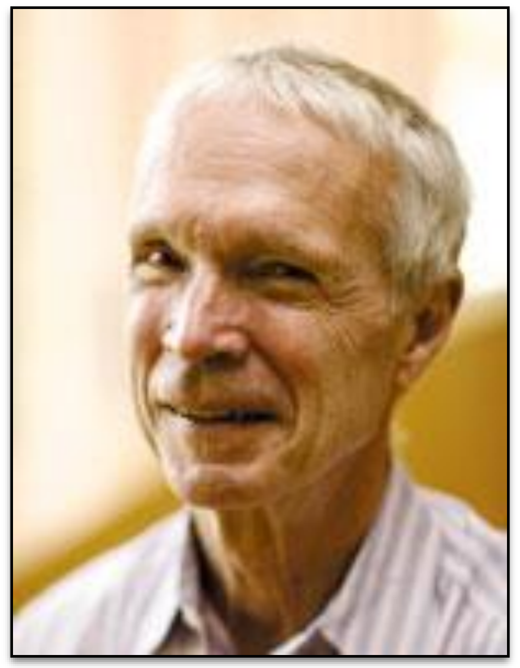




O’Reilly & McClelland, Hippocampus (1994)

2c. Hopfield networks

Hopfield networks



John Hopfield

- A basic model of associative memory recall, as in the CA3 region of the hippocampus.
- A Hopfield network is a recurrent network of "neurons" with synaptic connections set so to store memories as attractors (proposed in 1982).
- The network state evolves dynamically, typically toward some "attractor" state.
- A simple Hebb-like synaptic plasticity rule can imprint attractors in the network weights.
- Incredibly influential model in the history of computational neuroscience (attracted a bunch of physicists to the field).

Hopfield networks

- Network state evolves as:

$$if \sum_{j \neq i}^{N} [w_{ij} x_j(t) - \theta] > 0 \qquad then x_i \rightarrow +1$$

$$otherwise x_i \rightarrow -1$$

- There are two common flavours: synchronous or asynchronous updates.
- The weights follow a Hebbian-like rule:

$$w_{ij} = \frac{1}{N_{patterns}} \sum_{a}^{N_{patterns}} x_i^a x_j^a$$

- Usually the weights are symmetric: $w_{ij} = w_{ji}$
  and the connectivity is all-to-all.
- The network dynamics evolve to minimise an "energy":

$$E = -\frac{1}{2} \sum_{ij} w_{ij} x_i x_j$$

Hopfield network dynamics

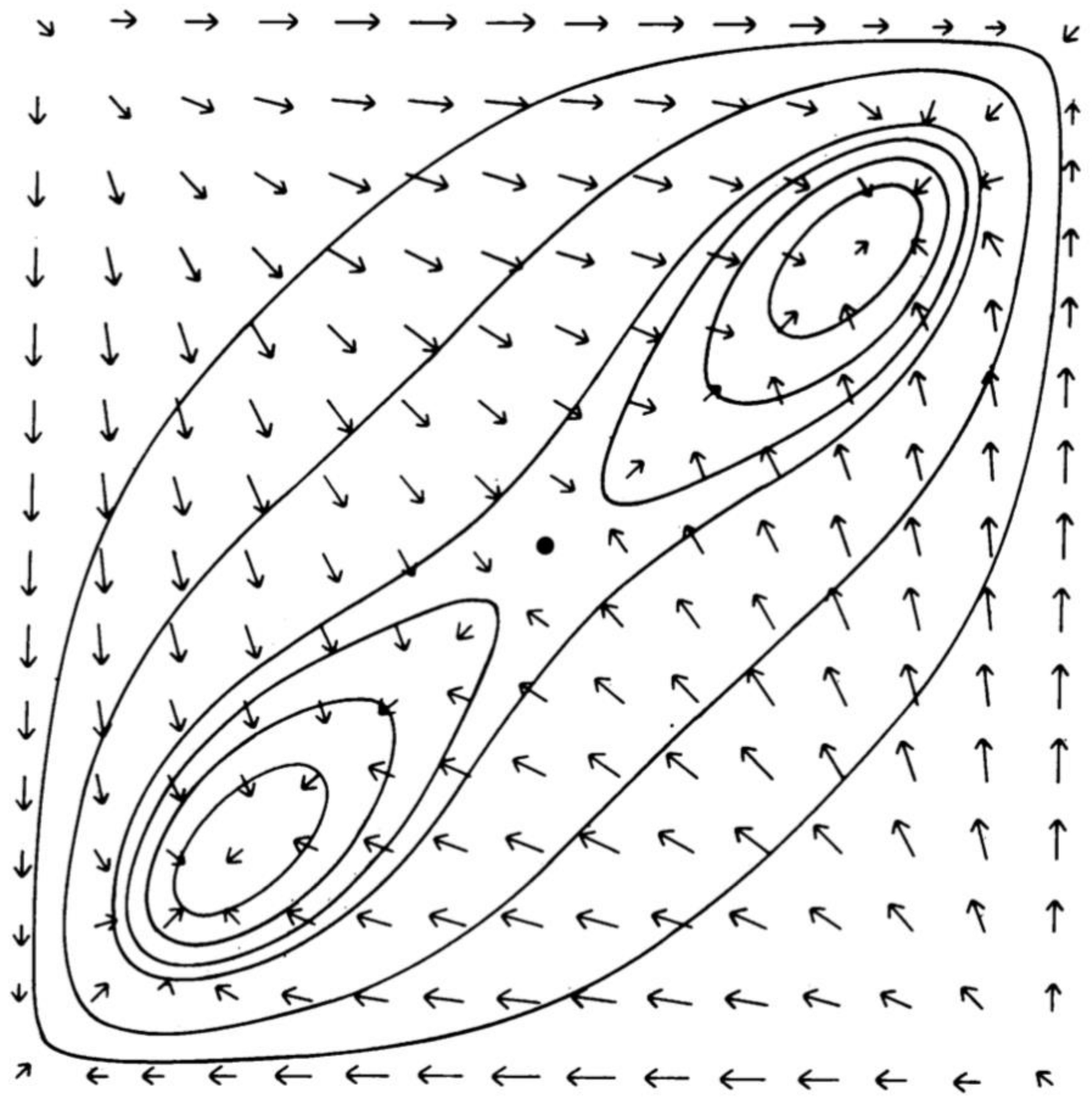


FIG. 3. An energy contour map for a two-neuron, two-stable-state system. The ordinate and abscissa are the outputs of the two neurons. Stable states are located near the lower left and upper right corners, and unstable extrema at the other two corners. The arrows show the motion of the state from Eq. 5. This motion is not in general perpendicular to the energy contours. The system parameters are $T_{12} = T_{21} = 1$, $\lambda = 1.4$, and $g(u) = (2/\pi)\tan^{-1}(\pi\lambda u/2)$. Energy contours are 0.449, 0.156, 0.017, −0.003, −0.023, and −0.041.

Hopfield, PNAS 1984

- Each of local minima in the energy landscape is known as an "attractor".
- The network can do pattern completion: retrieving the full pattern from a partial cue.
- The capacity of the network, or maximum number of attractors, is $\sim 0.14N$.

Content-addressable memory in Hopfield networks

Hertz, Krogh & Palmer, *Introduction to the Theory of Neural Computation* (1991)

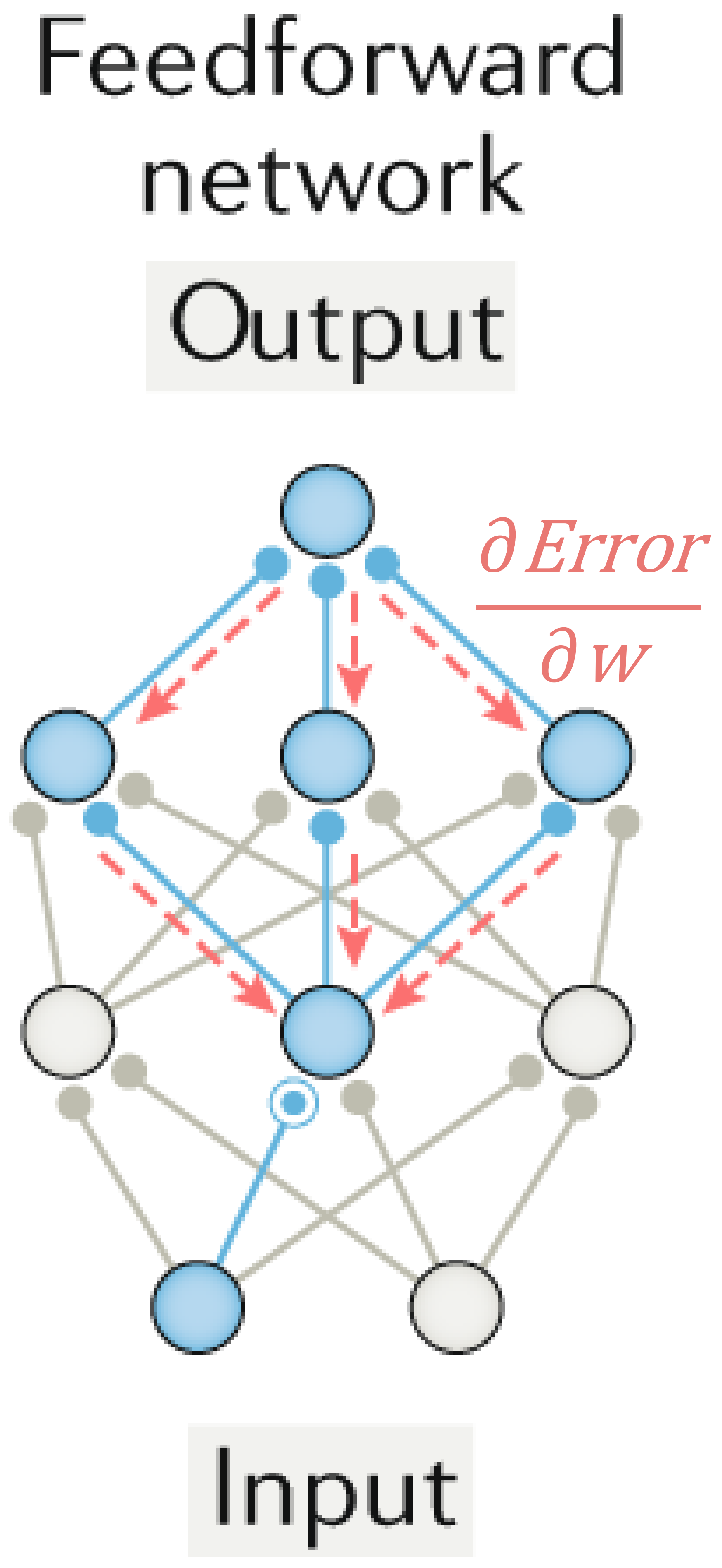
4. Links between learning in biological and artificial neural networks

The back-propagation learning rule

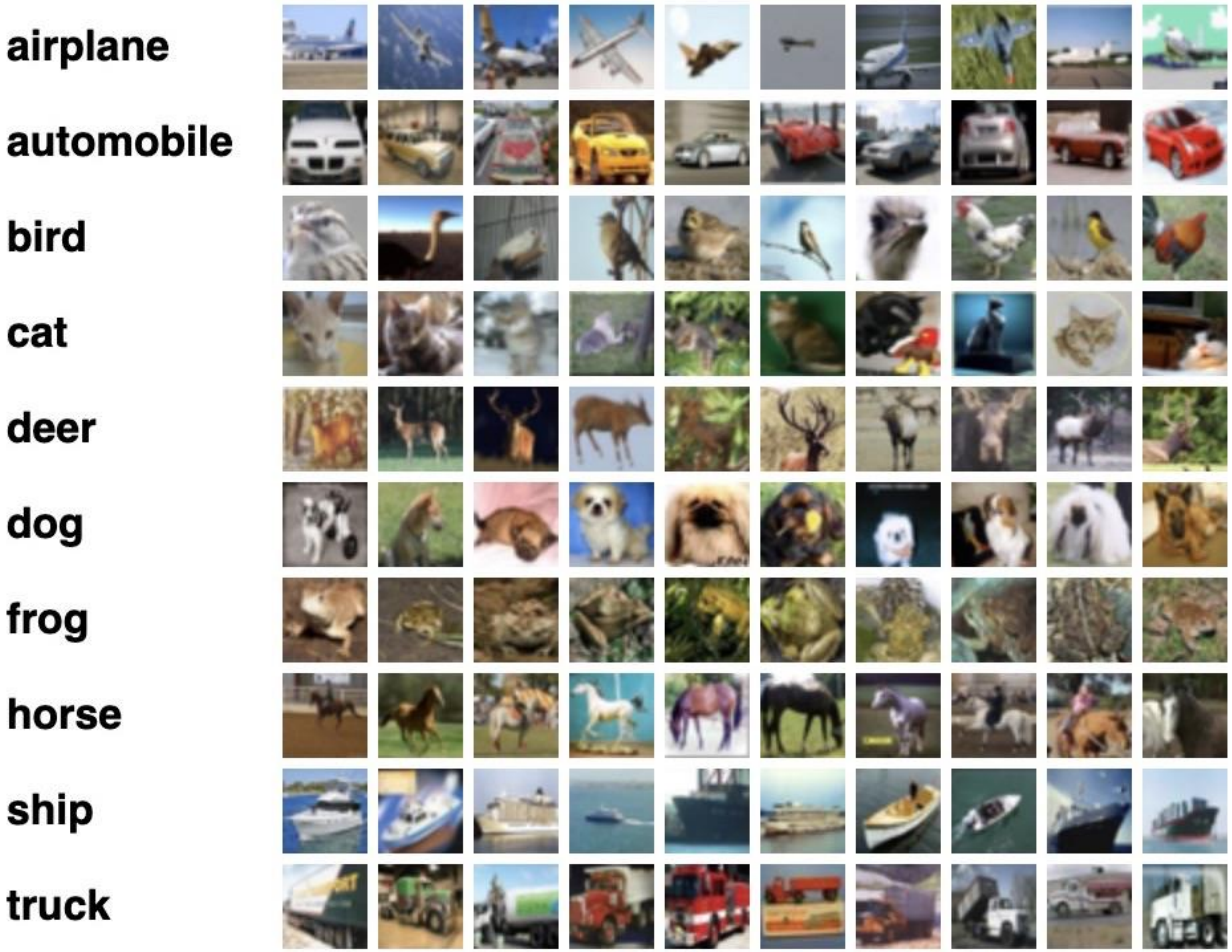
- Unlike Hebb's rule, in machine learning artificial neural networks are trained via supervised learning: performance errors are fed back into the system to adjust connection weights, so reducing future errors.

- The main method for doing this is called back-propagation (basically the calculus chain rule).

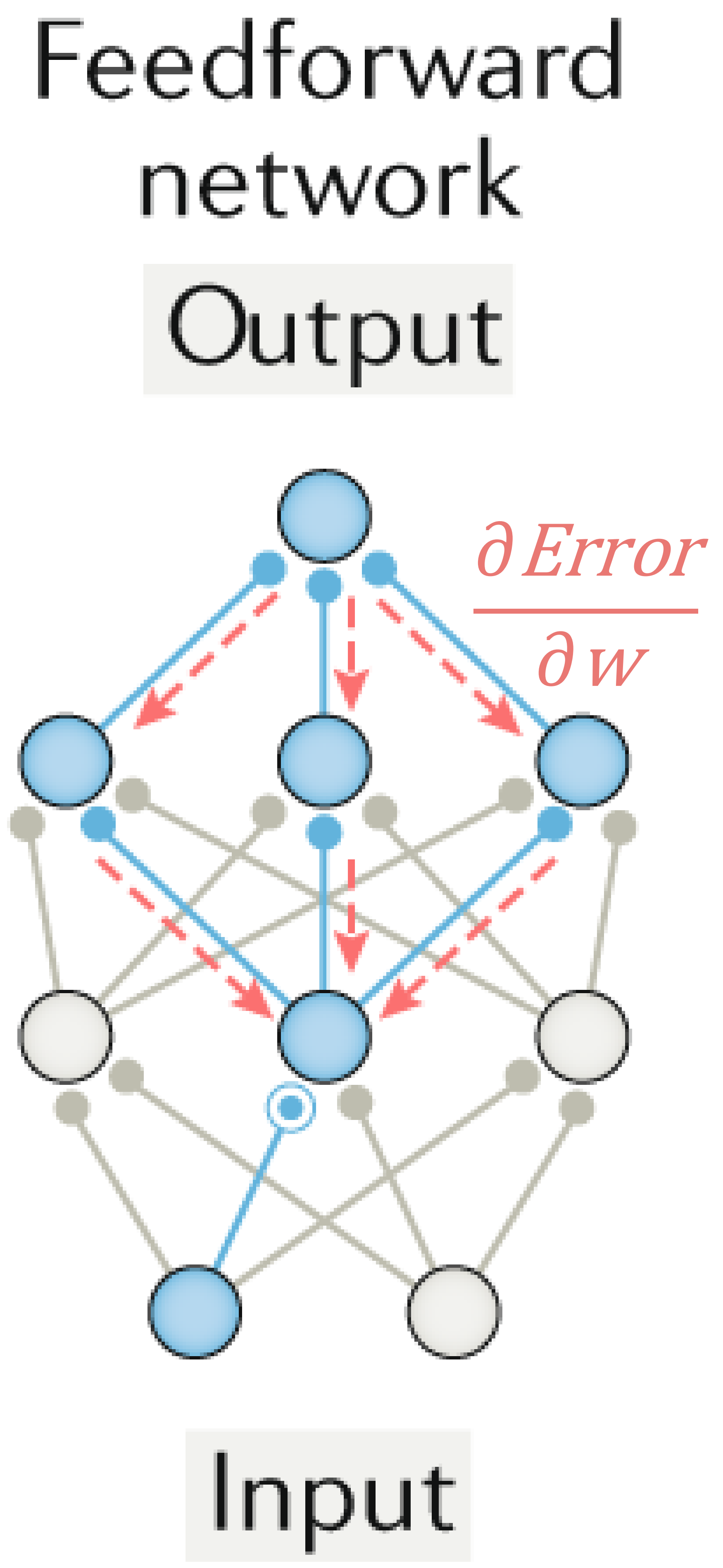
The back-propagation learning rule



Deep learning is everywhere



Issues with backprop as a model of brain learning



- Plasticity rules at each synapse needs non-local information.
- Requires symmetry in backwards and forwards weights.
- Error signals need to contain sign information (+/- error).
- Feedforward and feedback ‘passes’ need to be separated in time.

Example ideas for biologically plausible backprop

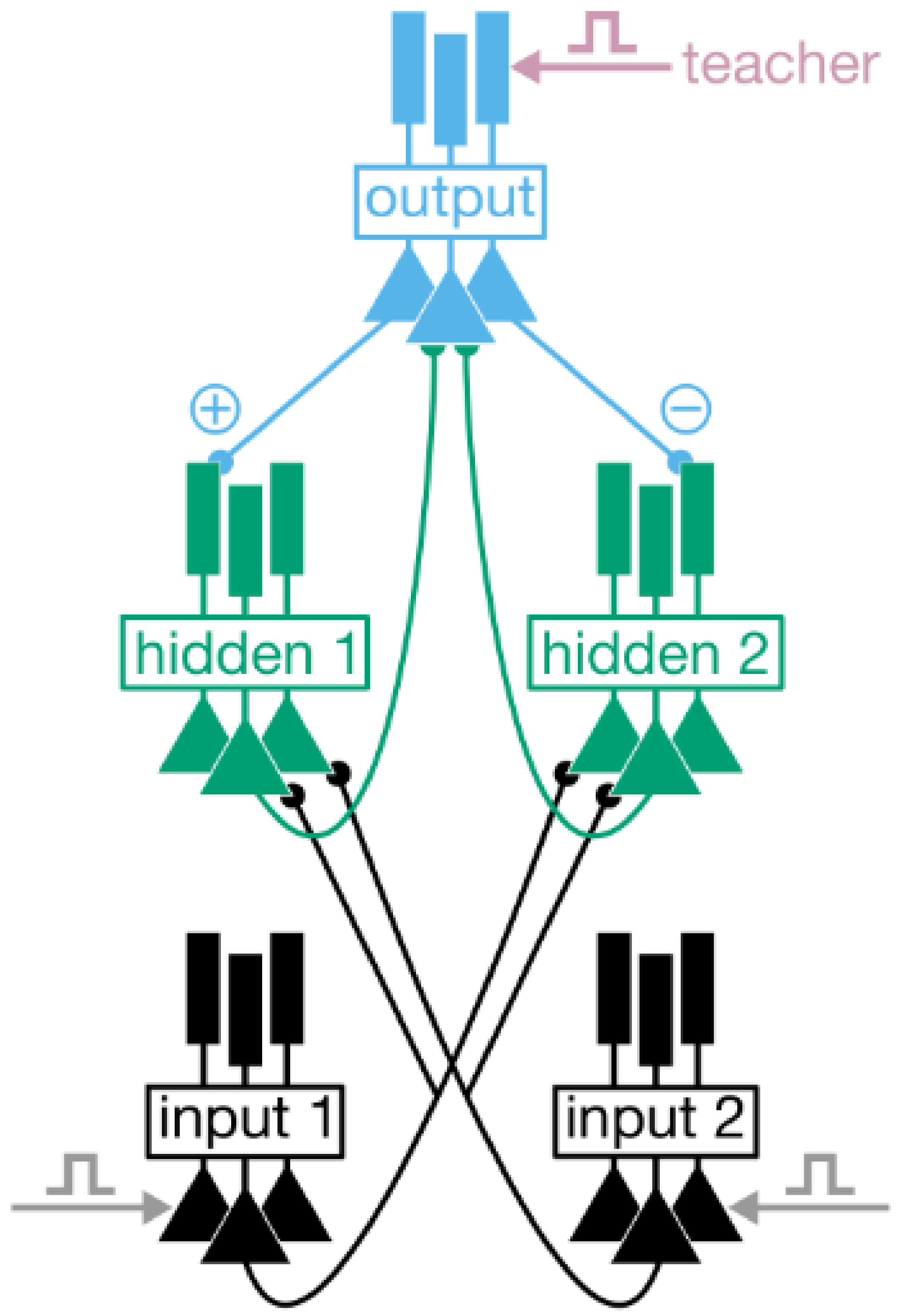
Feedback alignment even by random feedback.

"B pushes the network in roughly the same direction as backprop would"

$$\mathbf{e}^T W B \mathbf{e} > 0$$

Lillicrap et al, *Nature Comms* (2016)

Burst-prop: feedback signals to distal dendrites, driving bursts that gate plasticity.



Payeur et al, *Nature Neurosci* (2021)

Deep learning and brains summary

- Backpropagation of error gradients is a ubiquitous algorithm for learning in artificial neural networks.
- But the basic version in not biologically realistic.
- Current research is trying to dream up biologically plausible alternatives that approximate backprop.
- However the idea that the brain even learns by supervised gradient descent is controversial! Maybe unsupervised or reinforcement learning rules are all it takes.

Overall session summary

1. Synapses transmit signals between neurons. Changes in their strength underlie long-term memory.

2. Classic plasticity rules: Hebb, STDP.

3. The hippocampus as a Hopfield attractor network.

4. Does the brain do deep learning?? Maybe.

Thanks