UNIVERSITY OF OXFORD

MRC Brain Network Dynamics Unit

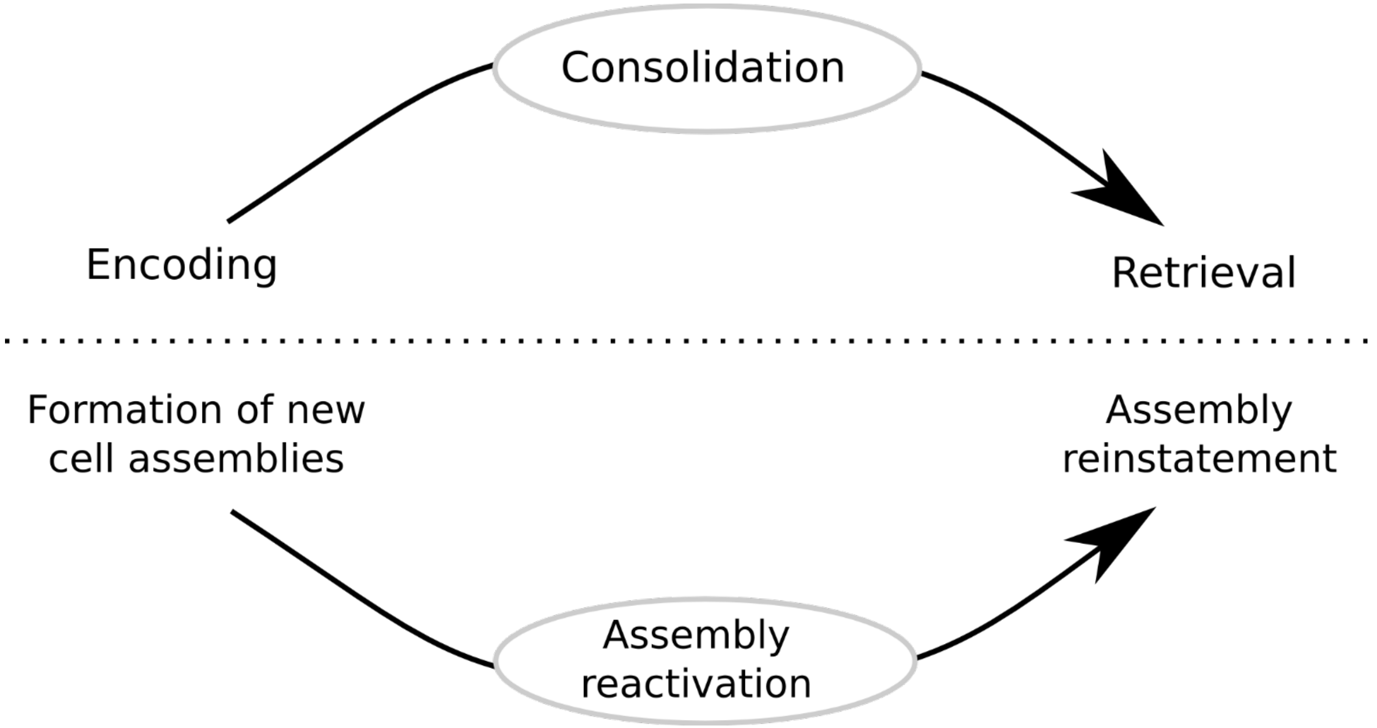
Baylor College of Medicine

Reactivation in Biological and Artificial Neural Networks

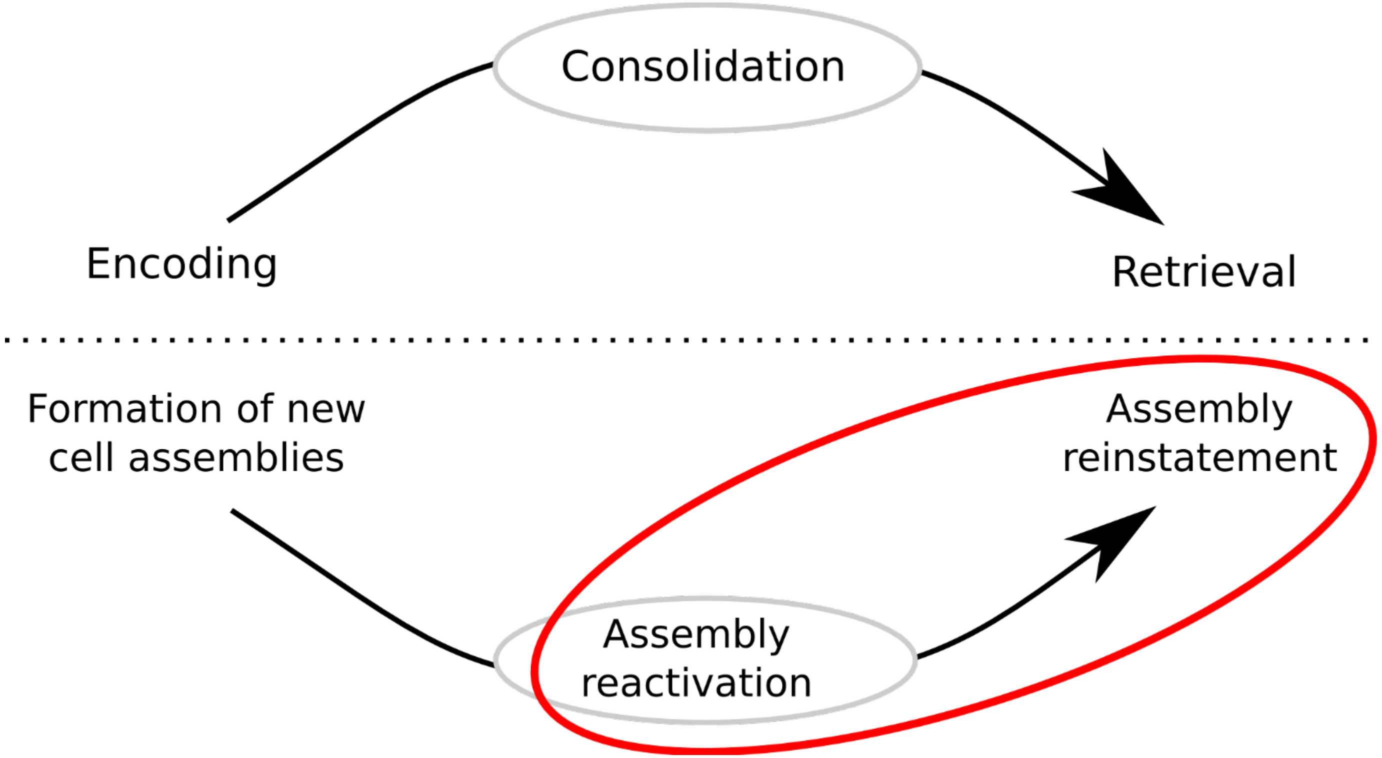
Gido van de Ven

August 2018 – UK-Korea Symposium

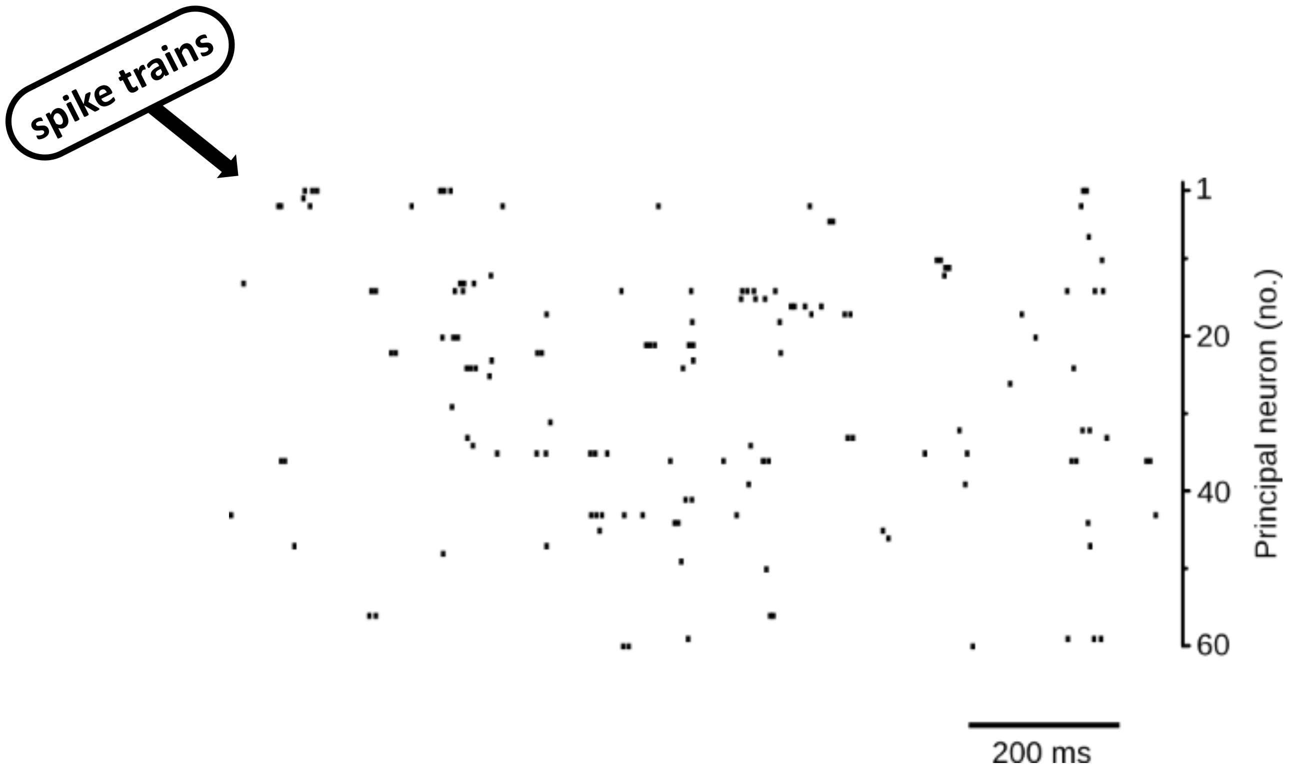
Memory consolidation in the brain: cell assembly / reactivation hypothesis

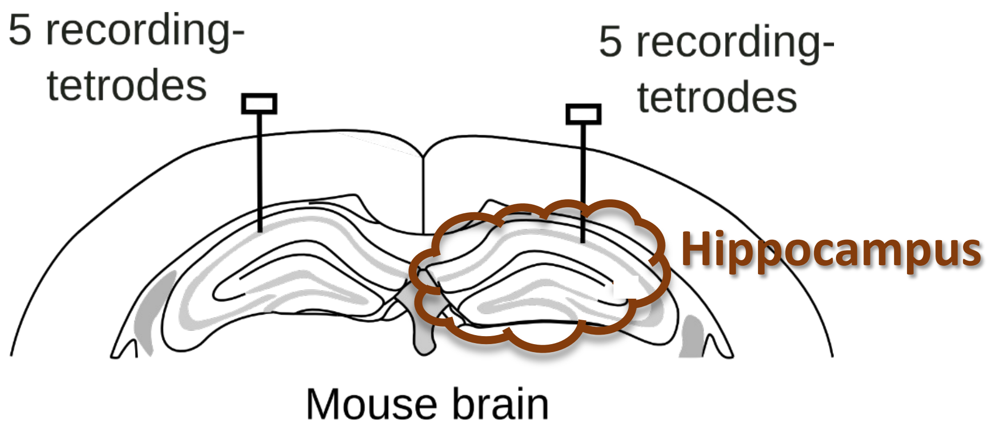


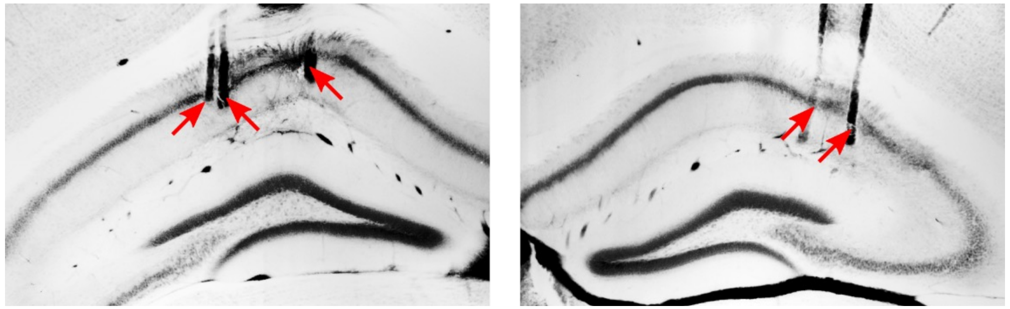
Memory consolidation in the brain: cell assembly / reactivation hypothesis



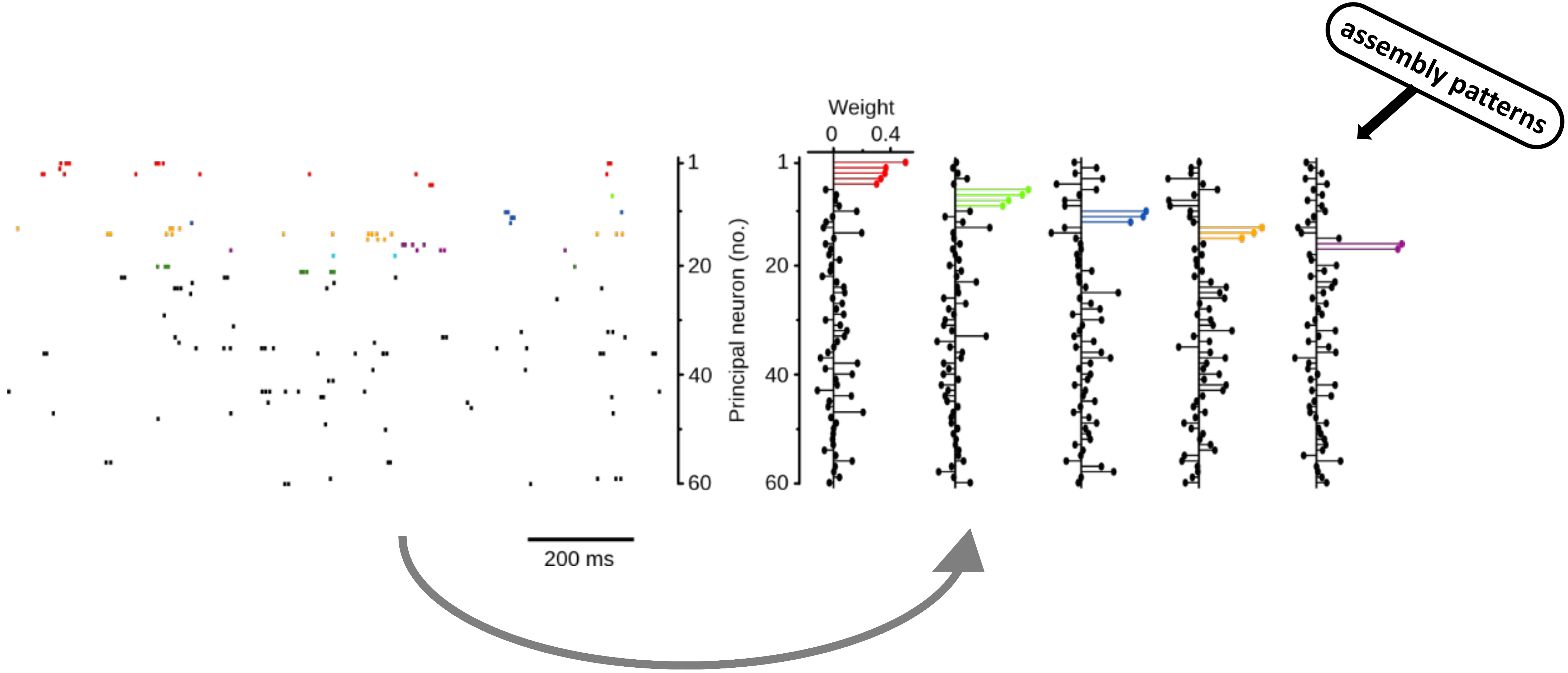
Identification of memory-representing “cell assemblies”





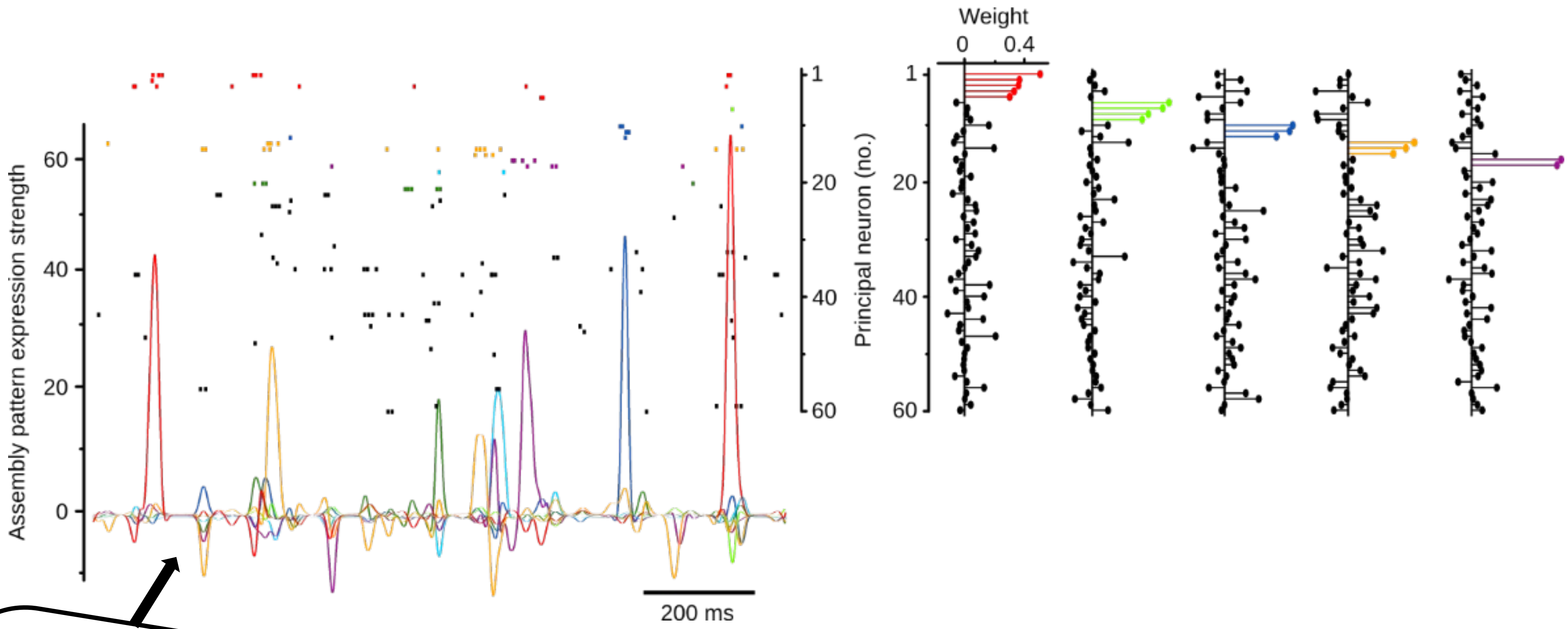


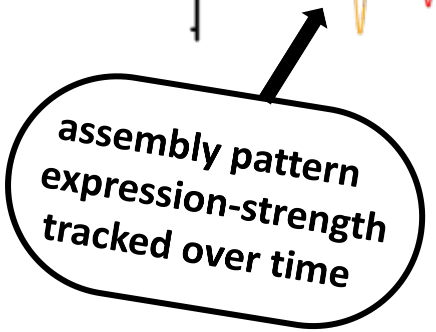
Identification of memory-representing “cell assemblies”



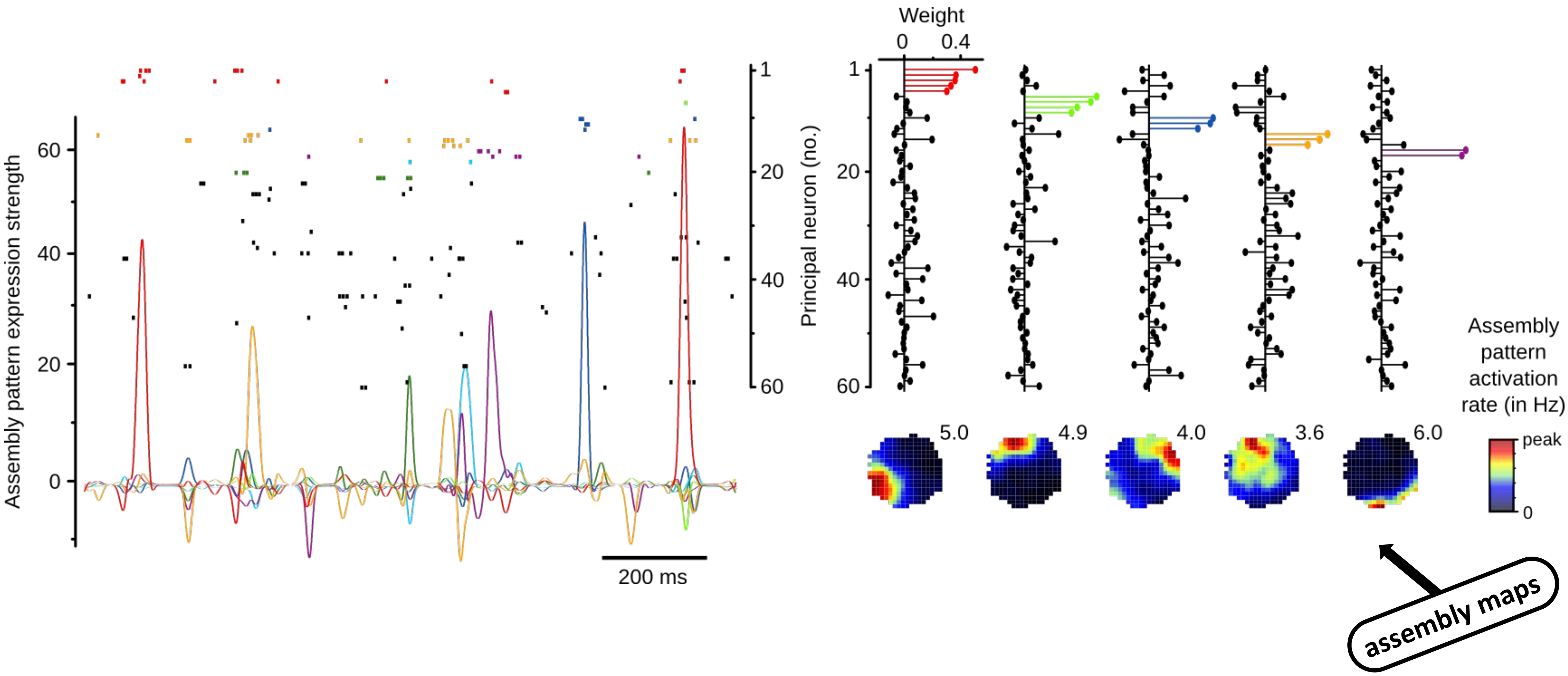
Assembly-detection method based on PCA and ICA

Identification of memory-representing “cell assemblies”

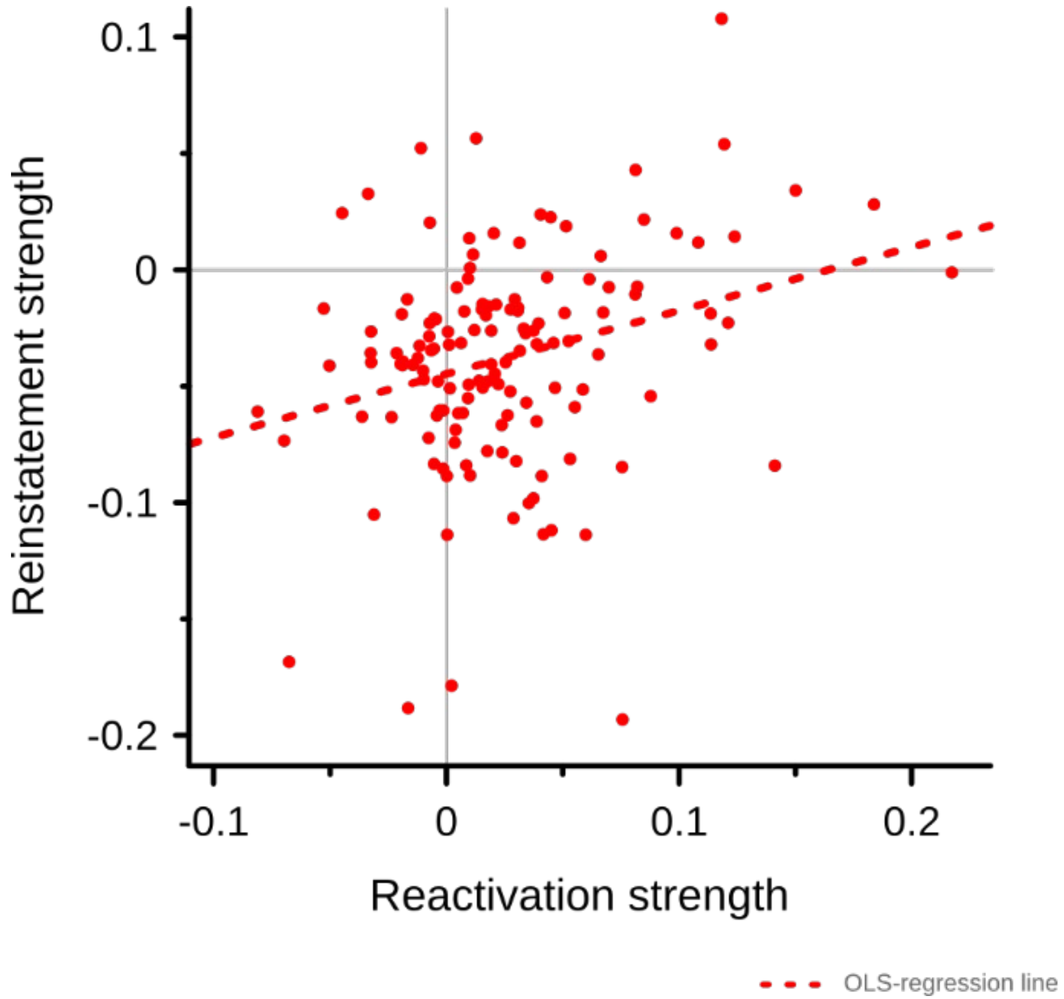


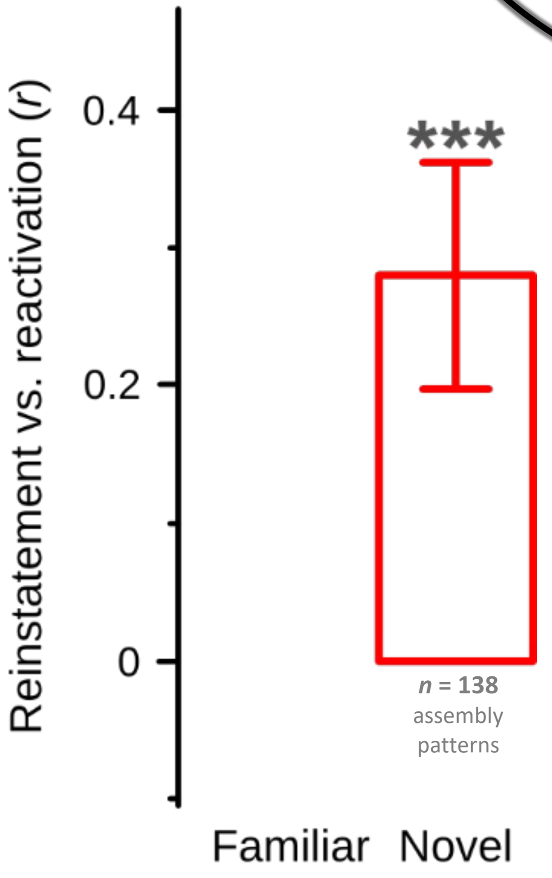


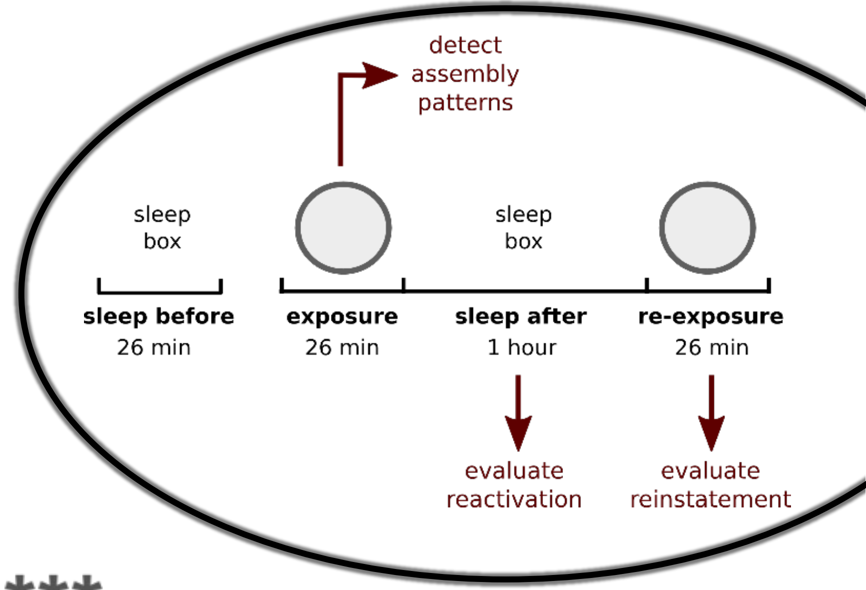
Identification of memory-representing “cell assemblies”



An assembly pattern's reactivation predicts it subsequent reinstatement

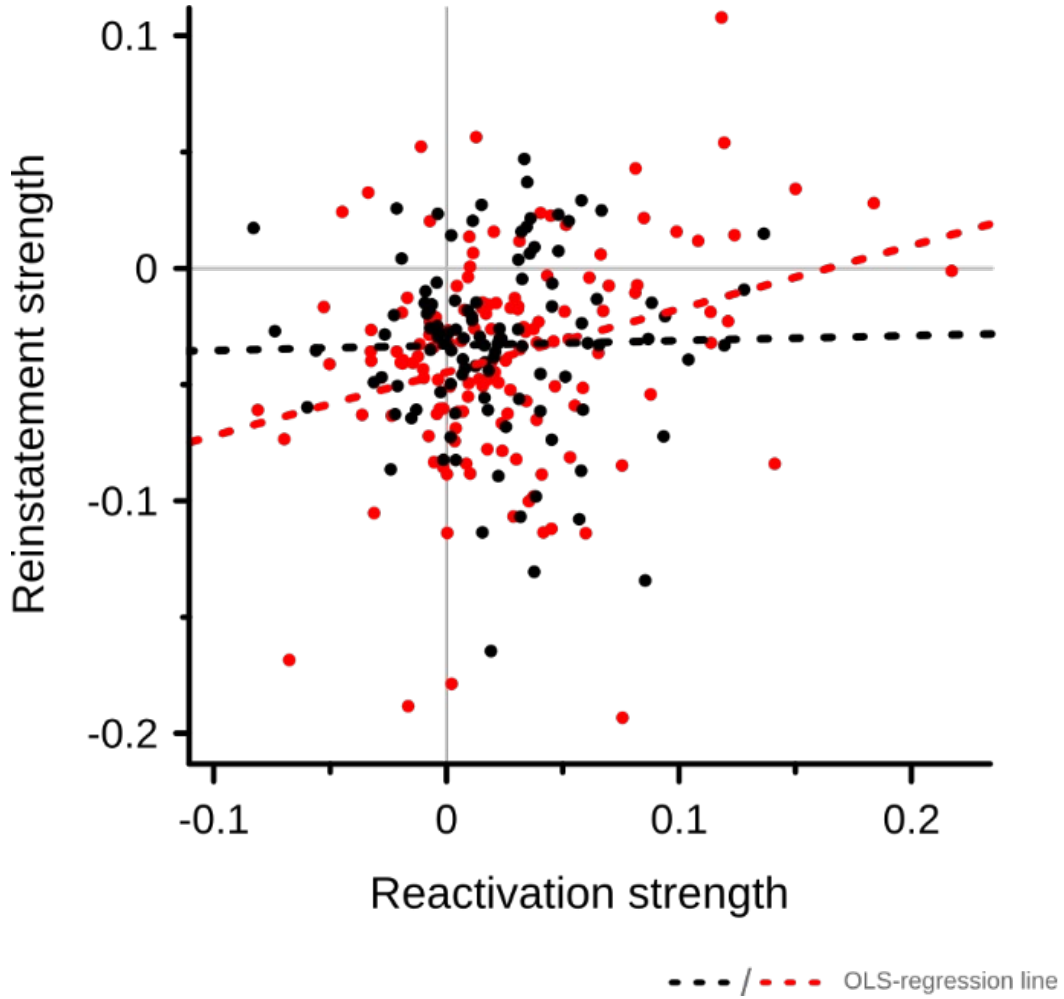


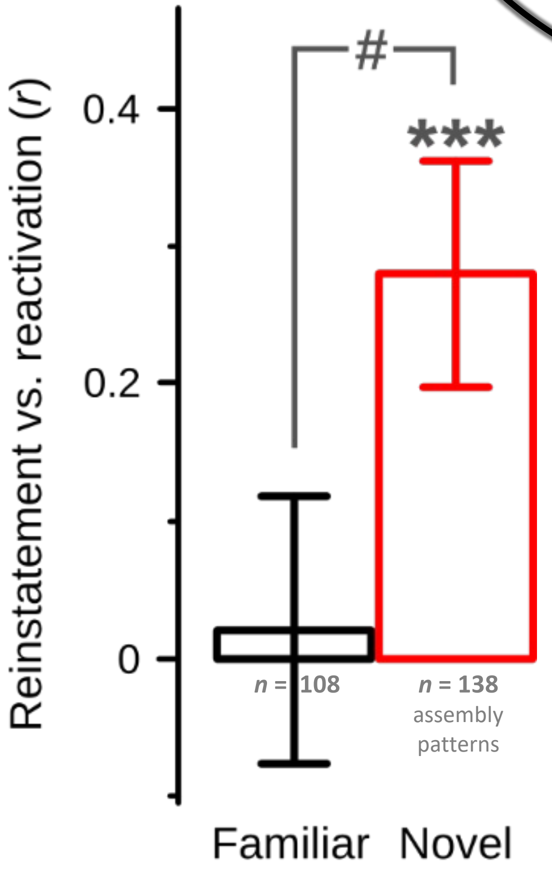


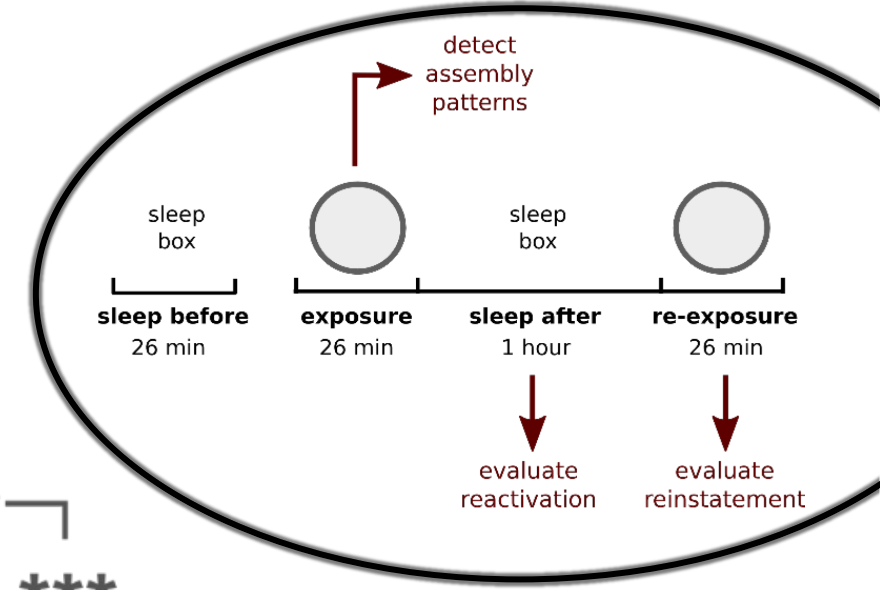


(based on 43 recording-blocks from 8 mice)

An assembly pattern's reactivation predicts it subsequent reinstatement

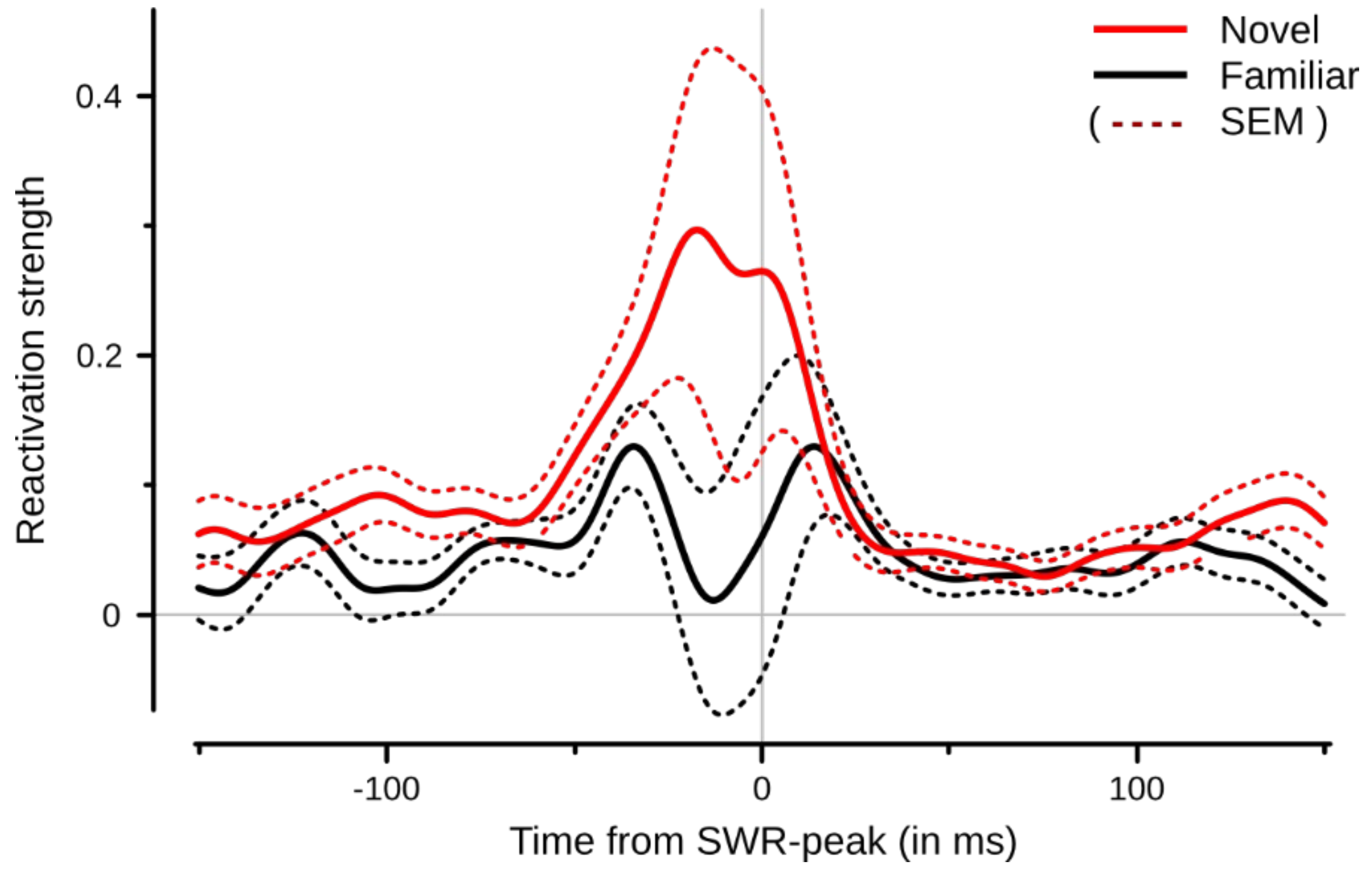






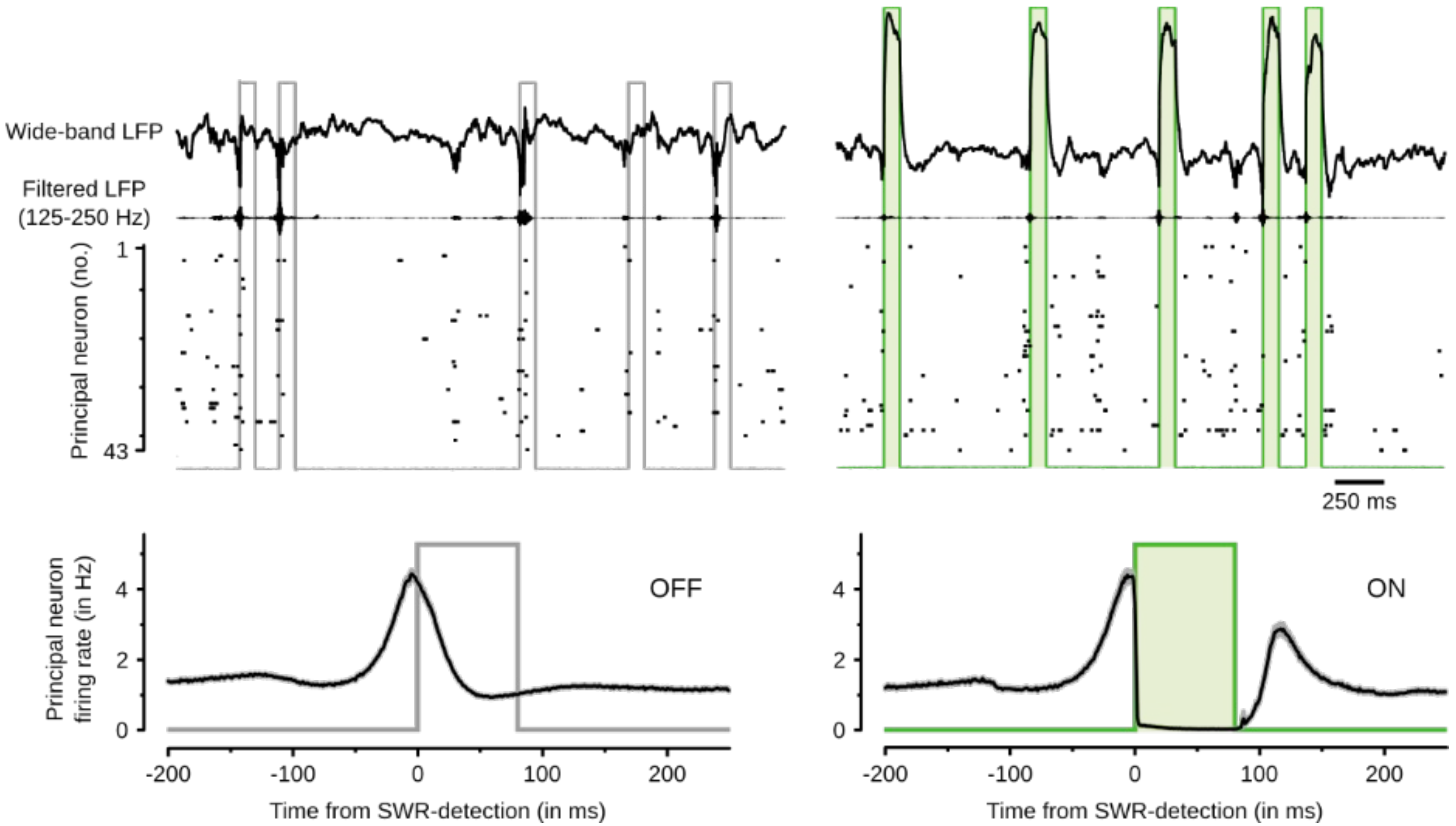
(based on 43 recording-blocks from 8 mice)

Selective disruption of reactivation?



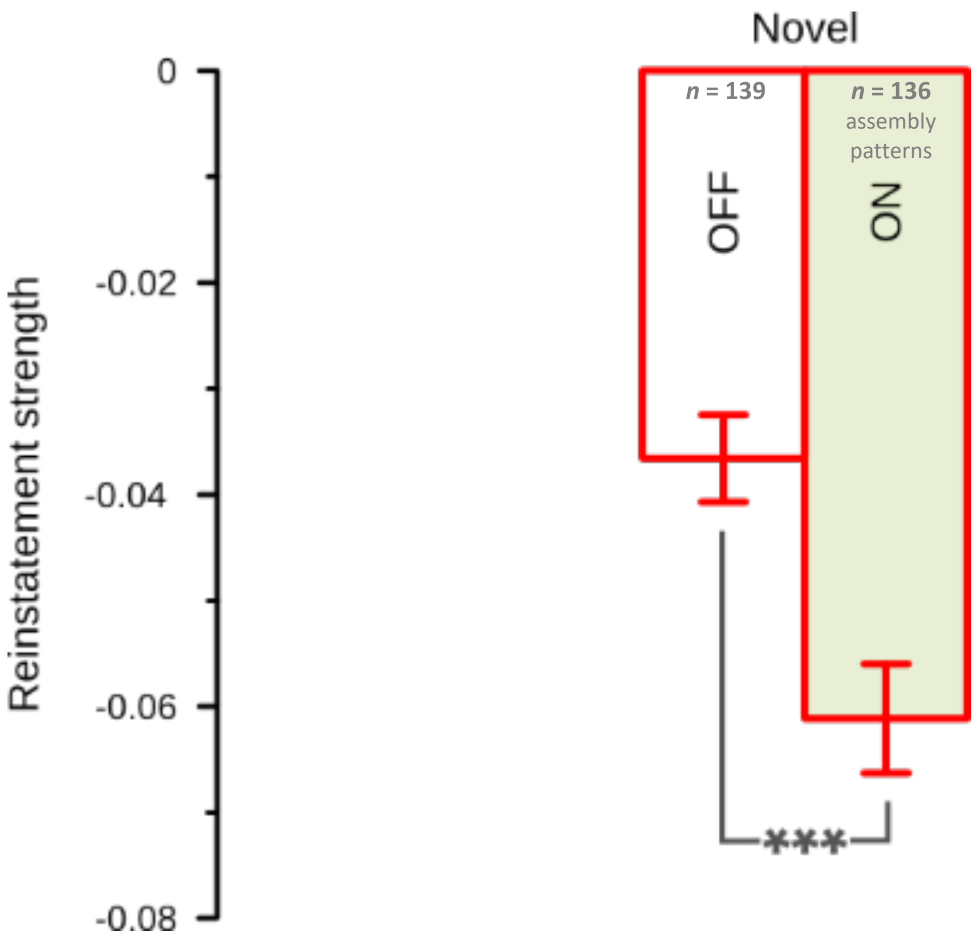
Novel: n = 139 assembly-patterns
Familiar: n = 108 assembly-patterns
(based on 43 recording-blocks from 8 mice)

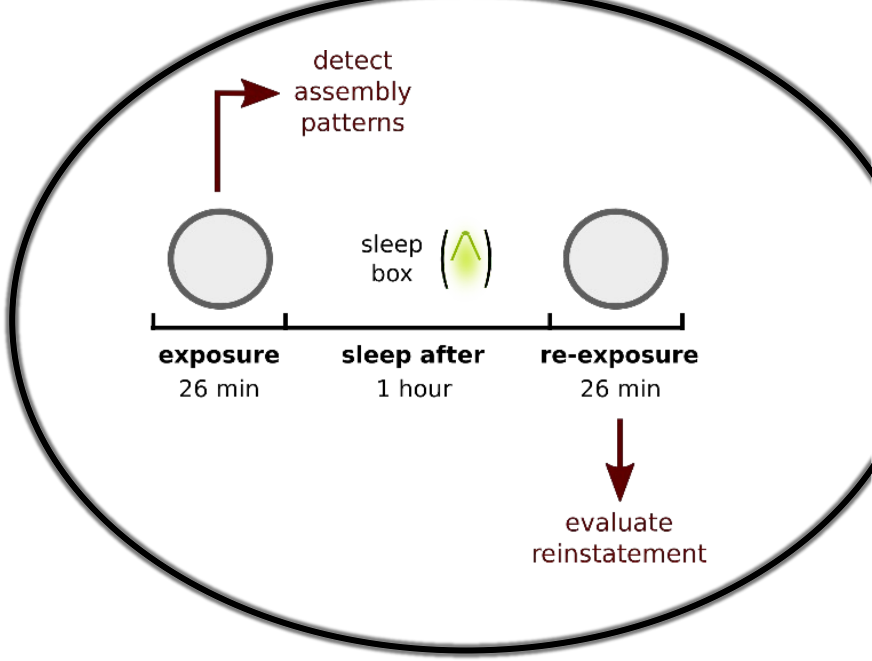
Selective disruption of reactivation: *optogenetic SWR silencing*



OFF: n = 1,988 neurons (from 43 sessions)
ON: n = 1,527 neurons (from 37 sessions)

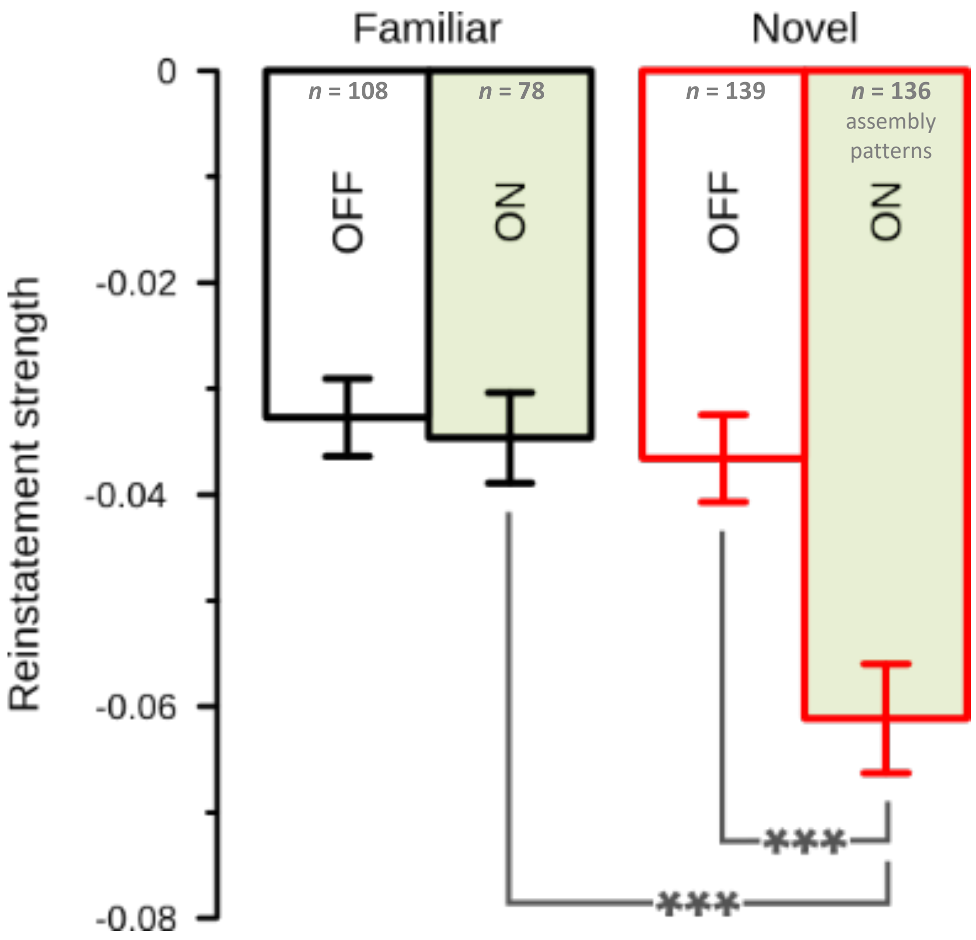
SWR-silencing impairs assembly pattern reinstatement



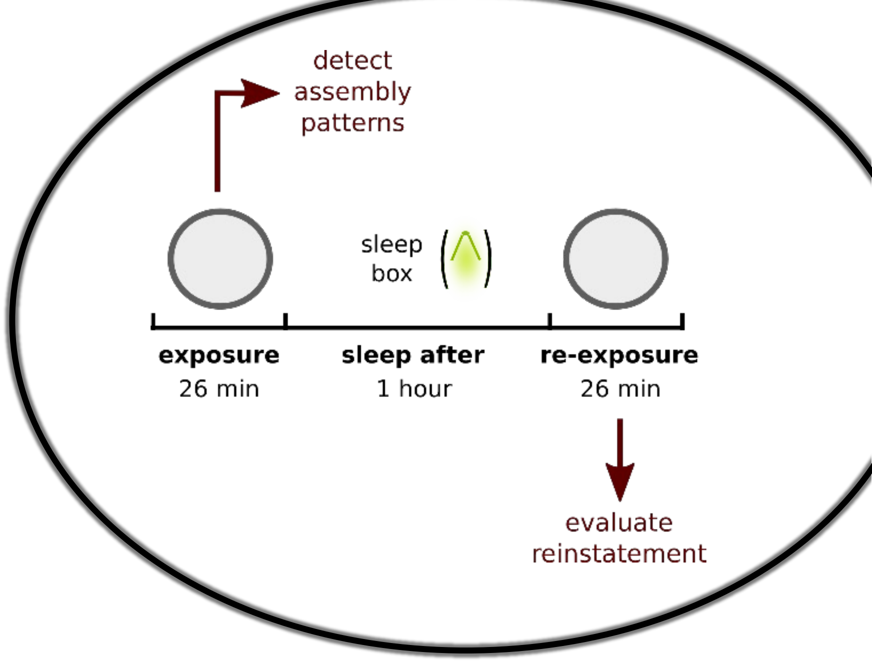


(based on 50 recording-blocks from 8 mice)

SWR-silencing impairs assembly pattern reinstatement



interaction SWR-silencing x enclosure type:
$F(1,318) = 5.05$, $P < 0.05$



(based on 80 recording-blocks from 8 mice)

Summary **Part I**

- In the brain, reactivation stabilizes recently-formed, memory-representing cell assembly patterns

Further details: van de Ven *et al.* (2016) *Neuron* [+ video abstract], or ask me for more slides!

But …

- *How* does reactivation stabilize these patterns?
- *Why* do memory-representations need to be gradually stabilized? Why are they not just stored “in one go”?

Approach

- Artificial neural networks as “model organism”

Summary **Part I**

- In the brain, reactivation stabilizes recently-formed, memory-representing cell assembly patterns

Further details: van de Ven *et al.* (2016) *Neuron* [+ video abstract], or ask me for more slides!

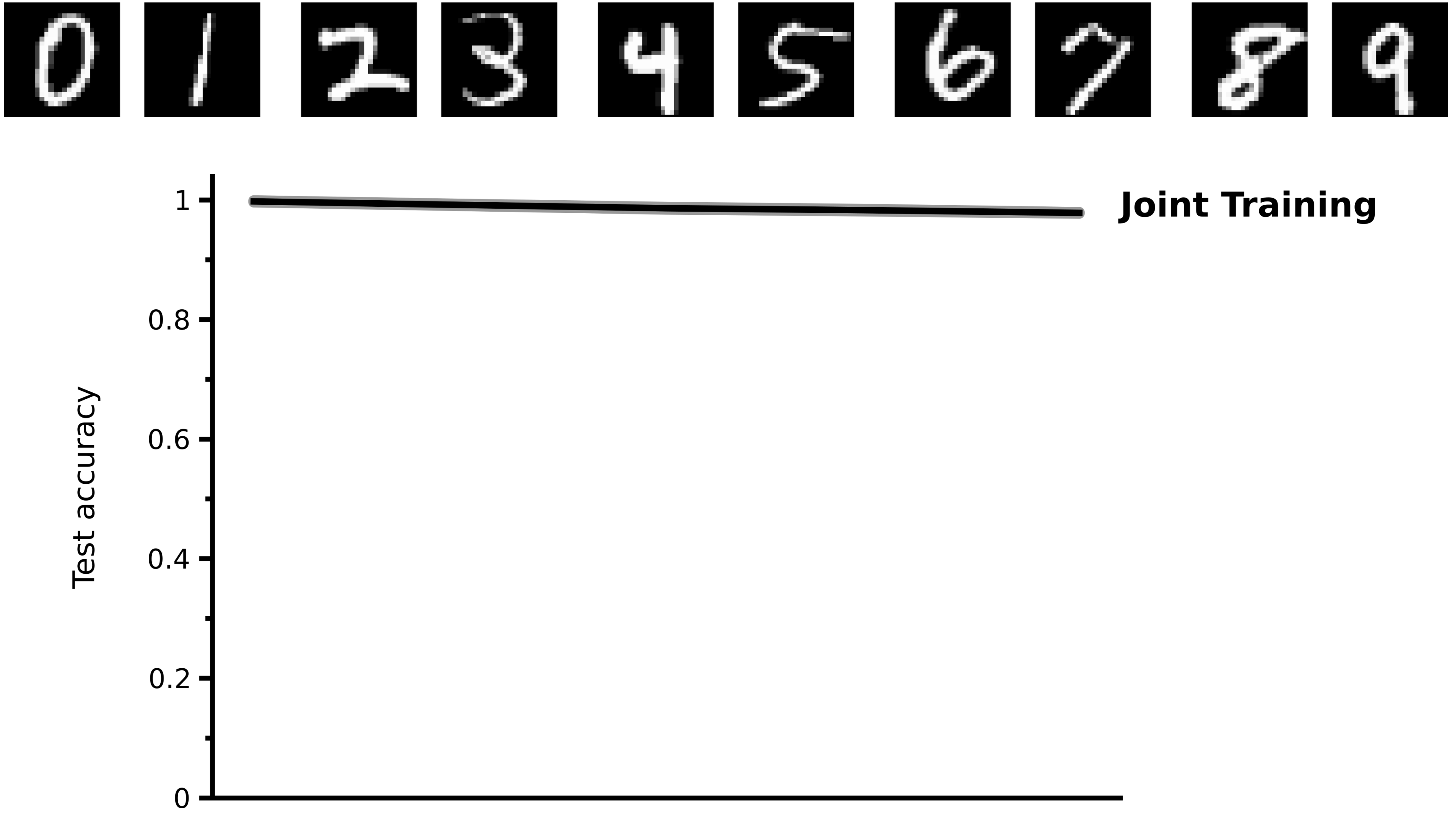
But …

- *How* does reactivation stabilize these patterns?
- *Why* do memory-representations need to be gradually stabilized? Why are they not just stored “in one go”?

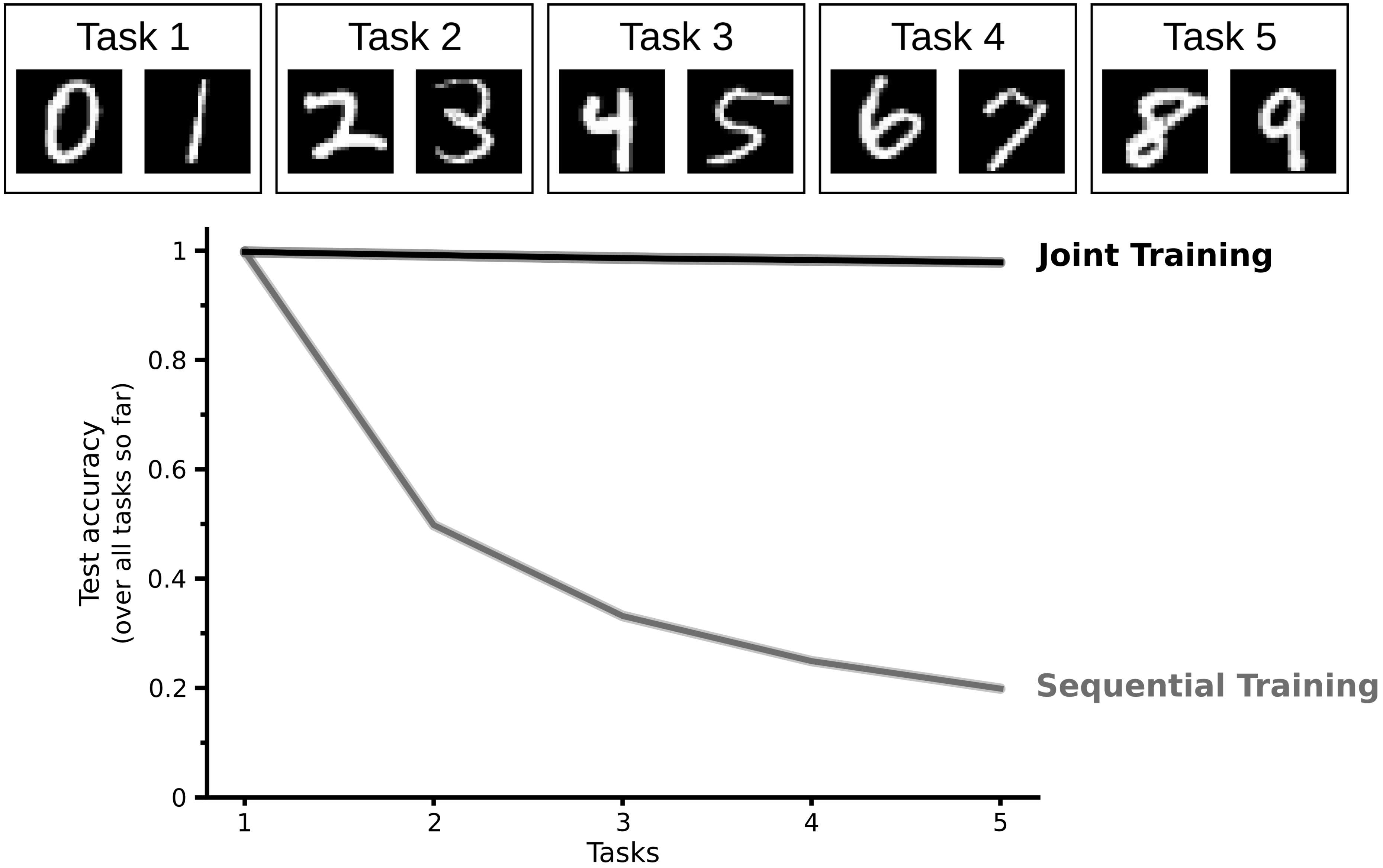
Approach

- Artificial neural networks as “model organism”

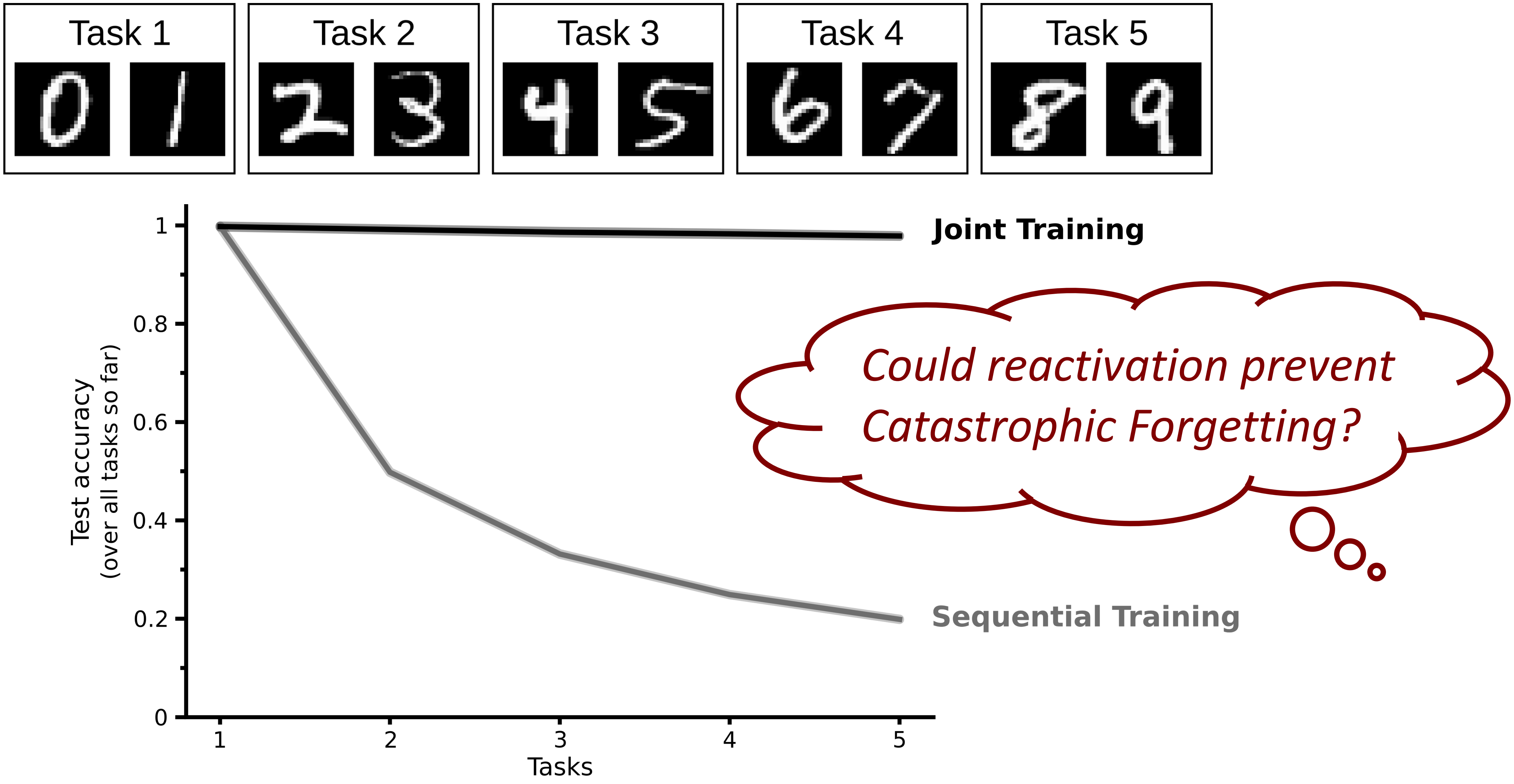
Catastrophic Forgetting in Artificial Neural Networks



Catastrophic Forgetting in Artificial Neural Networks

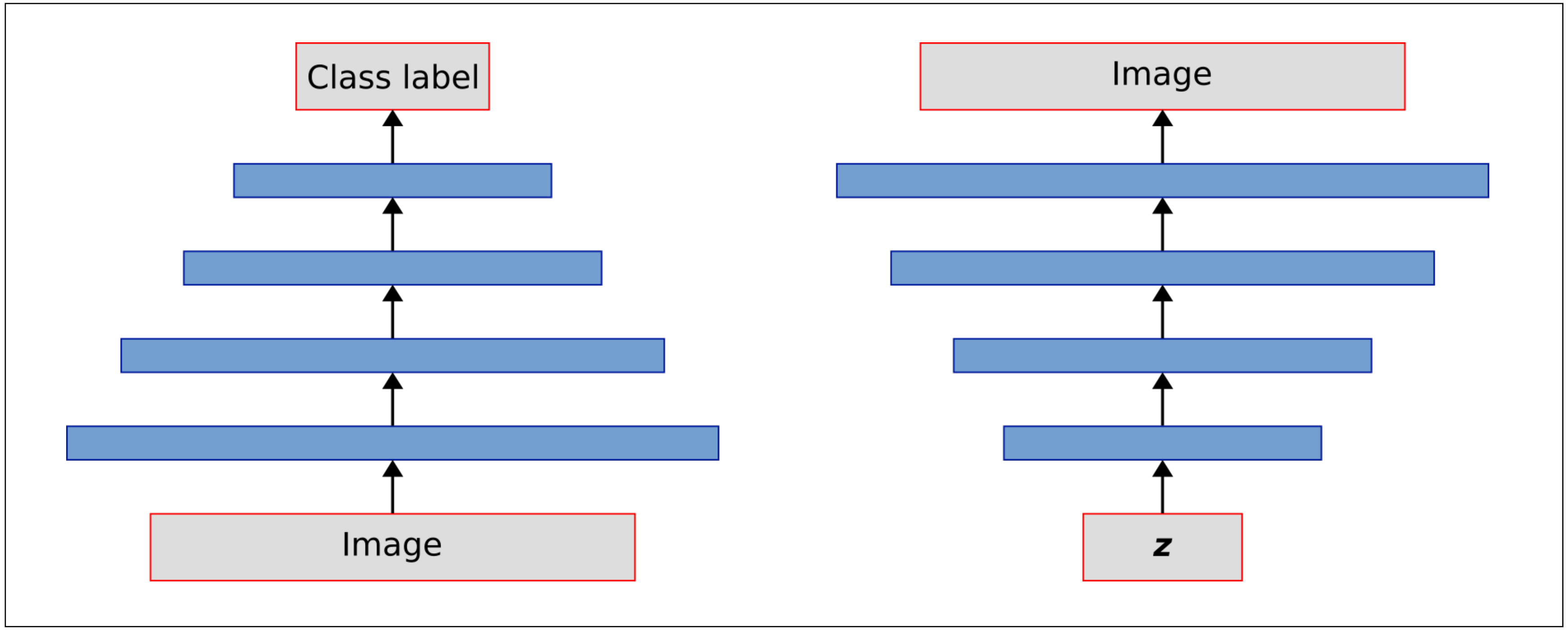


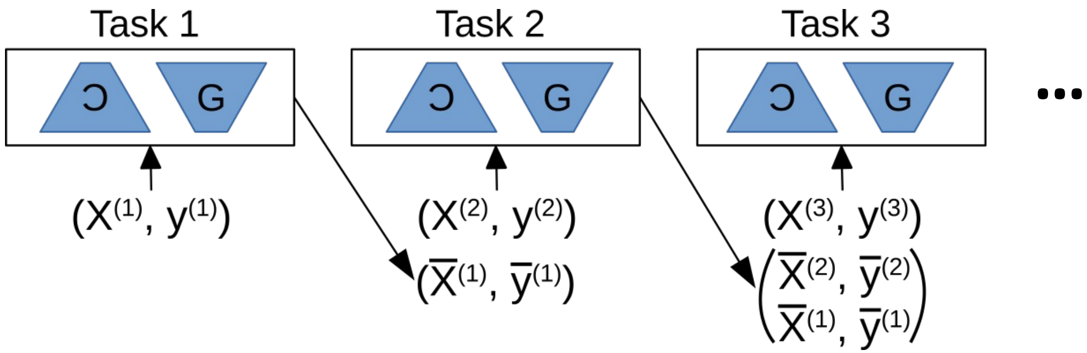
Catastrophic Forgetting in Artificial Neural Networks



Generative Replay

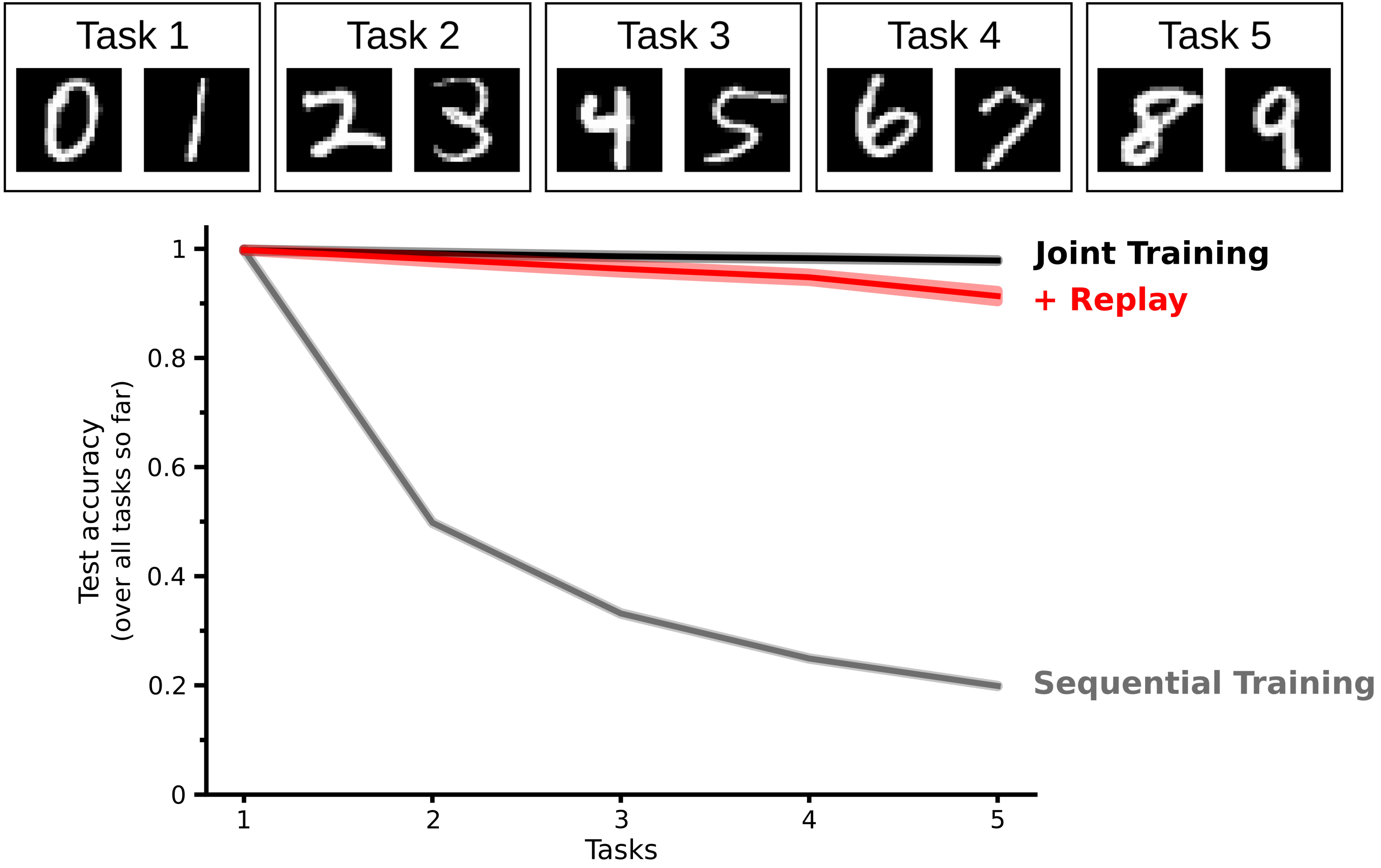
Classifier | **Generator** (e.g., GAN or VAE)



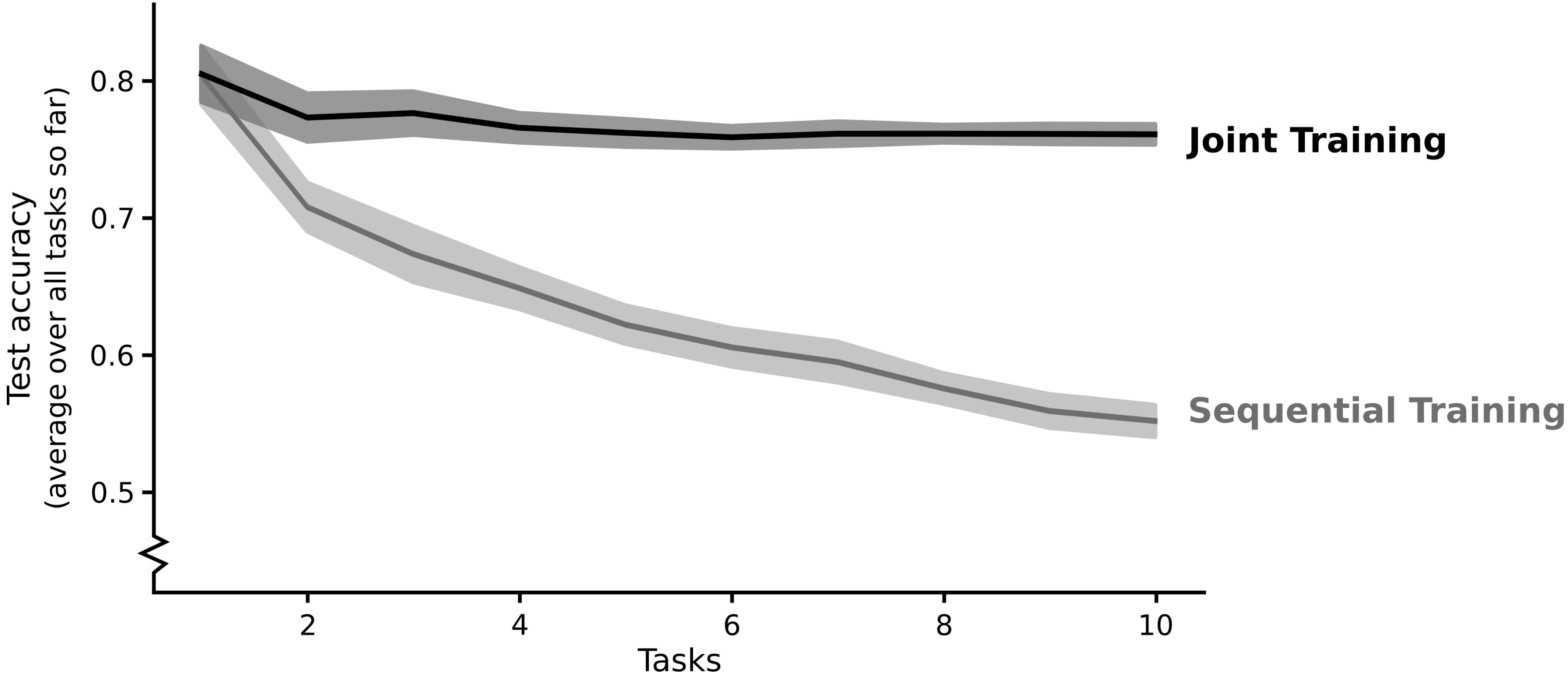


Shin et al. (2017) *NIPS*

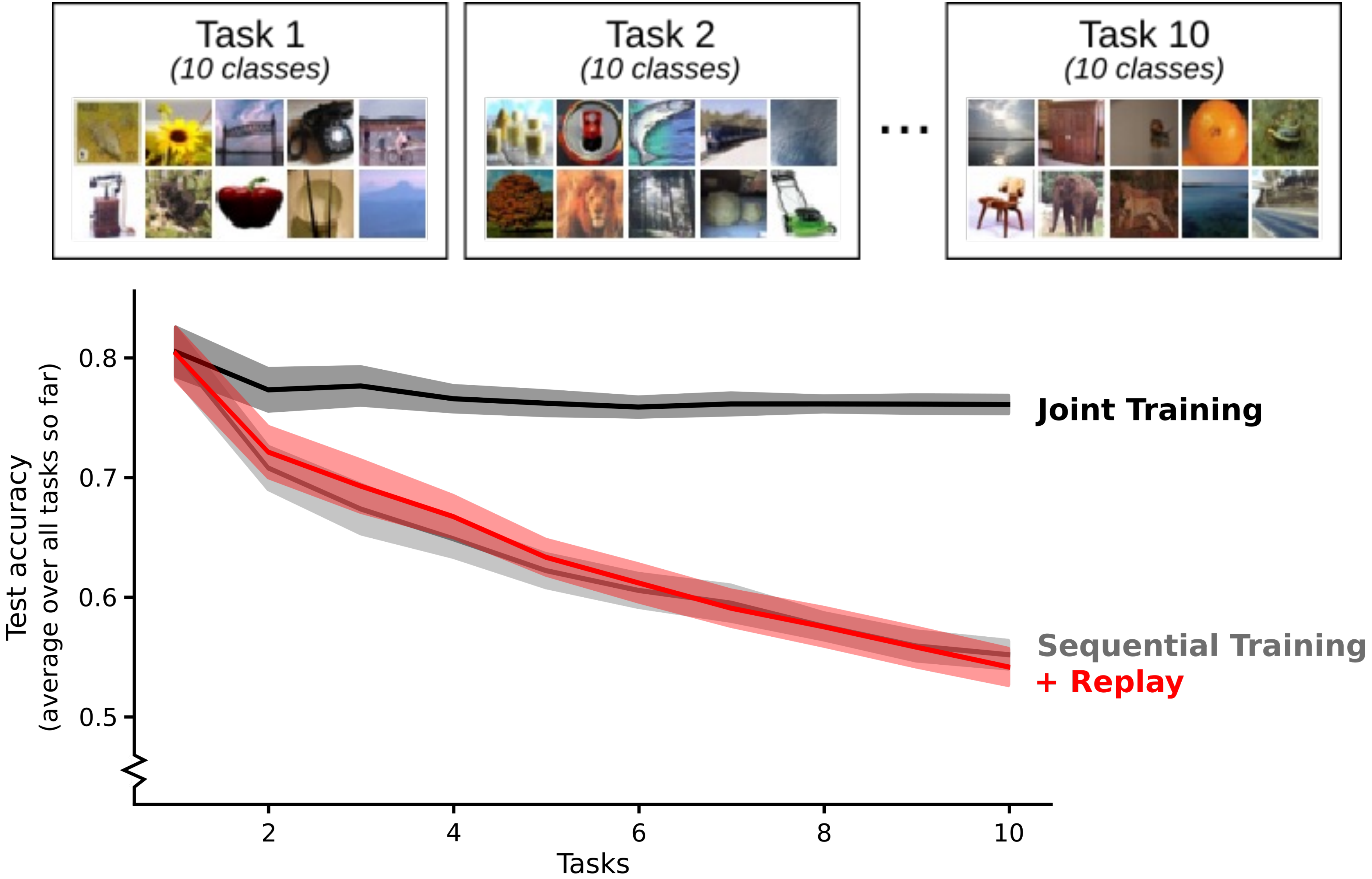
Generative Replay can prevent Catastrophic Forgetting



Is Generative Replay scalable to more complicated inputs?



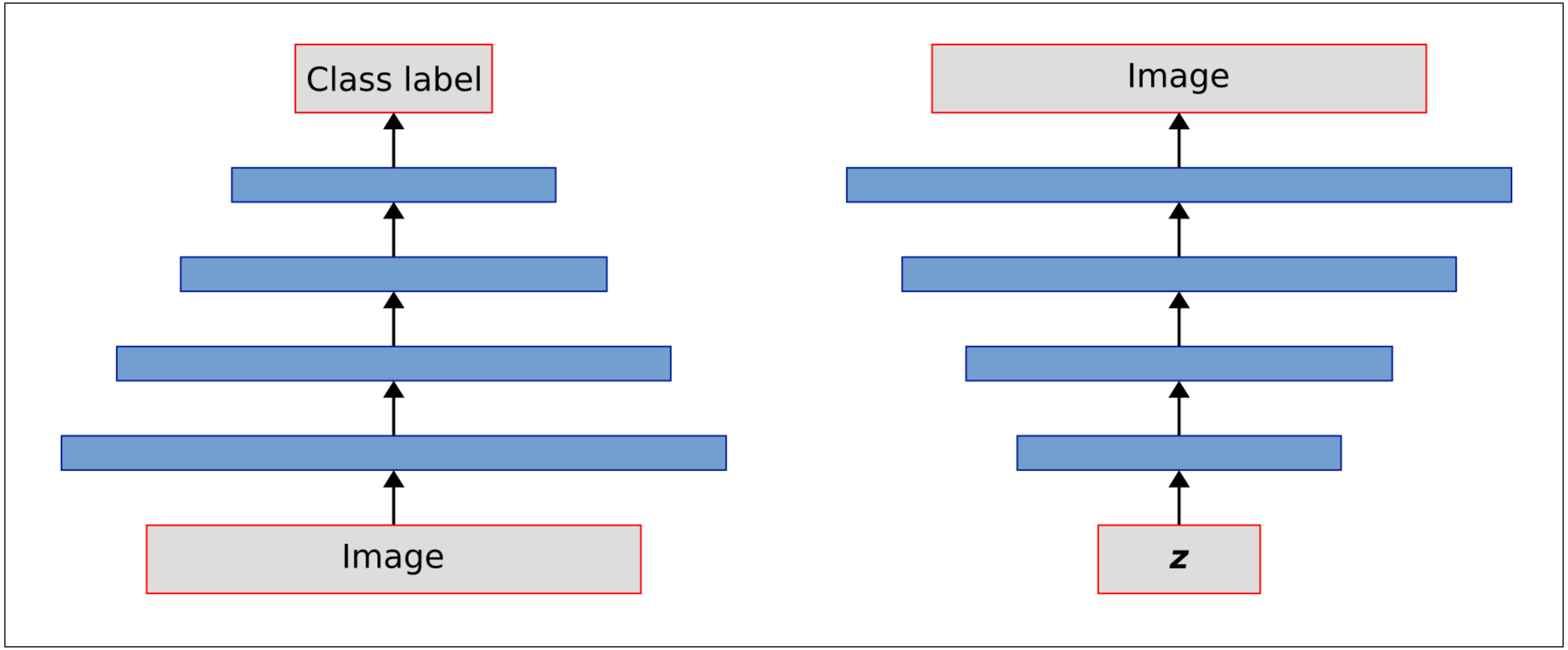
Is Generative Replay scalable to more complicated inputs?



Brain-inspired modifications to Generative Replay:

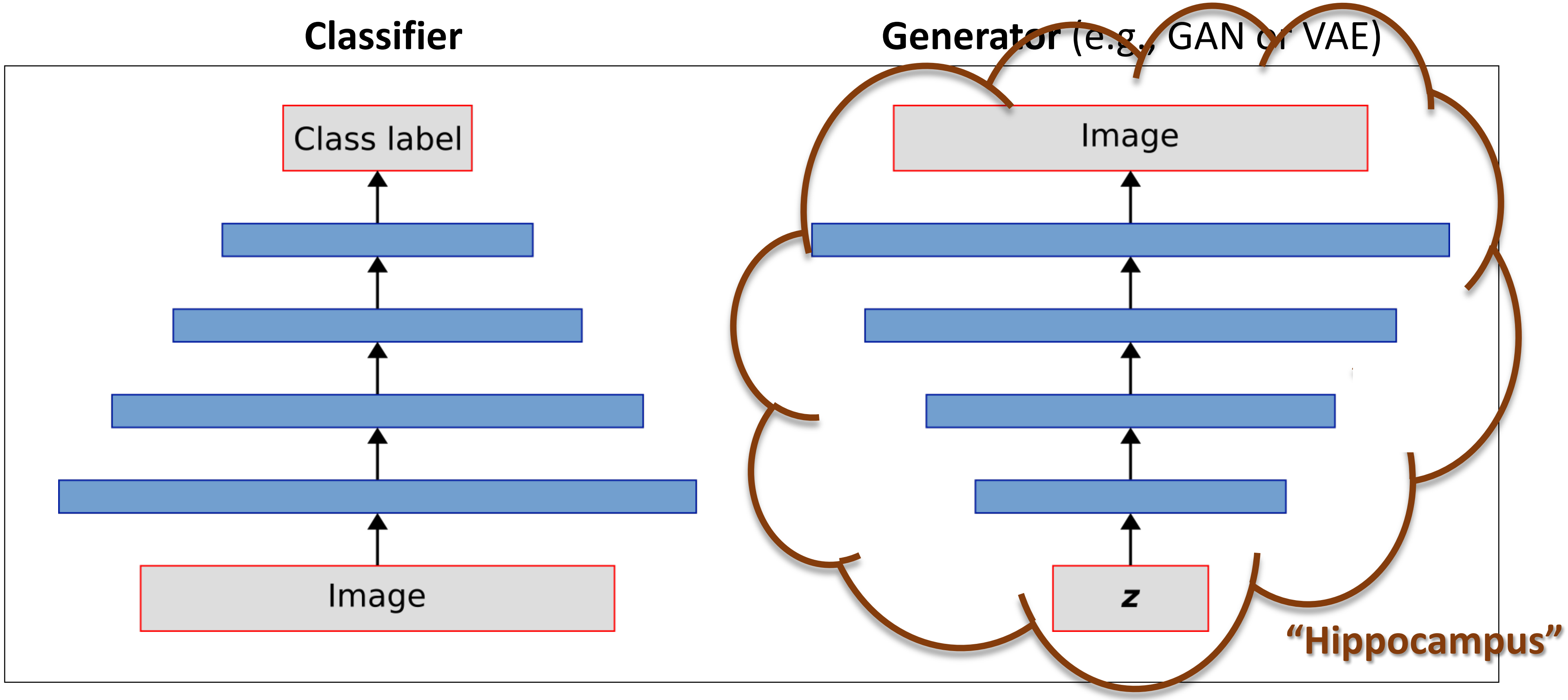
- through feedback connections
- replay hidden representations

Classifier **Generator** (e.g., GAN or VAE)



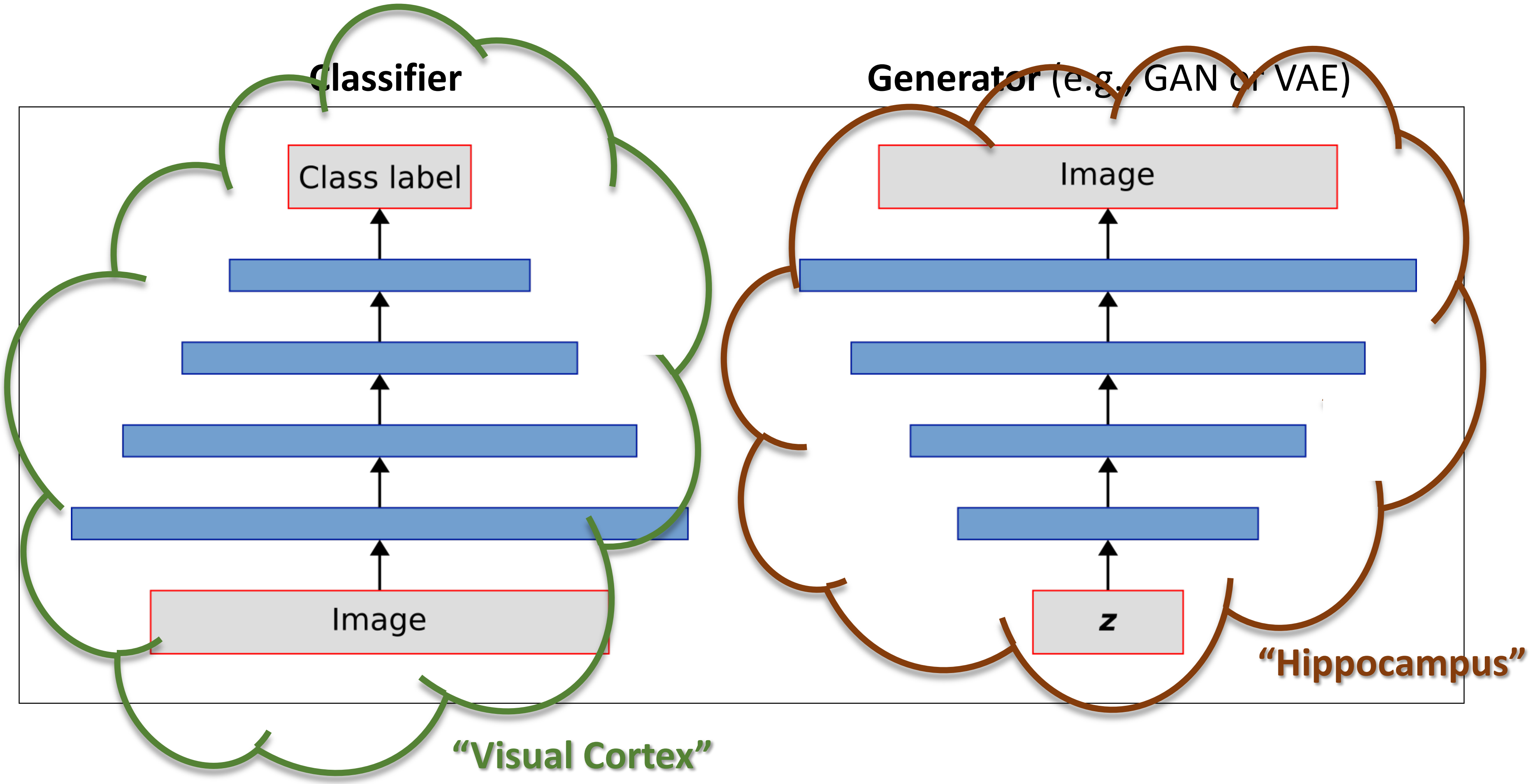
Brain-inspired modifications to Generative Replay:

- through feedback connections
- replay hidden representations



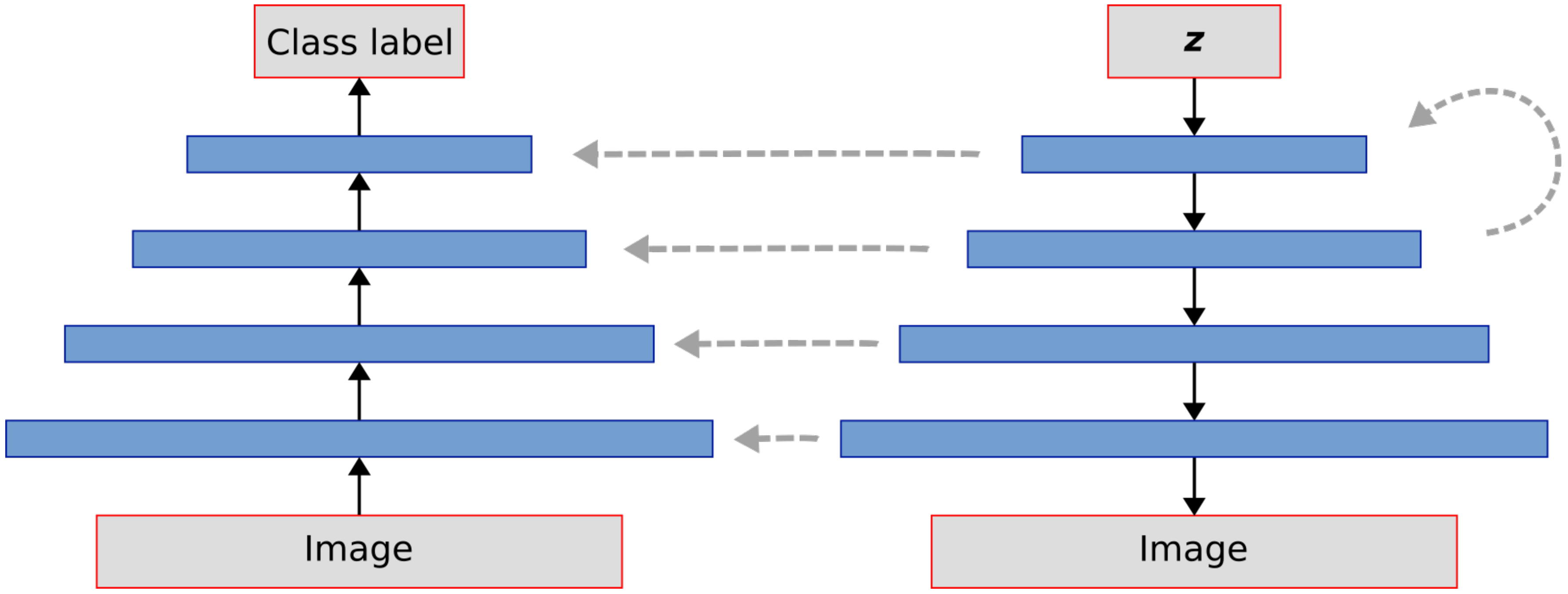
Brain-inspired modifications to Generative Replay:

- through feedback connections
- replay hidden representations



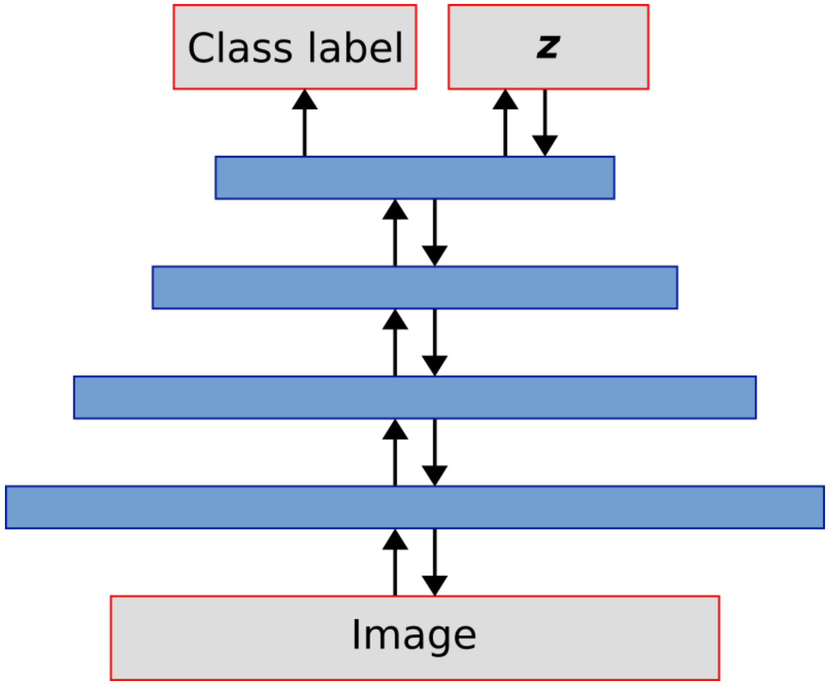
Brain-inspired modifications to Generative Replay:

- through feedback connections
- replay hidden representations



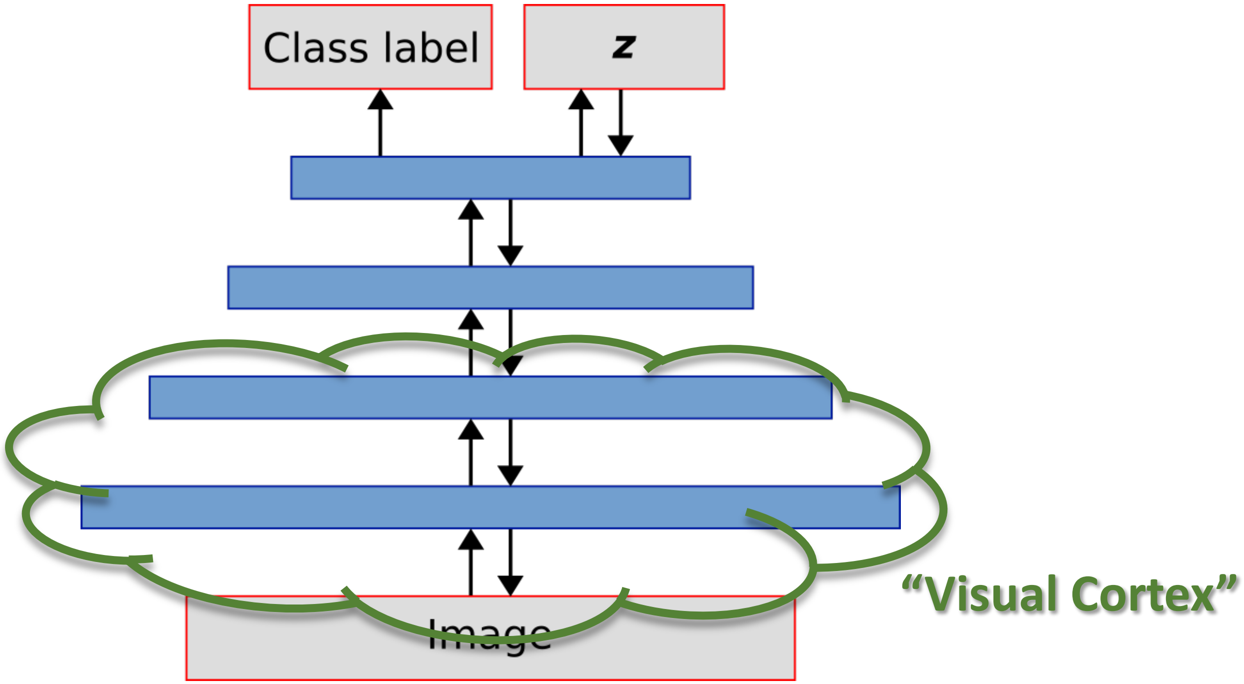
Brain-inspired modifications to Generative Replay:

- through feedback connections
- replay hidden representations



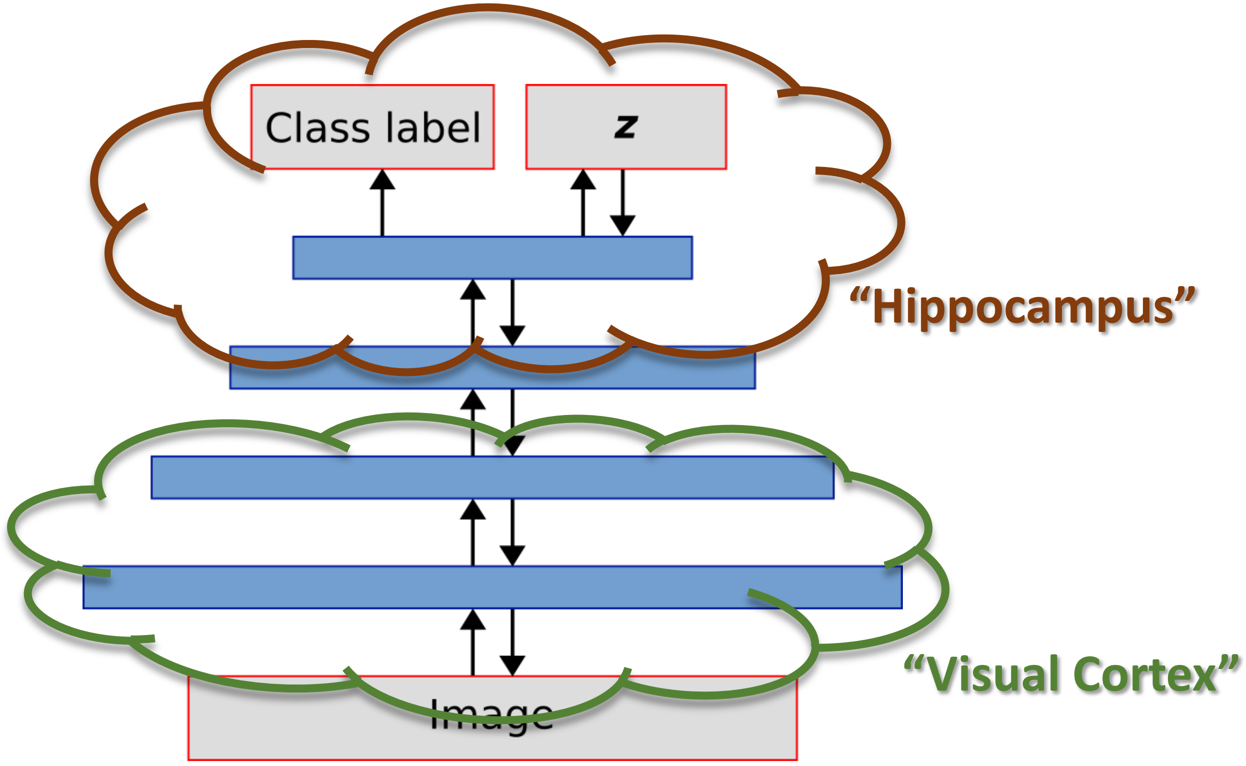
Brain-inspired modifications to Generative Replay:

- through feedback connections
- replay hidden representations



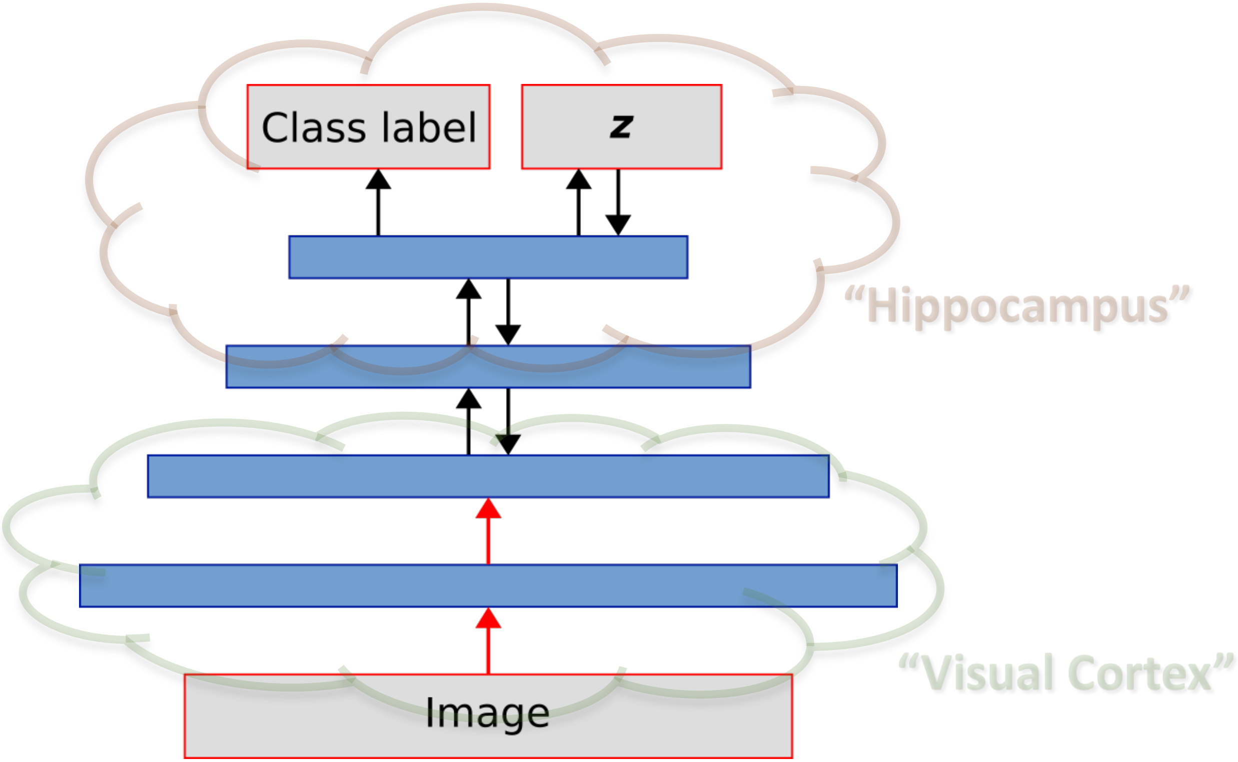
Brain-inspired modifications to Generative Replay:

- through feedback connections
- replay hidden representations

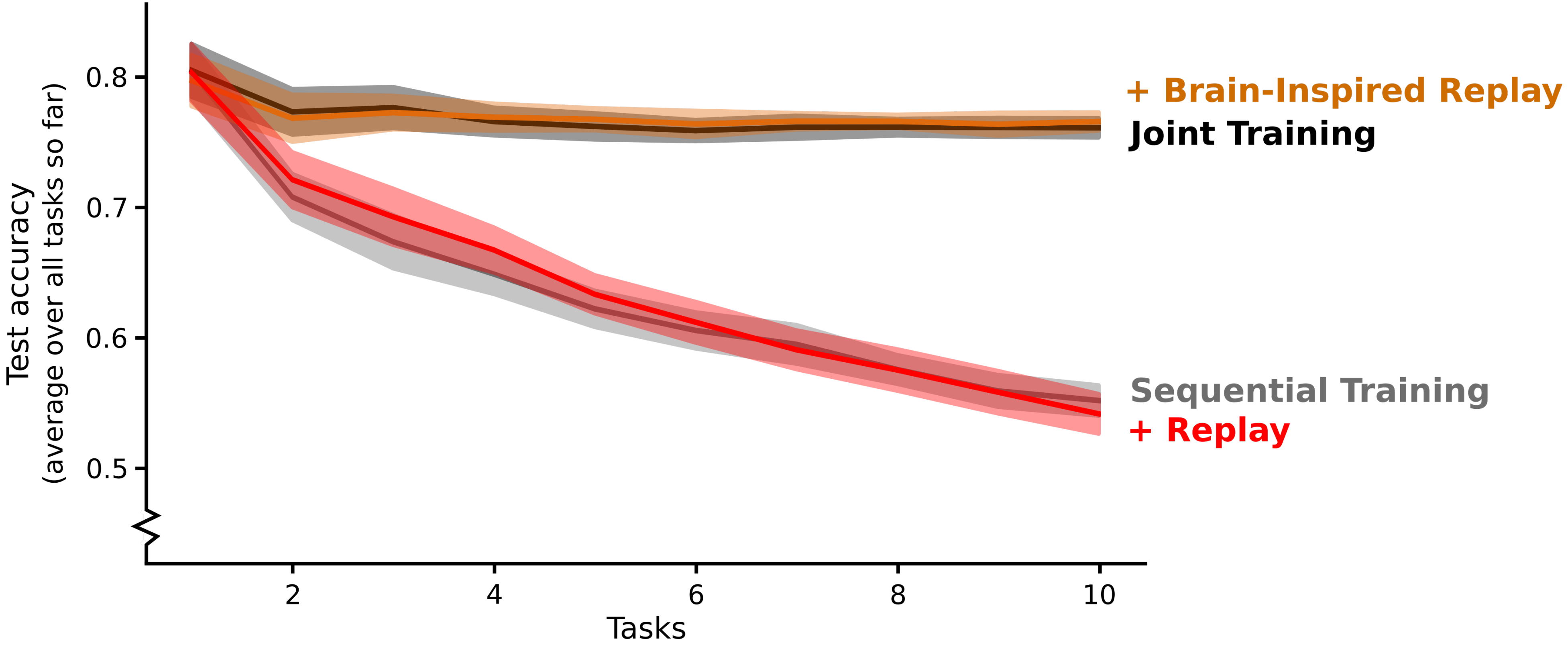


Brain-inspired modifications to Generative Replay:

- through feedback connections
- replay hidden representations



Brain-inspired modifications help Generative Replay scale



Summary **Part II**

- Generative Replay can successfully reduce Catastrophic Forgetting in artificial neural networks
- Modelling it after the brain helps to make this strategy scalable

Further details: see poster B1 tomorrow!

Contributors & Funding

University of Oxford
David Dupret
Stephanie Trouche
Colin McNamara
Kevin Allen

Baylor College of Medicine, Houston
Andreas Tolias

UNIVERSITY OF OXFORD
MRC Medical Research Council
IBRO
DARPA
IARPA