



university of
groningen

Forgetting in Brains and Machines

Gido M. van de Ven

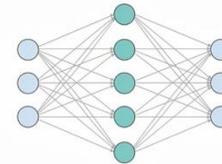
March 2026

Overview

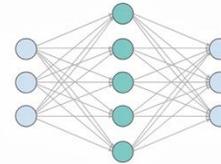
- Replay's role in memory consolidation



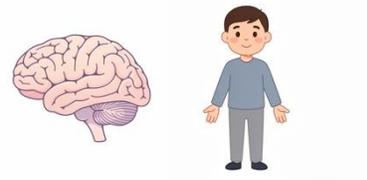
- Replay to prevent catastrophic forgetting



- Temporary forgetting in artificial neural networks



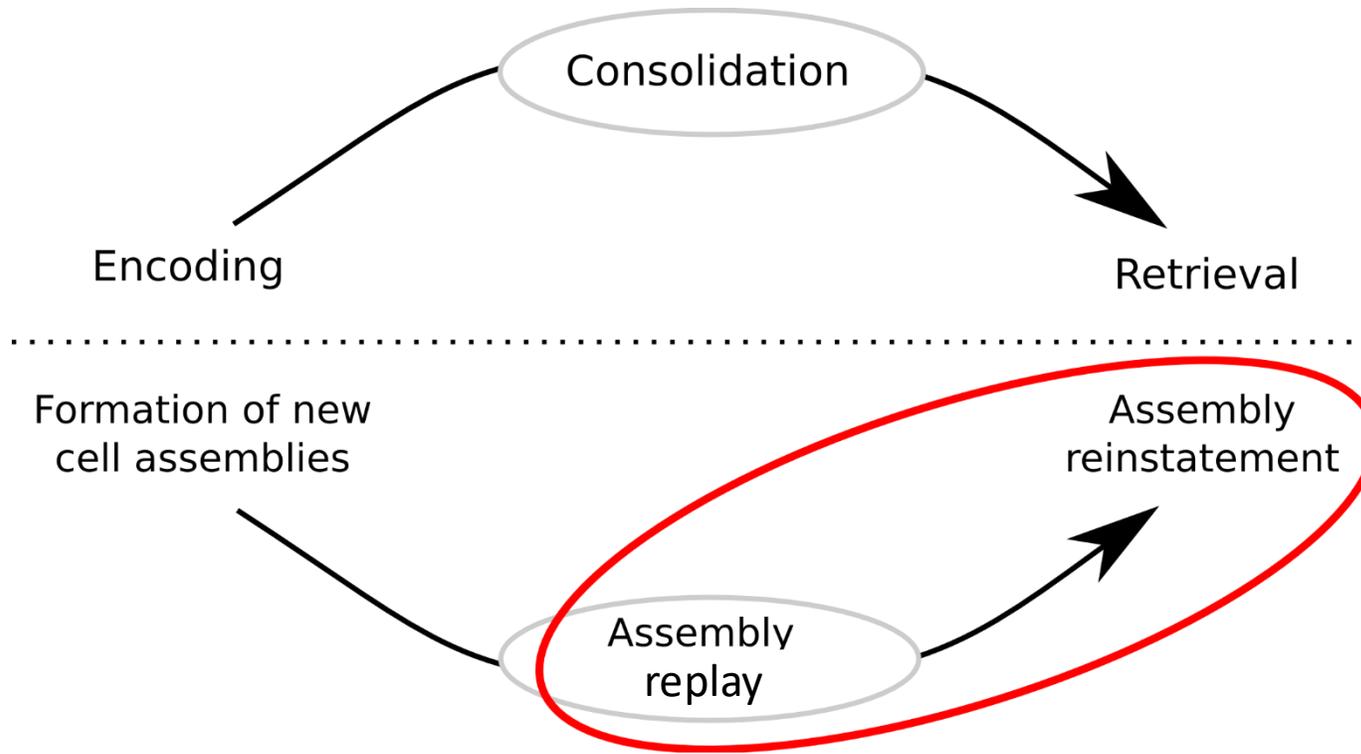
- Temporary forgetting in humans?



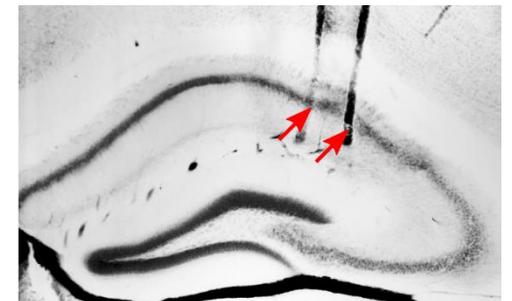
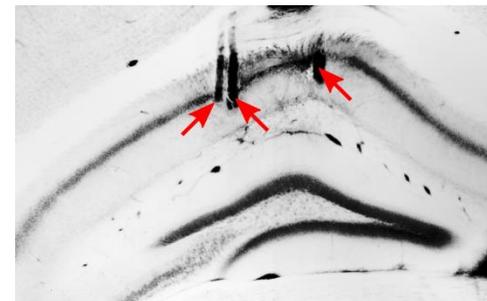
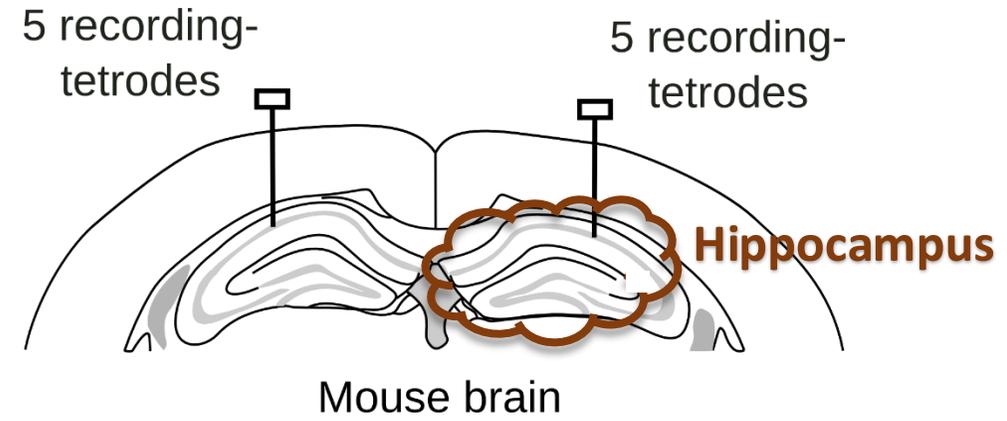
Memory consolidation: cell assembly / replay hypothesis



Prof. David Dupret



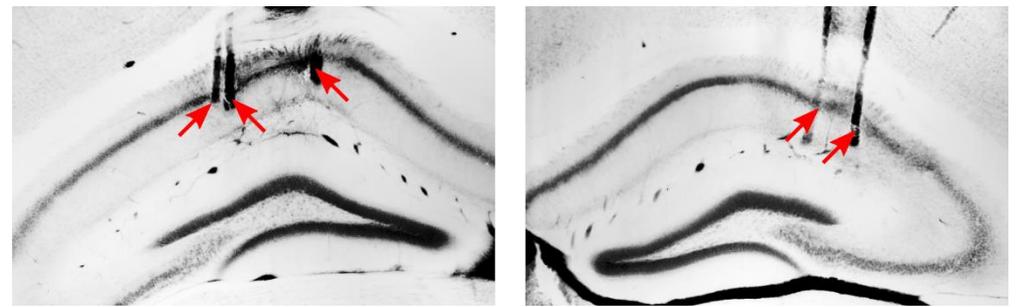
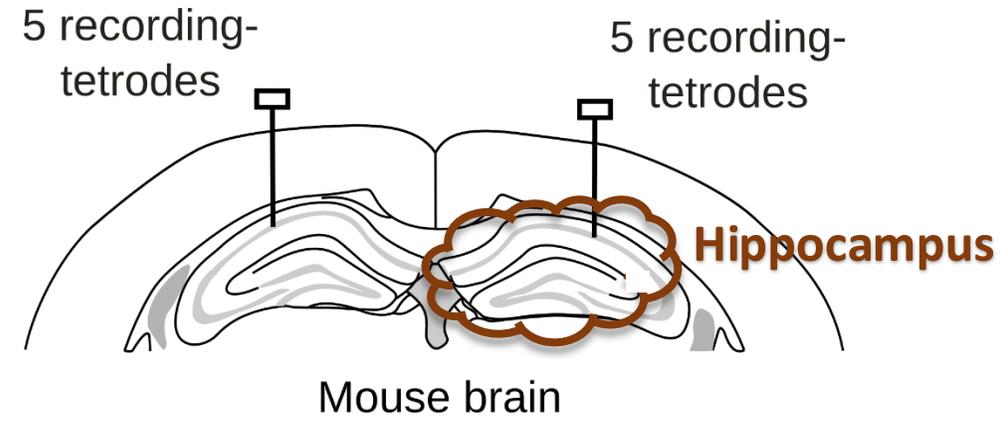
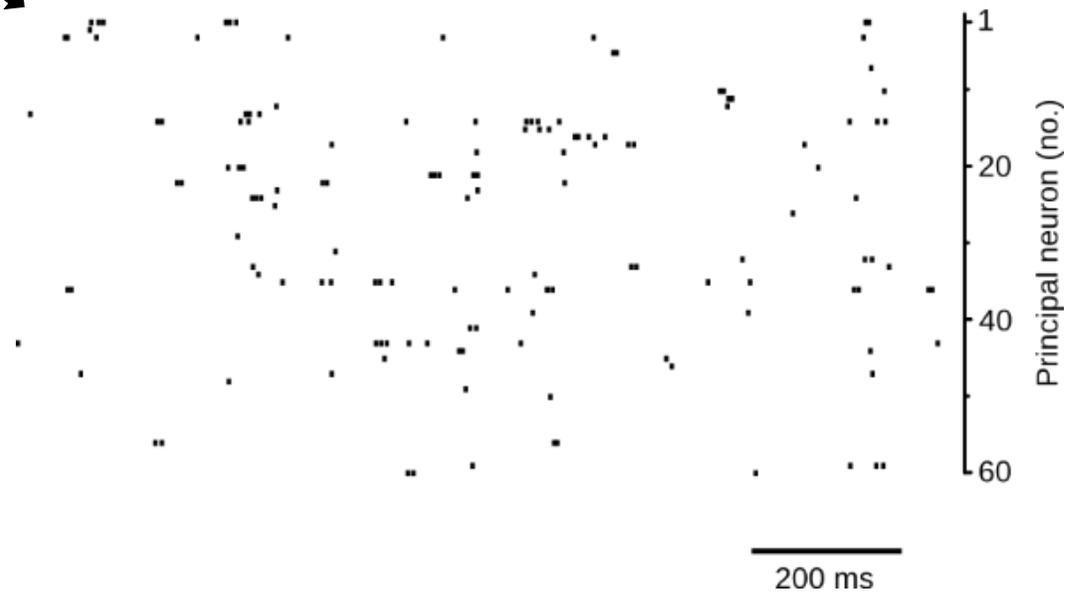
Identification of cell assembly patterns



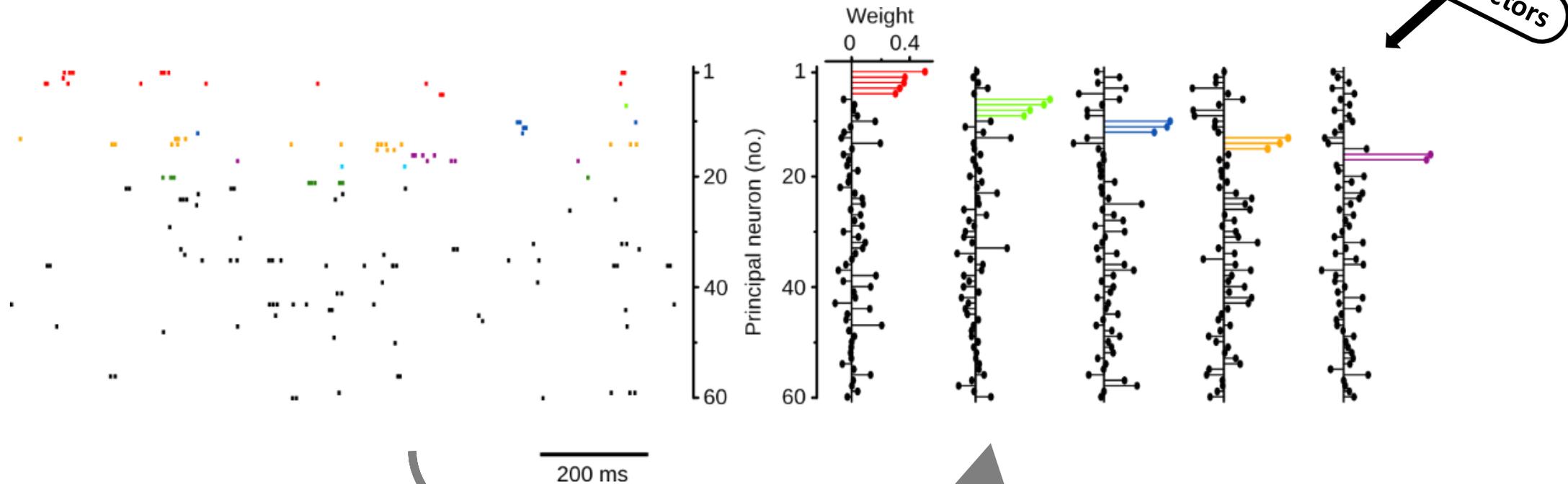
Identification of cell assembly patterns



spike trains

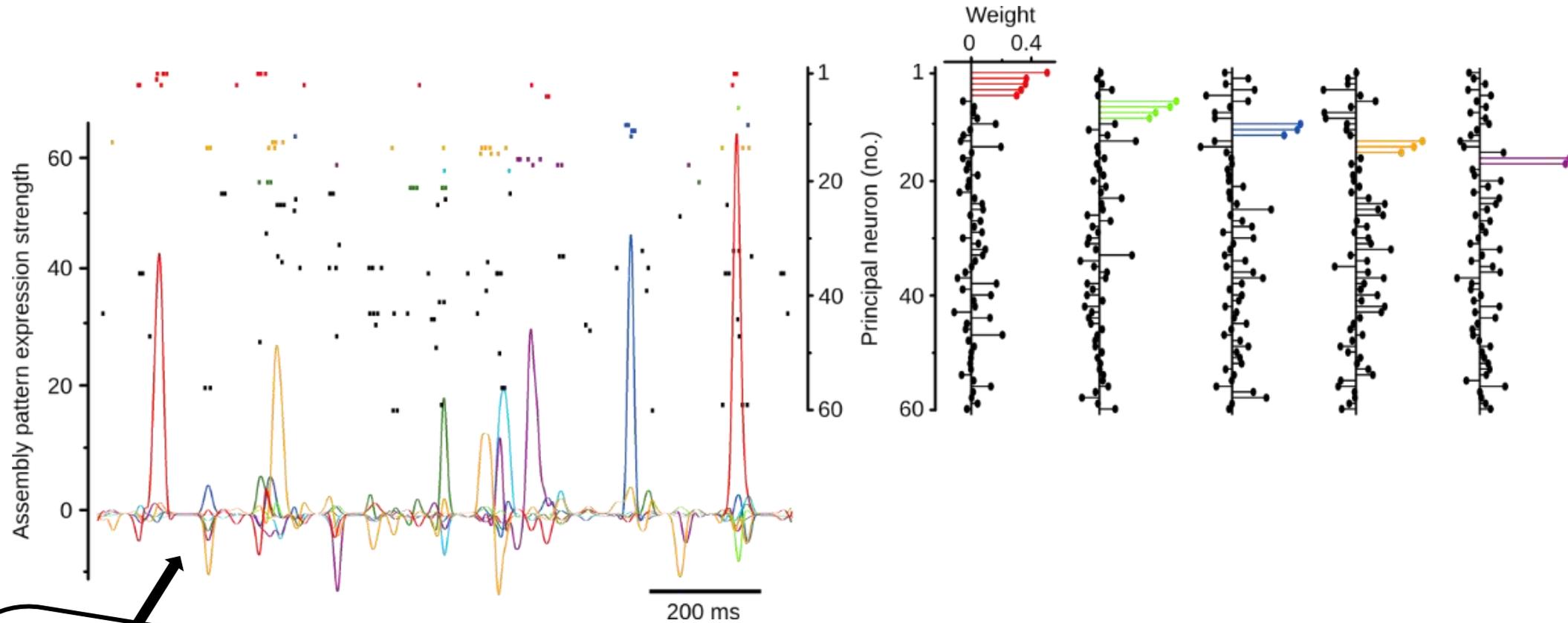


Identification of cell assembly patterns



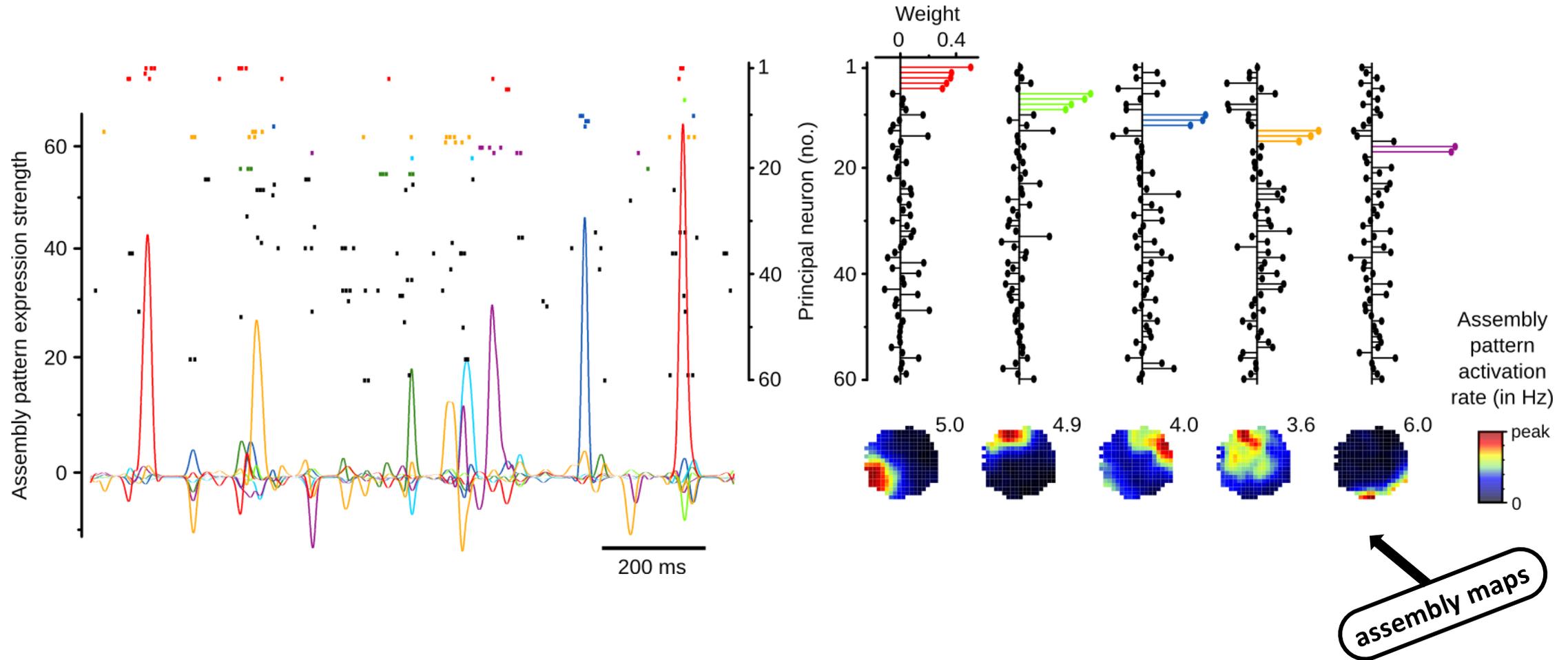
*Assembly-detection method
based on PCA and ICA*

Identification of cell assembly patterns

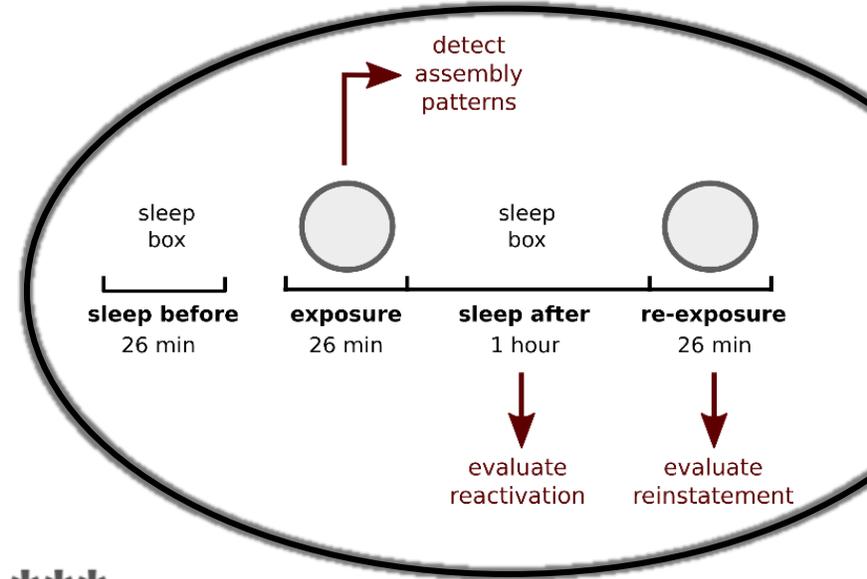
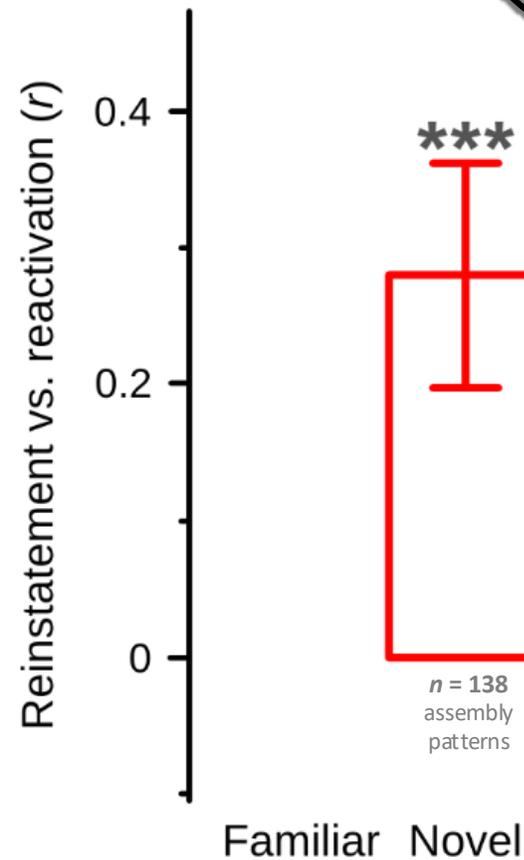
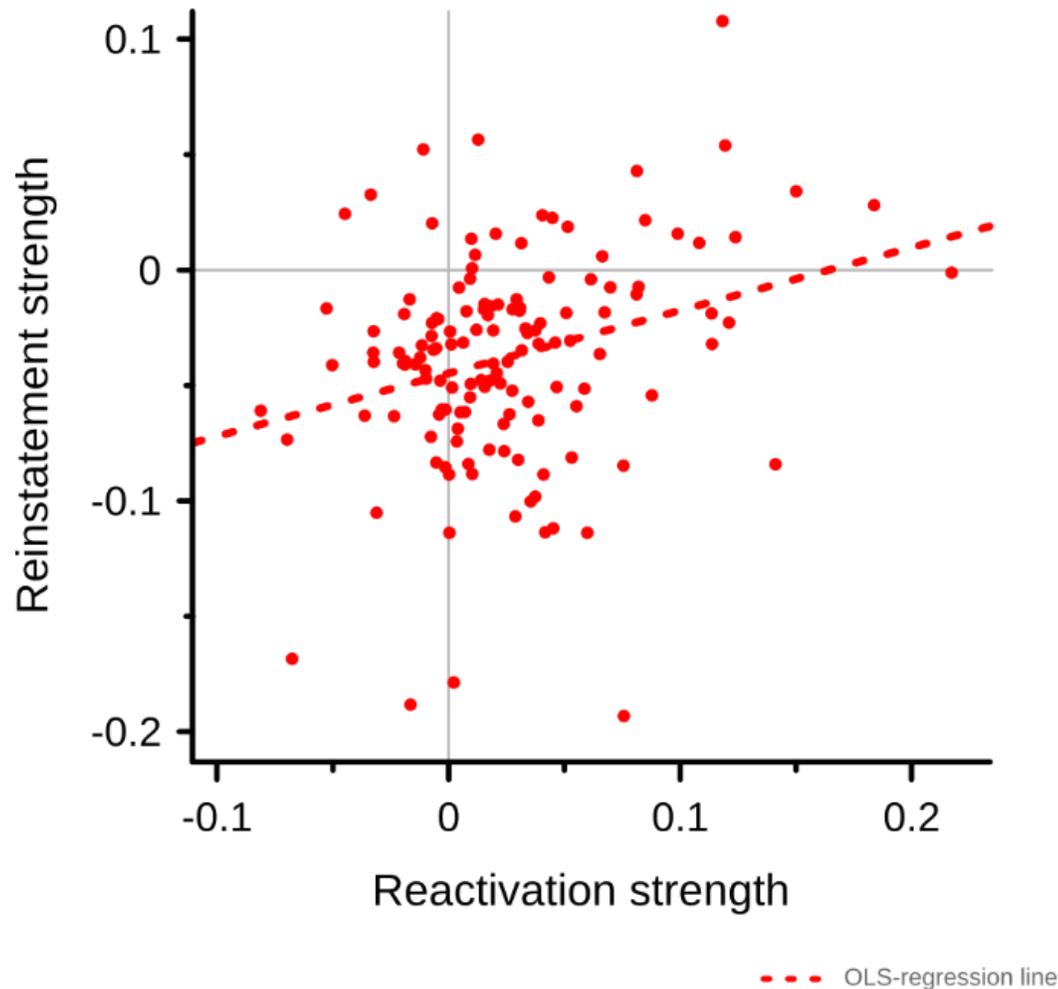


assembly pattern expression-strength tracked over time

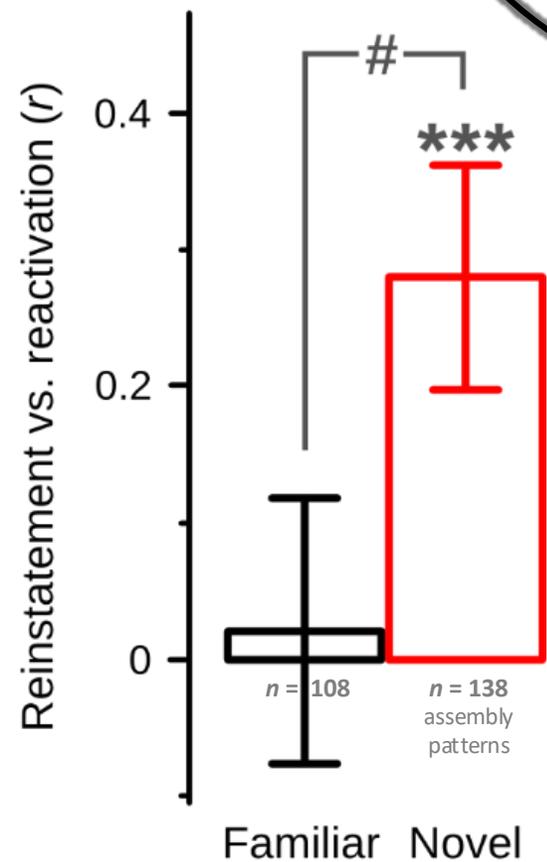
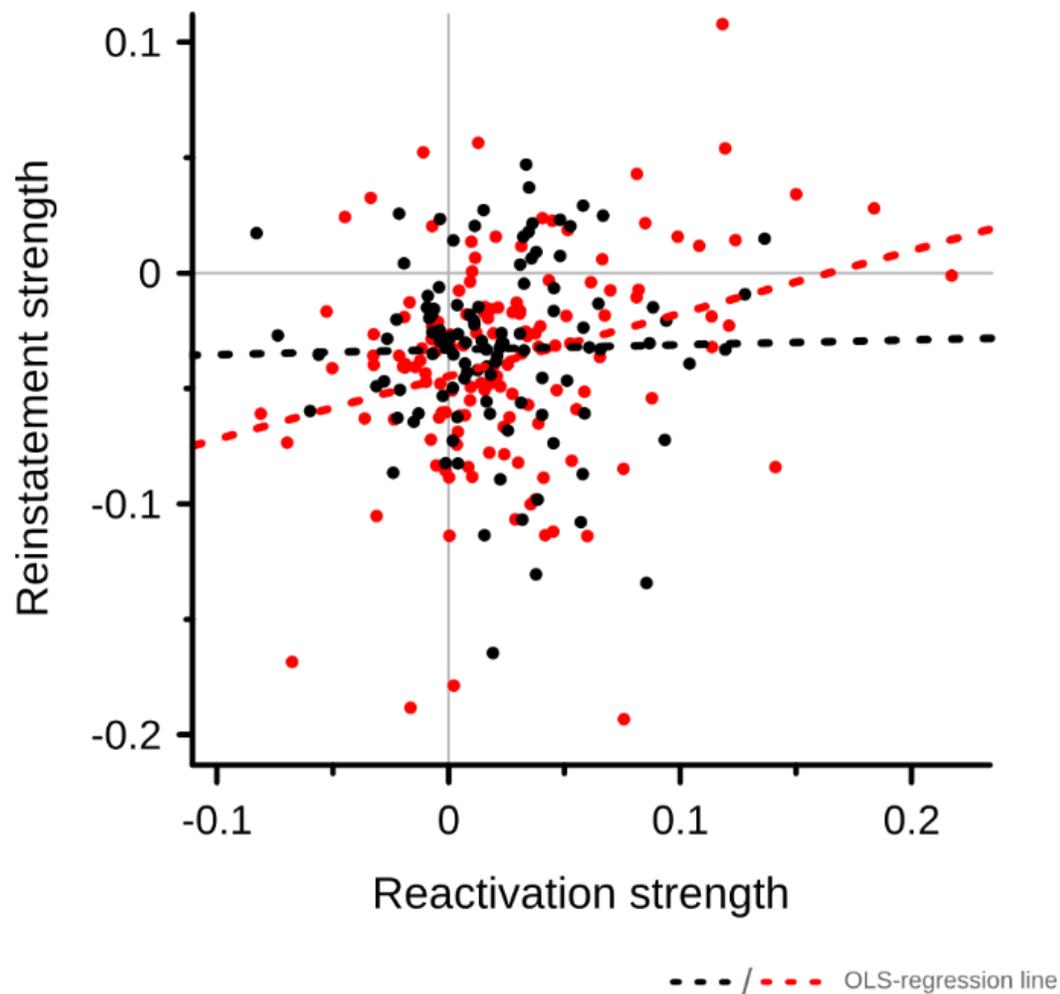
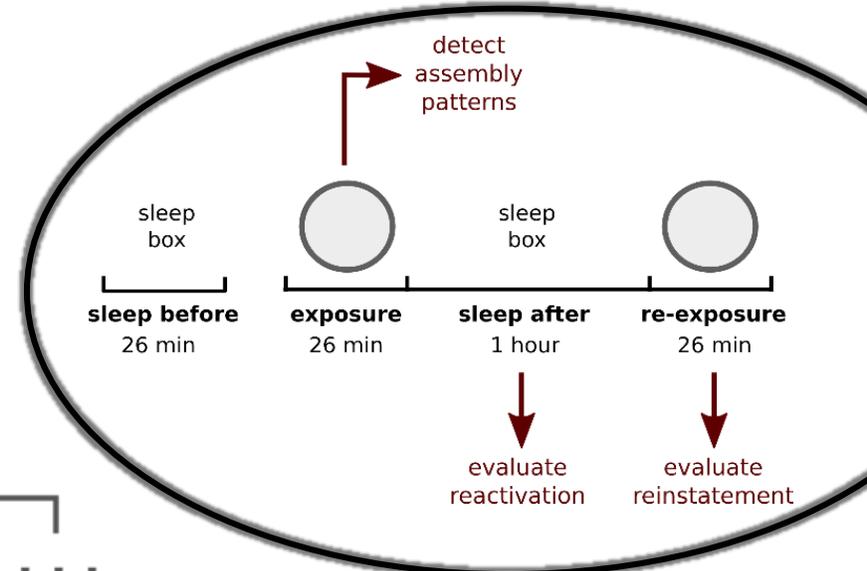
Identification of cell assembly patterns



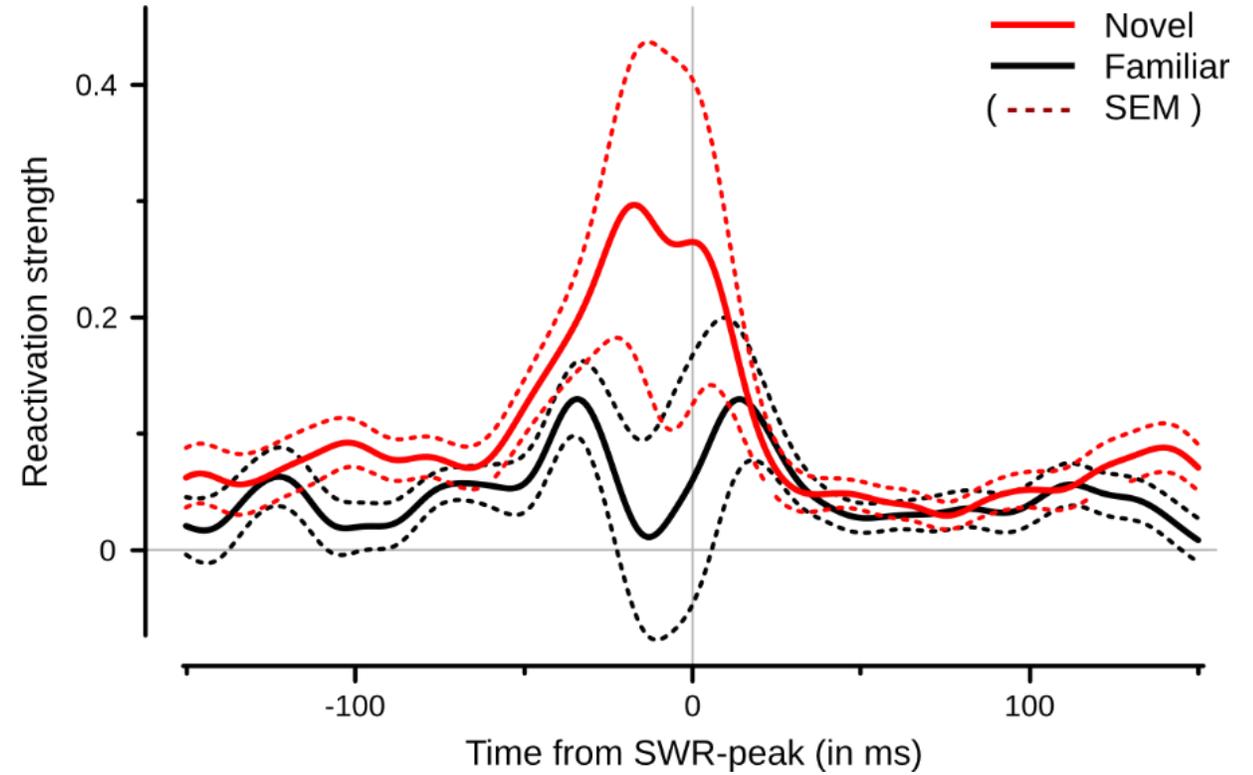
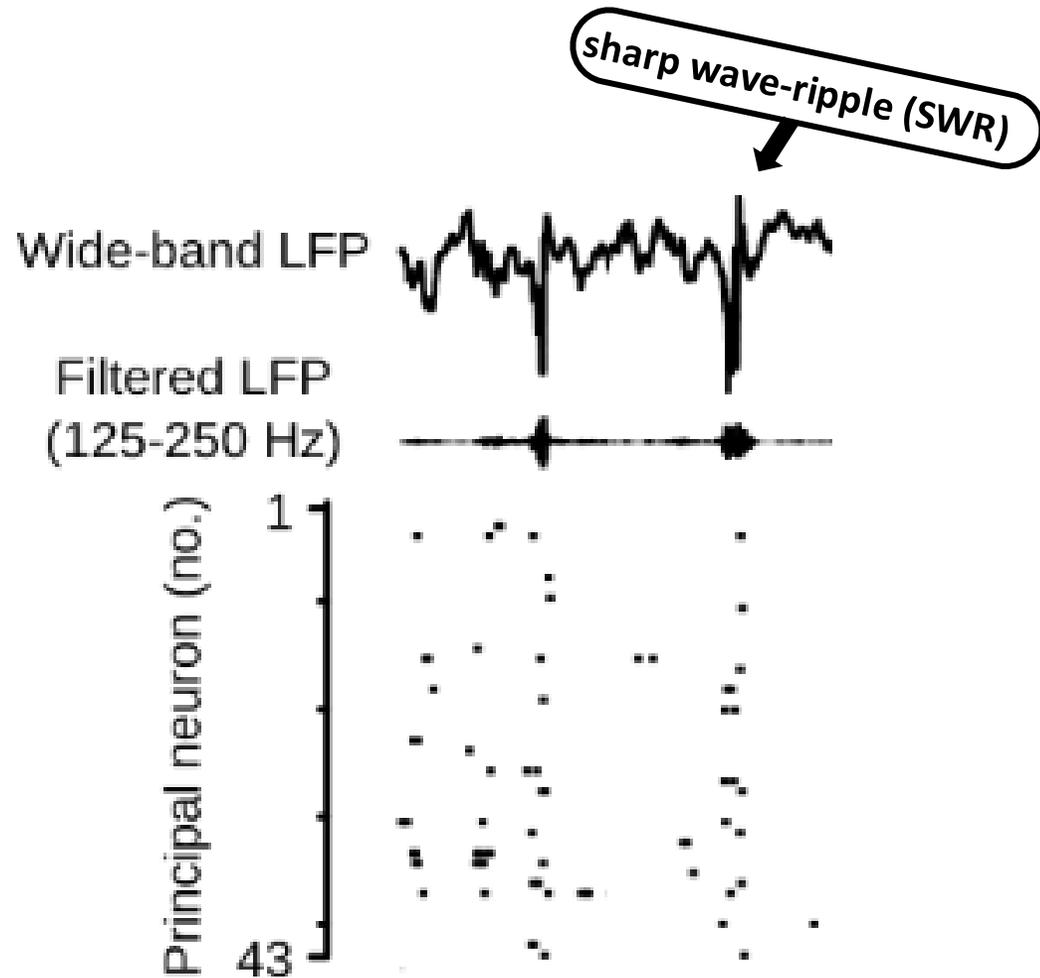
An assembly pattern's replay predicts it subsequent reinstatement



An assembly pattern's replay predicts it subsequent reinstatement

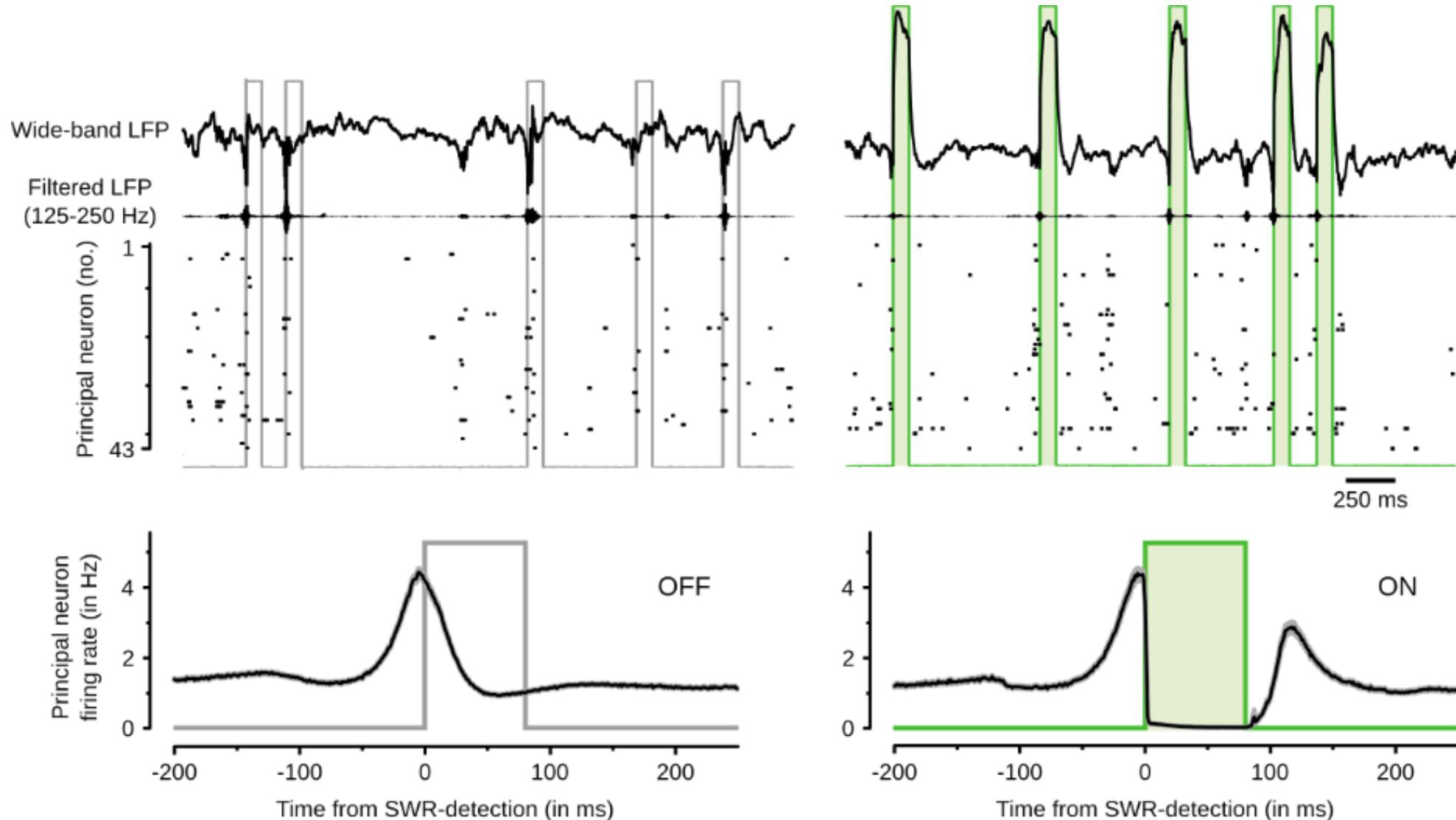


Selective disruption of replay?



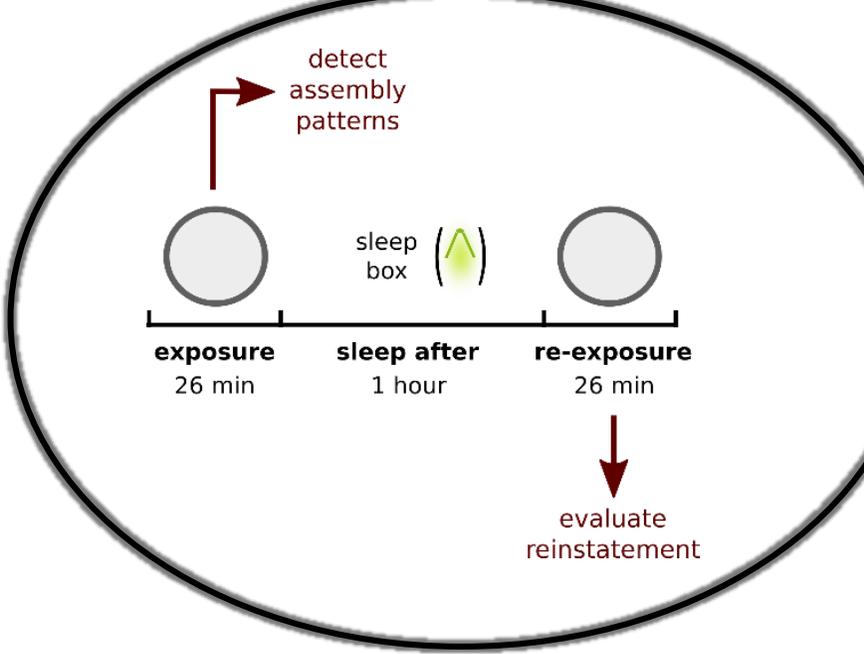
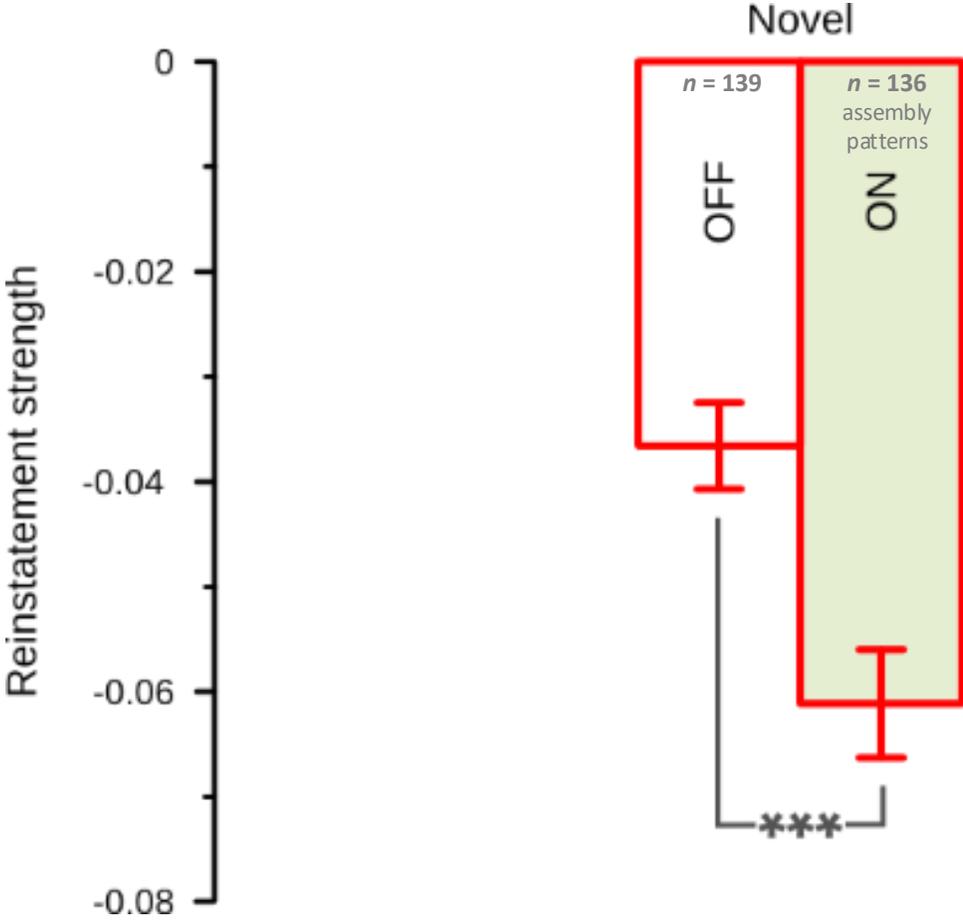
Novel: $n = 139$ assembly-patterns
Familiar: $n = 108$ assembly-patterns
(based on 43 recording-blocks from 8 mice)

Selective disruption of replay: *optogenetic SWR silencing*



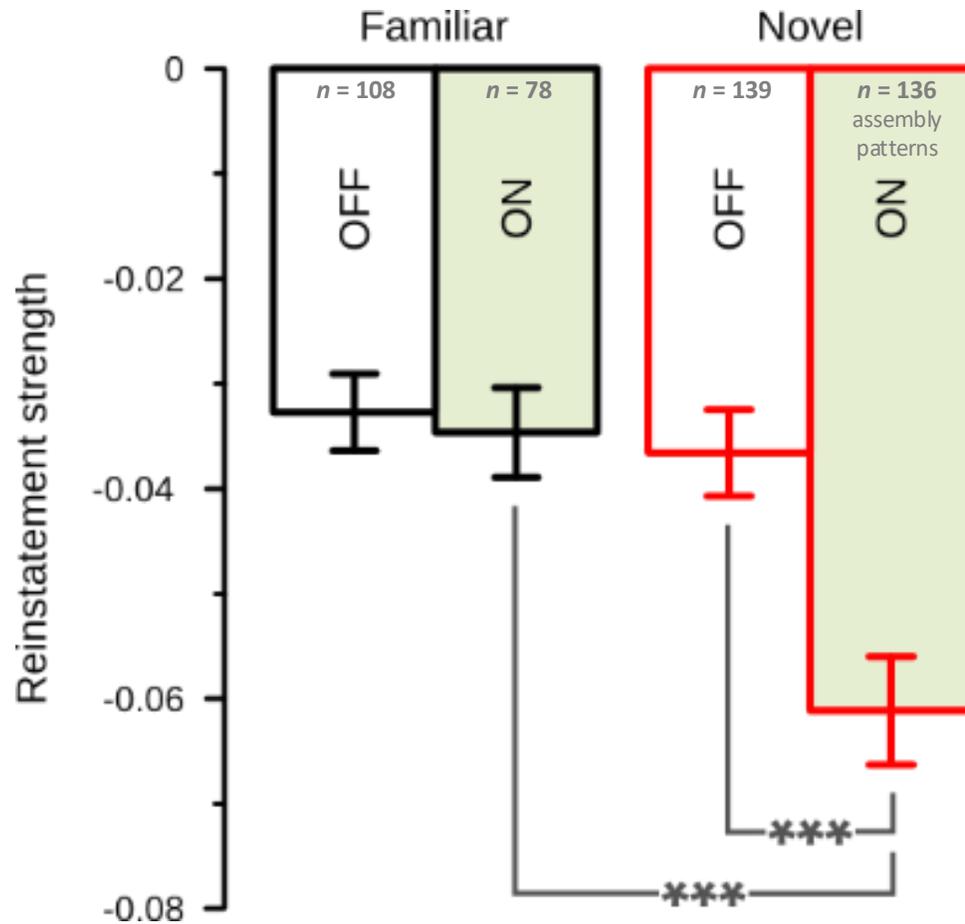
OFF: $n = 1,988$ neurons (from 43 sessions)
ON: $n = 1,527$ neurons (from 37 sessions)

SWR-silencing impairs assembly pattern reinstatement

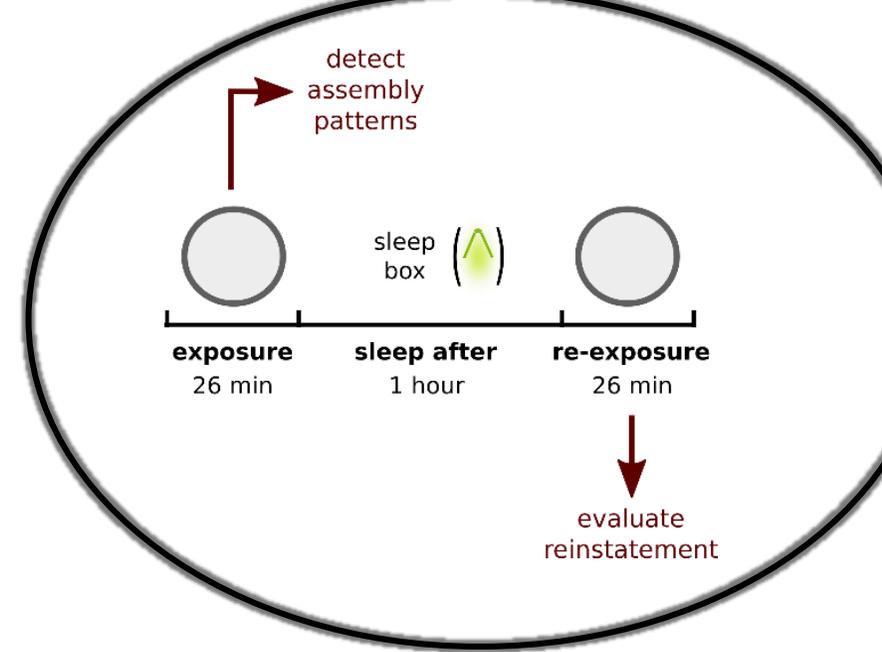


(based on 50 recording-blocks from 8 mice)

SWR-silencing impairs assembly pattern reinstatement



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





Summary Part I

- In the brain, replay stabilizes newly-formed, memory-representing cell assembly patterns

Further details: van de Ven *et al.* (2016) *Neuron* **92**, pp. 968-974 [+ video abstract]



Summary Part I

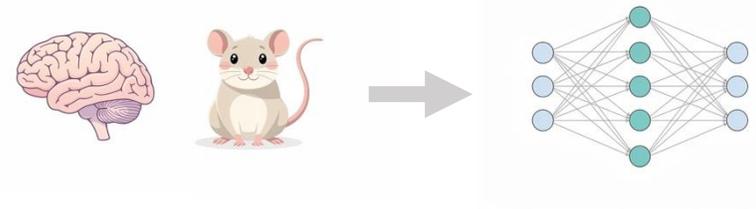
- In the brain, replay stabilizes newly-formed, memory-representing cell assembly patterns

Further details: van de Ven *et al.* (2016) *Neuron* 92, pp. 968-974 [+ video abstract]

But ...

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

Summary Part I



- In the brain, replay stabilizes newly-formed, memory-representing cell assembly patterns

Further details: van de Ven *et al.* (2016) *Neuron* 92, pp. 968-974 [+ video abstract]

But ...

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

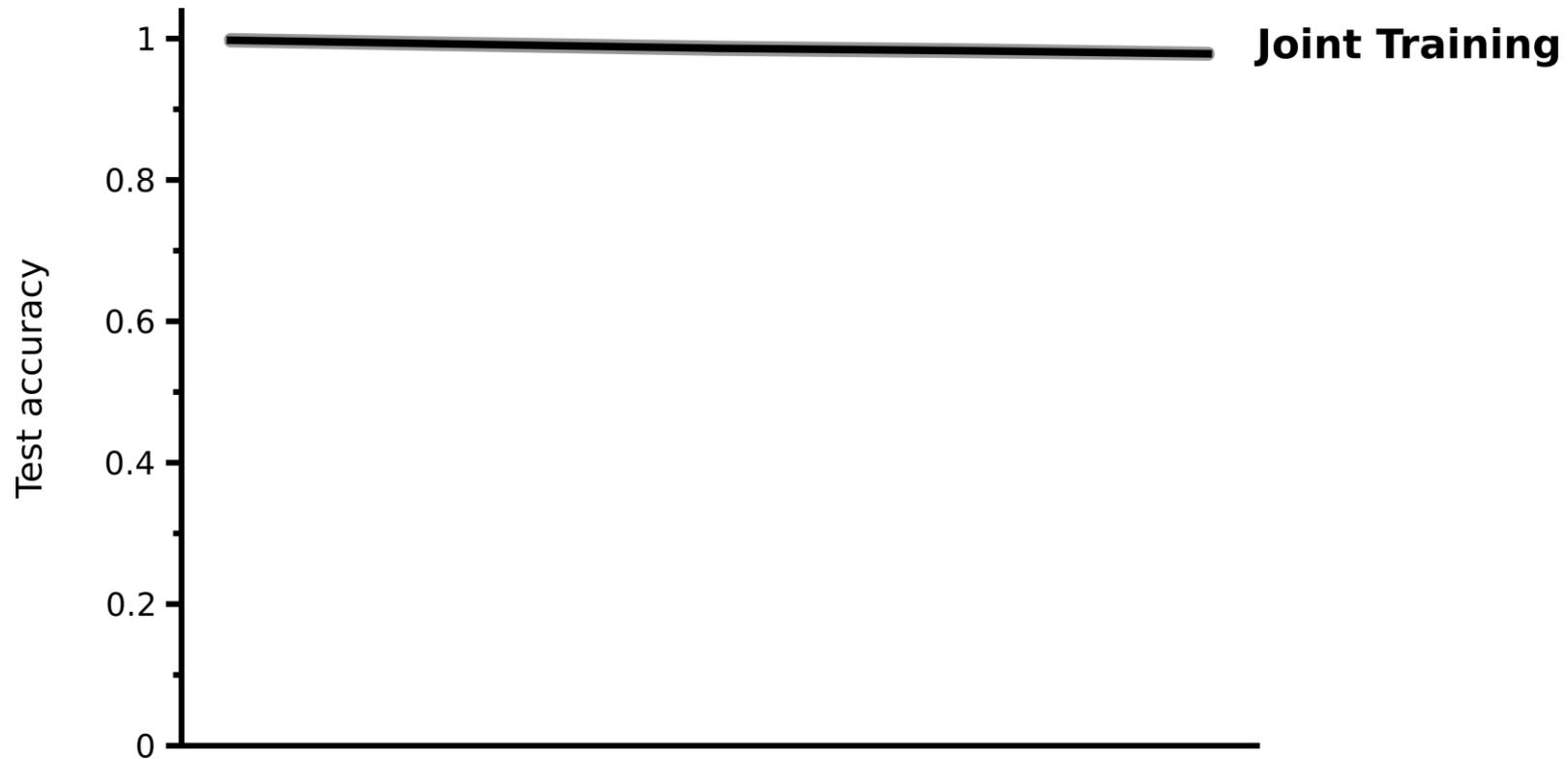
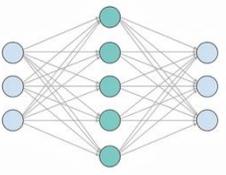
Approach

- Artificial neural networks as “model organism”

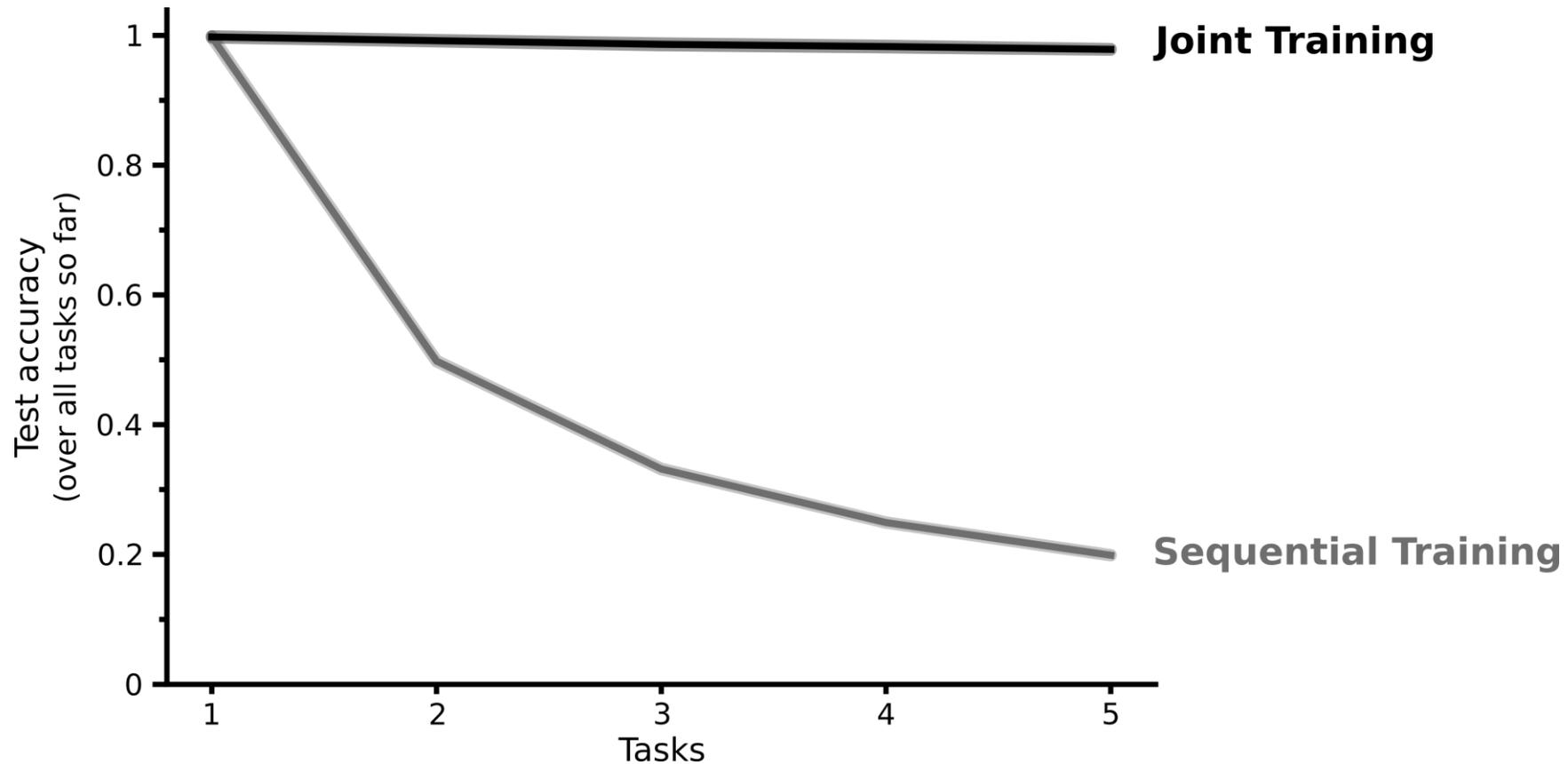
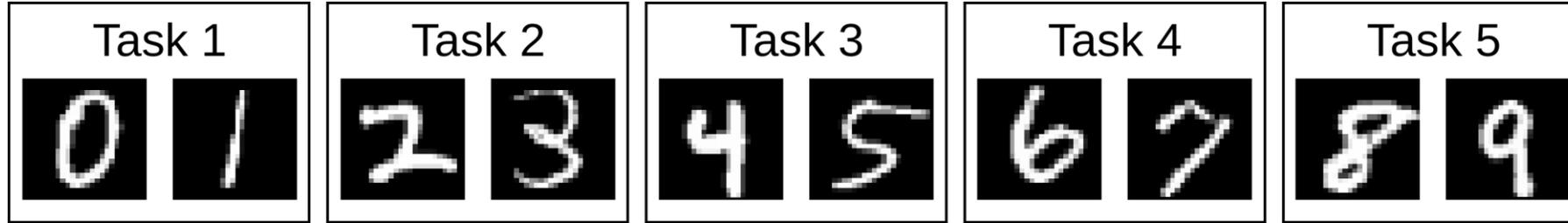
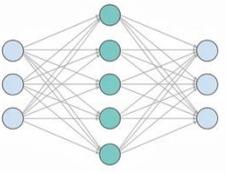


Prof. Andreas Tolias

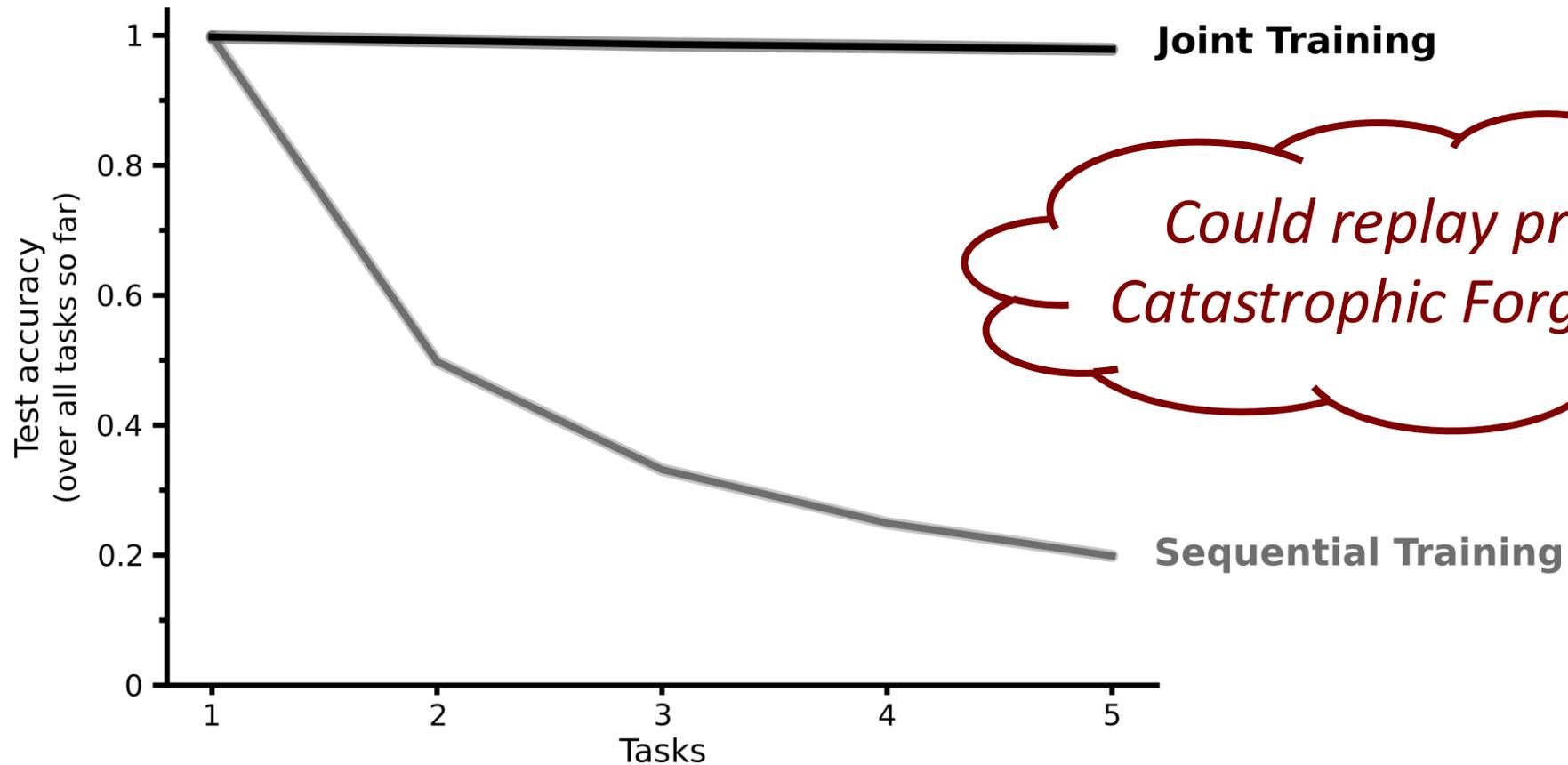
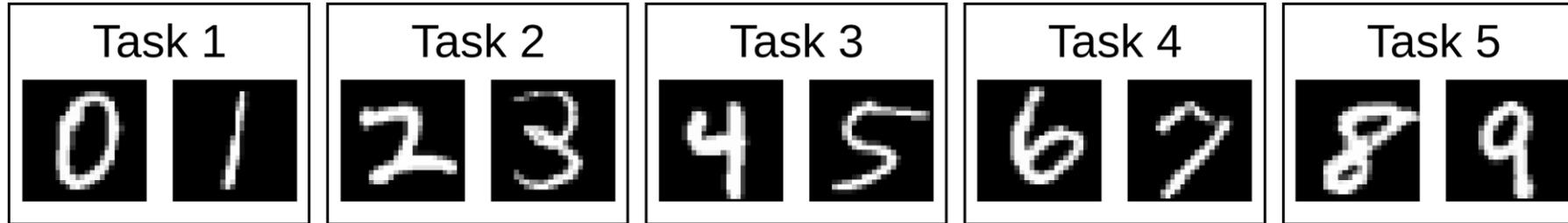
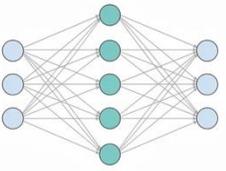
Catastrophic Forgetting in Artificial Neural Networks



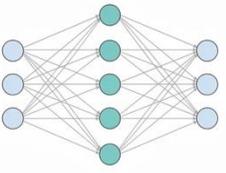
Catastrophic Forgetting in Artificial Neural Networks



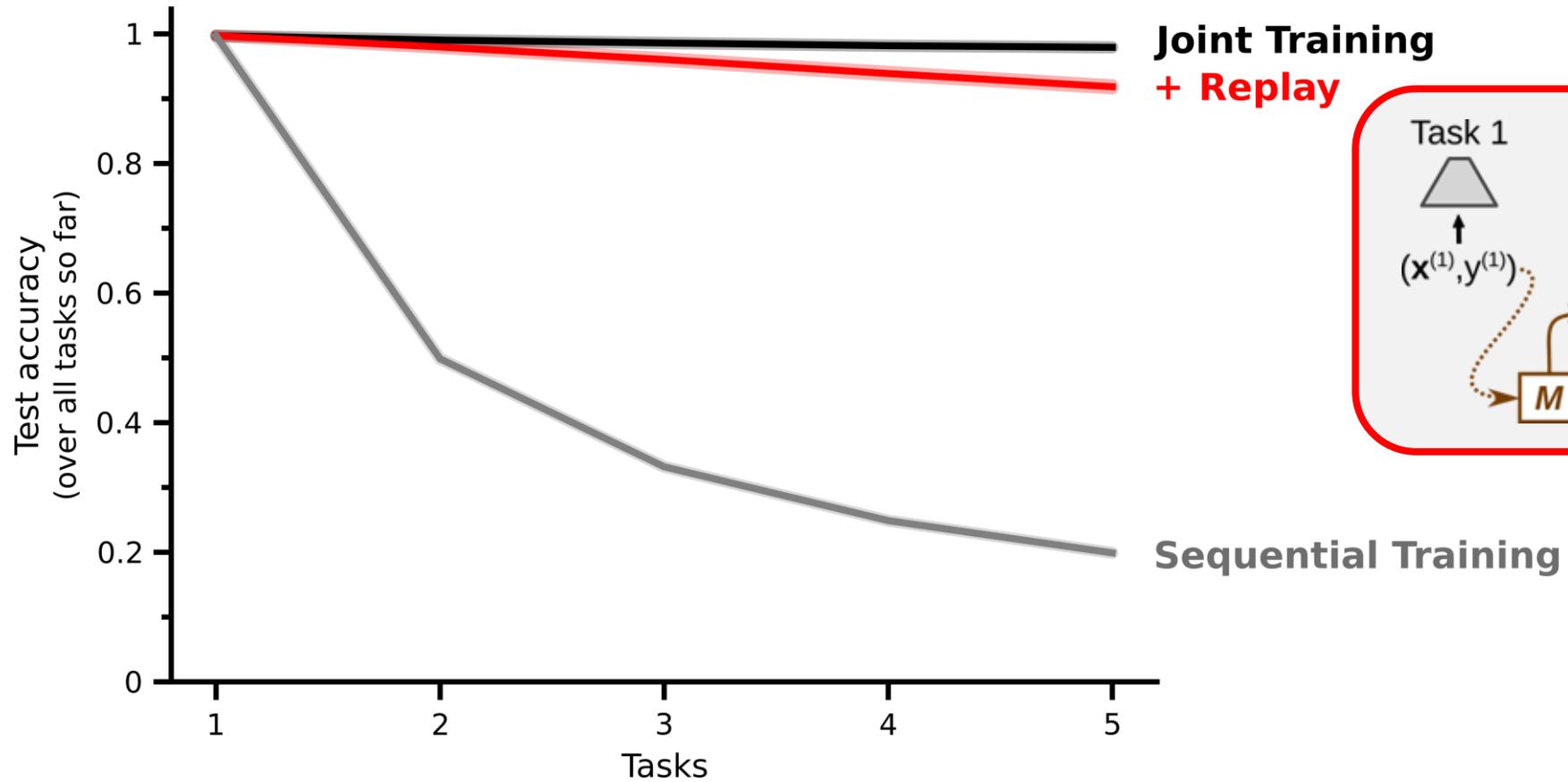
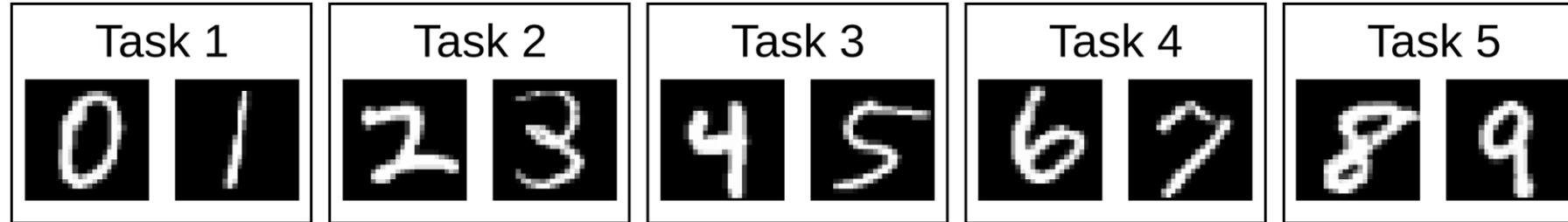
Catastrophic Forgetting in Artificial Neural Networks



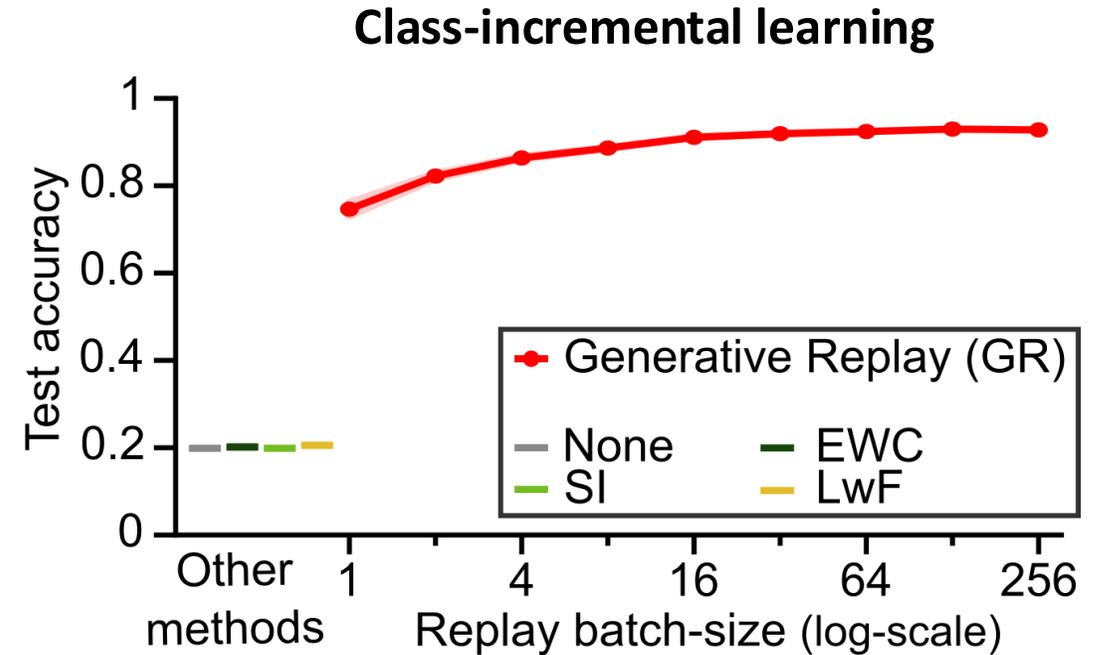
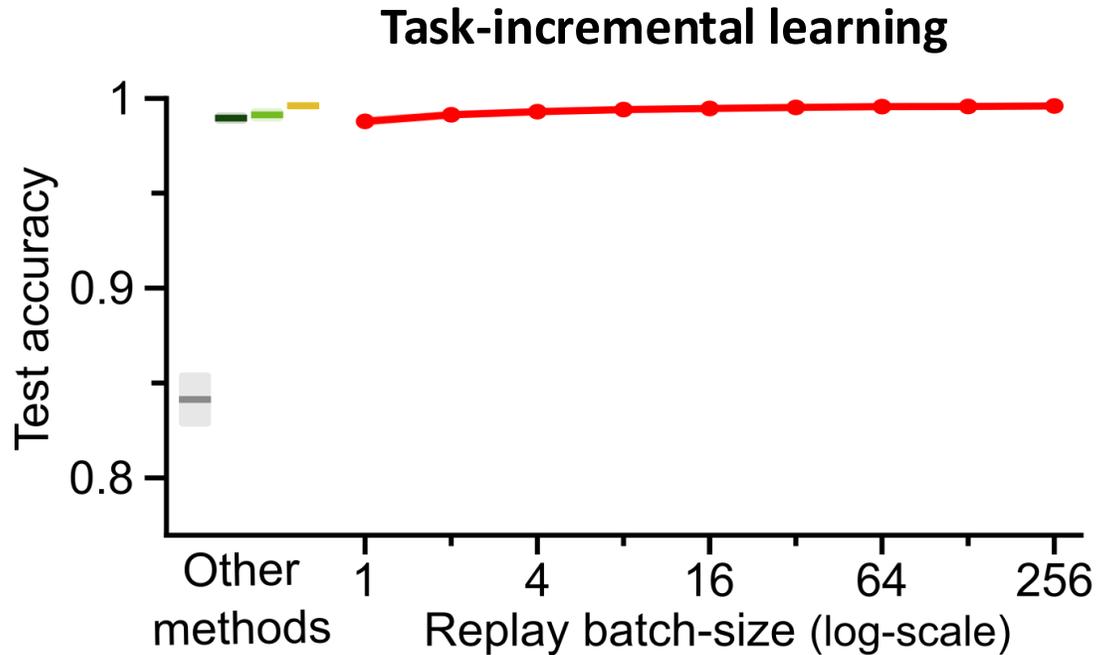
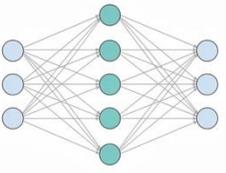
*Could replay prevent
Catastrophic Forgetting?*



Replay can prevent Catastrophic Forgetting

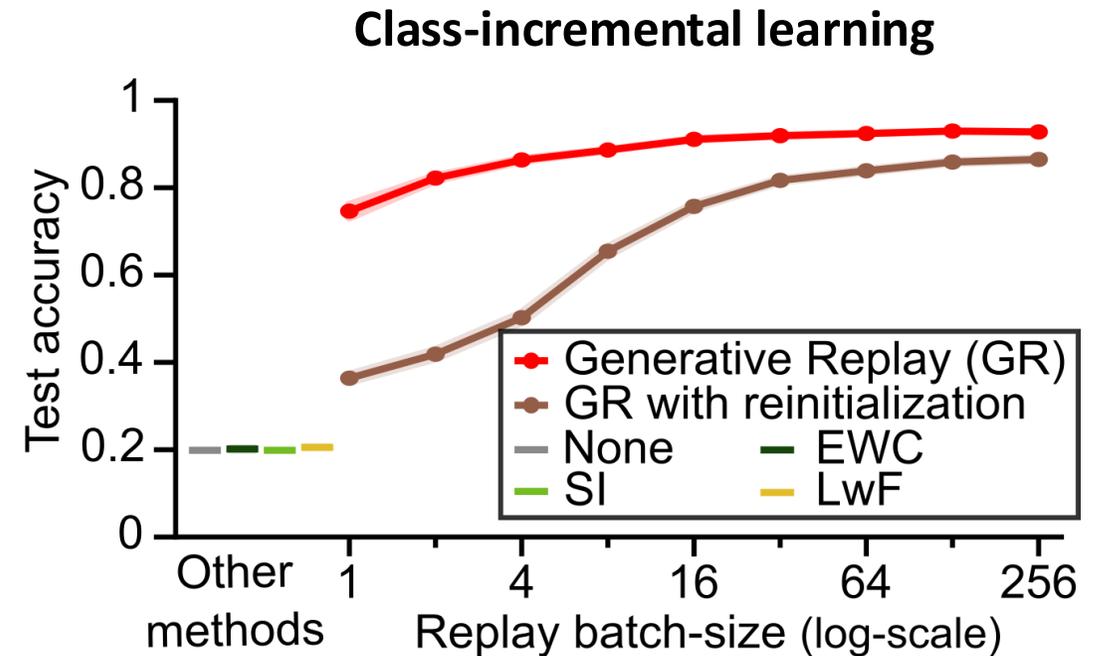
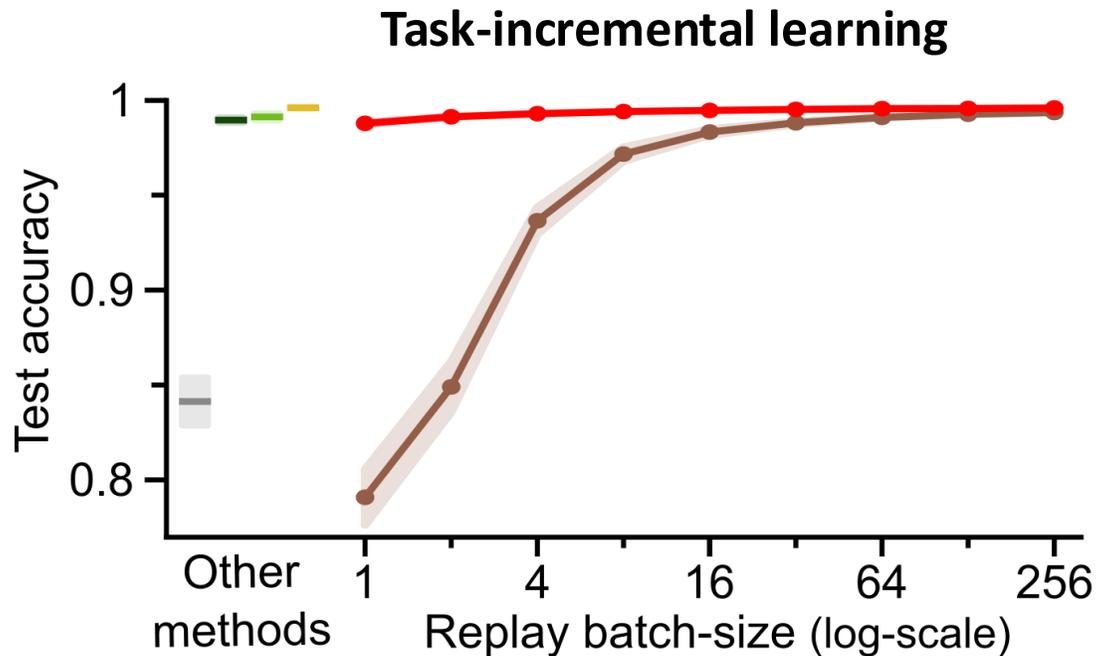
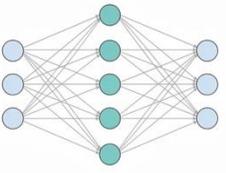


Efficiency: How much replay is needed?



- Fully replaying past tasks is not needed, replaying just a few samples could suffice

Efficiency: How much replay is needed?



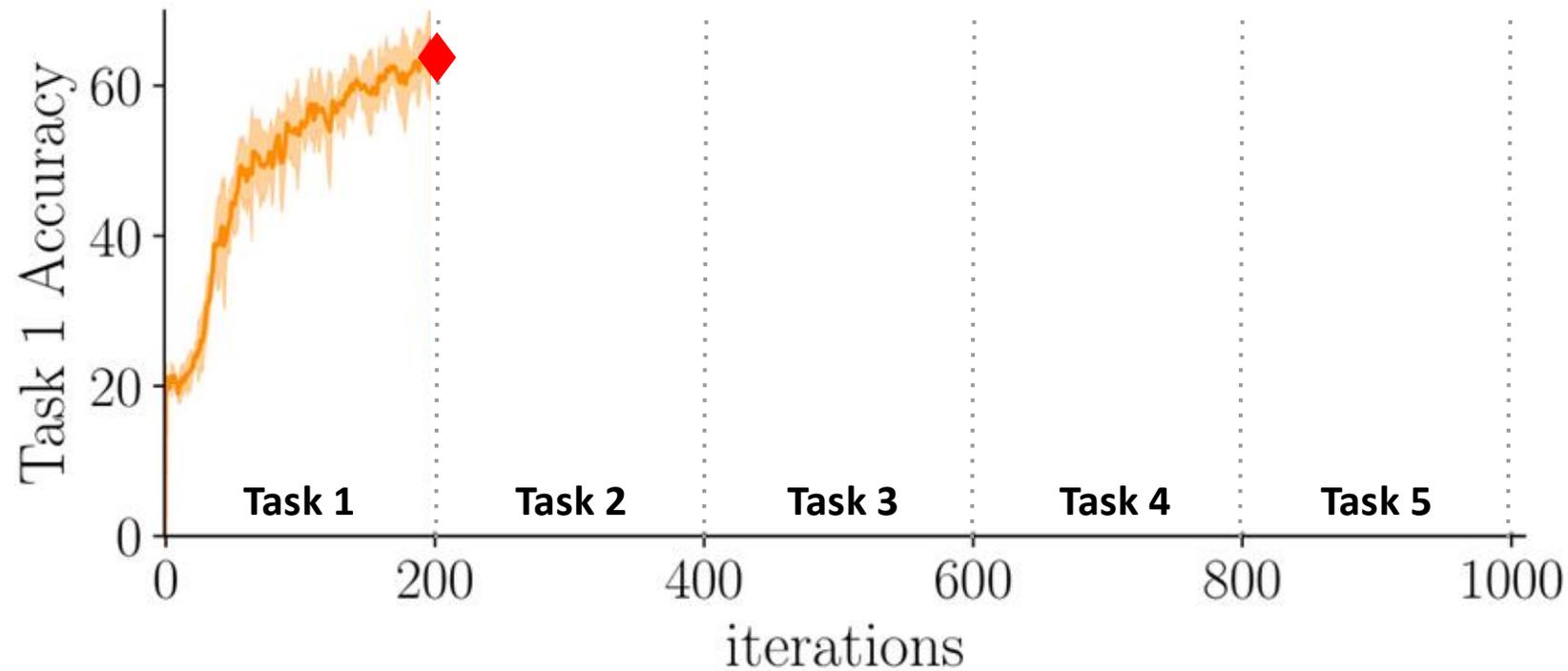
- Fully replaying past tasks is not needed, replaying just a few samples could suffice
- Because “learning something” is harder than “not forgetting it”

Does replay prevent forgetting?

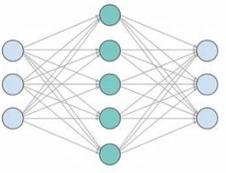


Prof. Tinne Tuytelaars Matthias De Lange

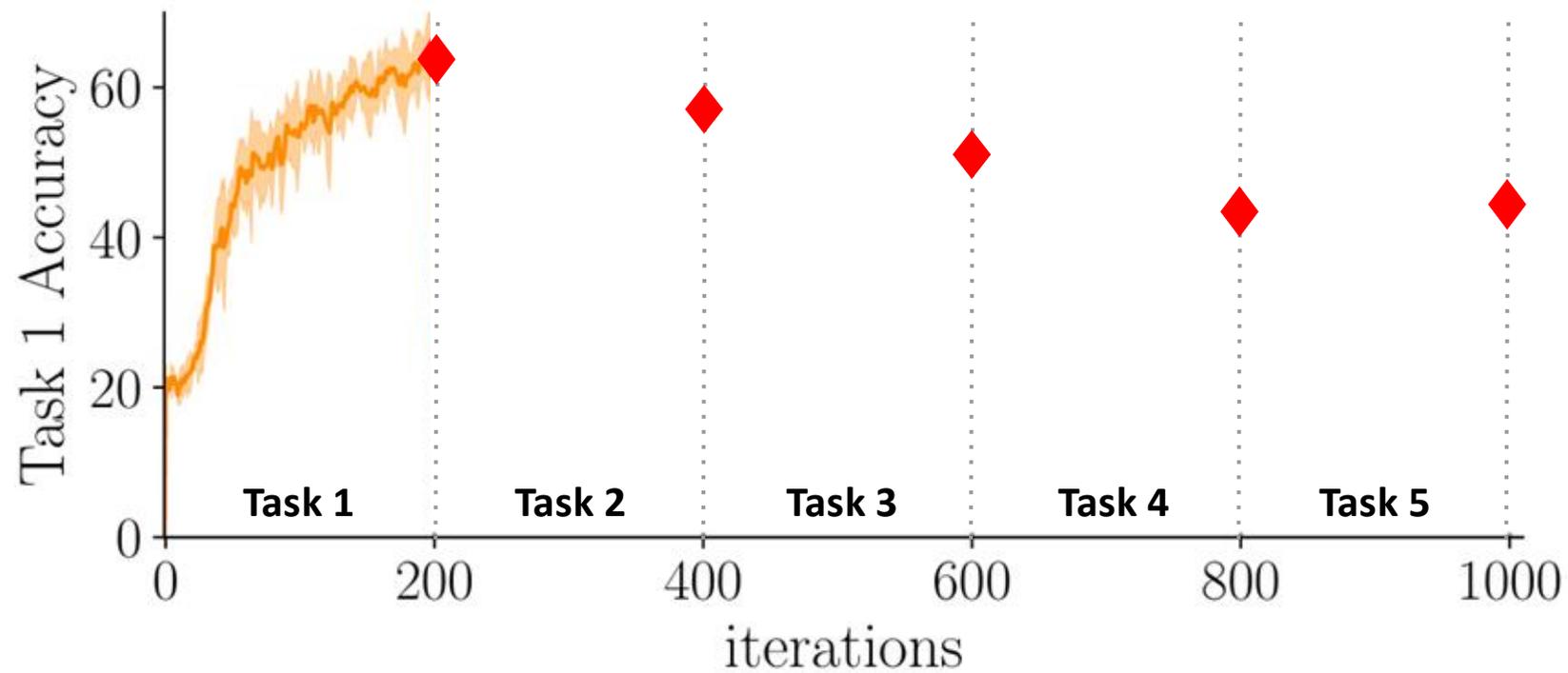
Replay, Class-incremental Split MinilmageNet



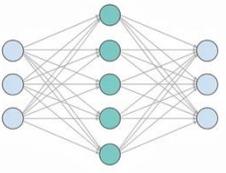
Does replay prevent forgetting?



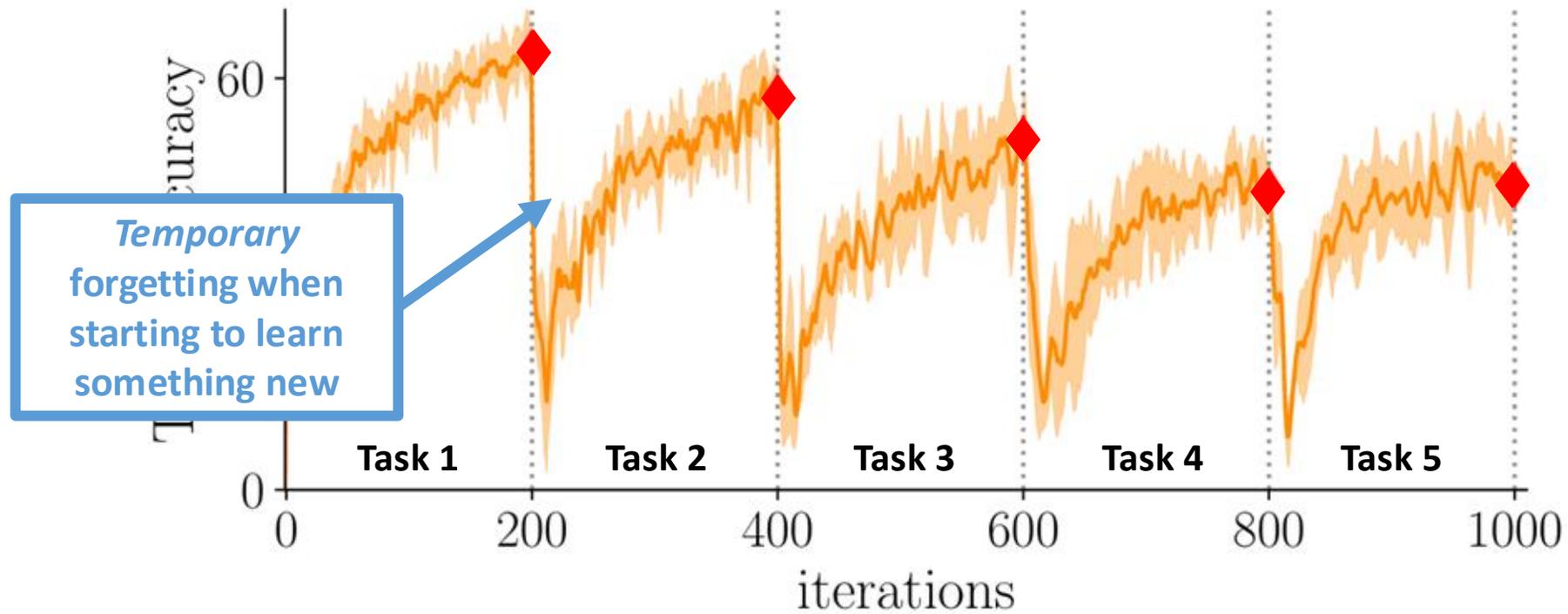
Replay, Class-incremental Split MinImageNet



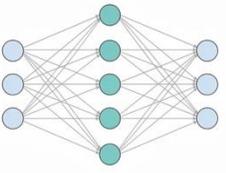
Does replay prevent forgetting?



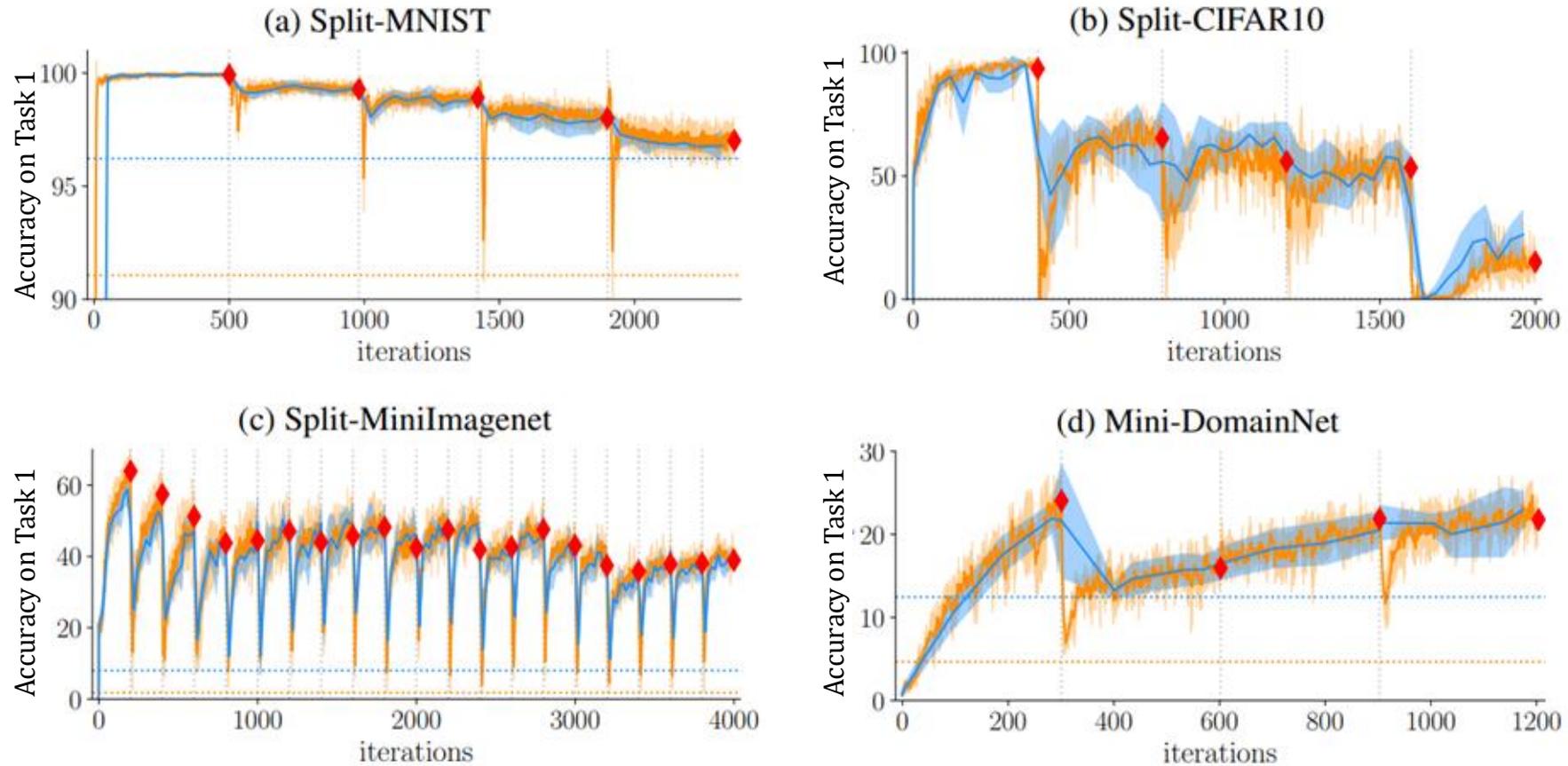
Replay, Class-incremental Split MiniImageNet



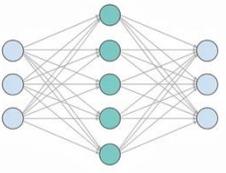
The stability gap is consistently observed



Replay, Class-incremental on ...

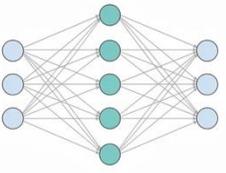


Why should we care?



- Scientifically interesting:
 - Insight into how replay works
- Problematic for safety-critical applications
 - Worst-case performance might be important
 - Could be exploited by adversarial agent with control over training stream
- Seems highly inefficient
 - Preventing forgetting seems more efficient than having to re-learn

How to avoid the stability gap?

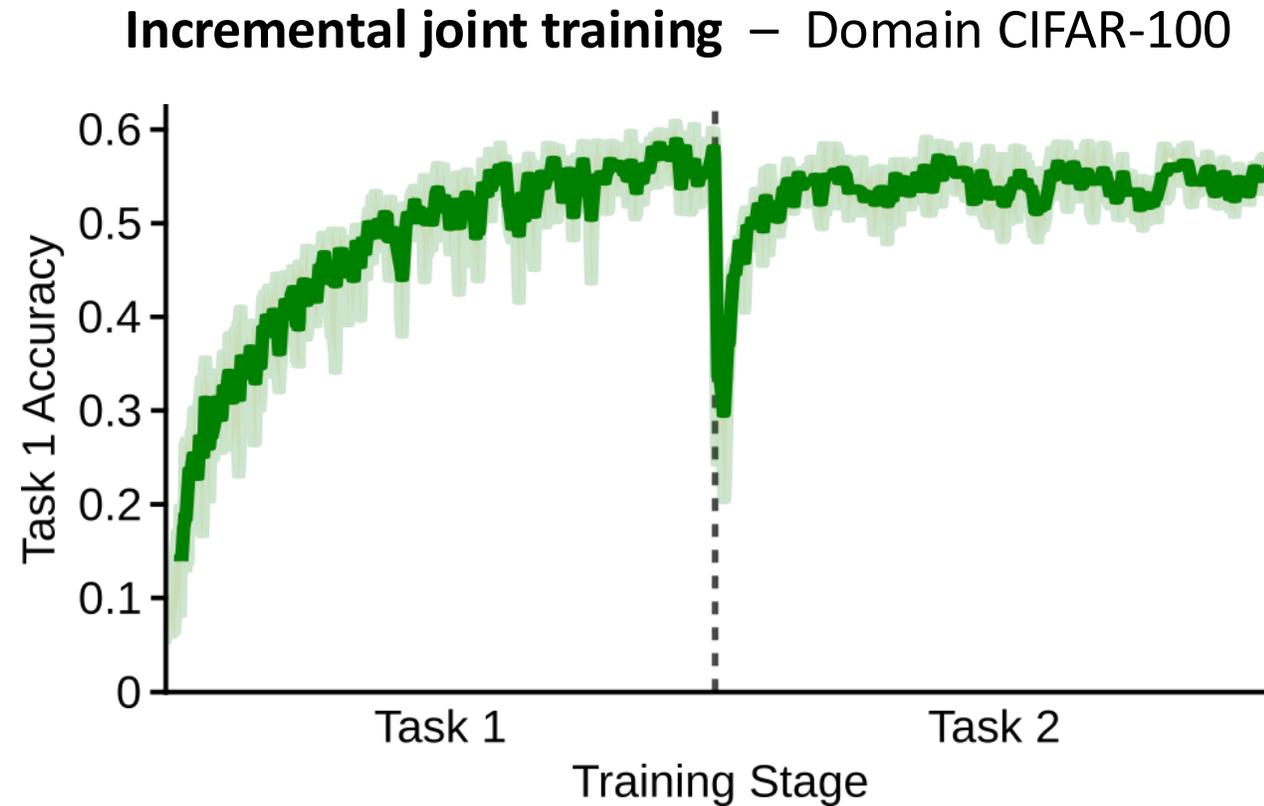


- By increasing the amount of replay?

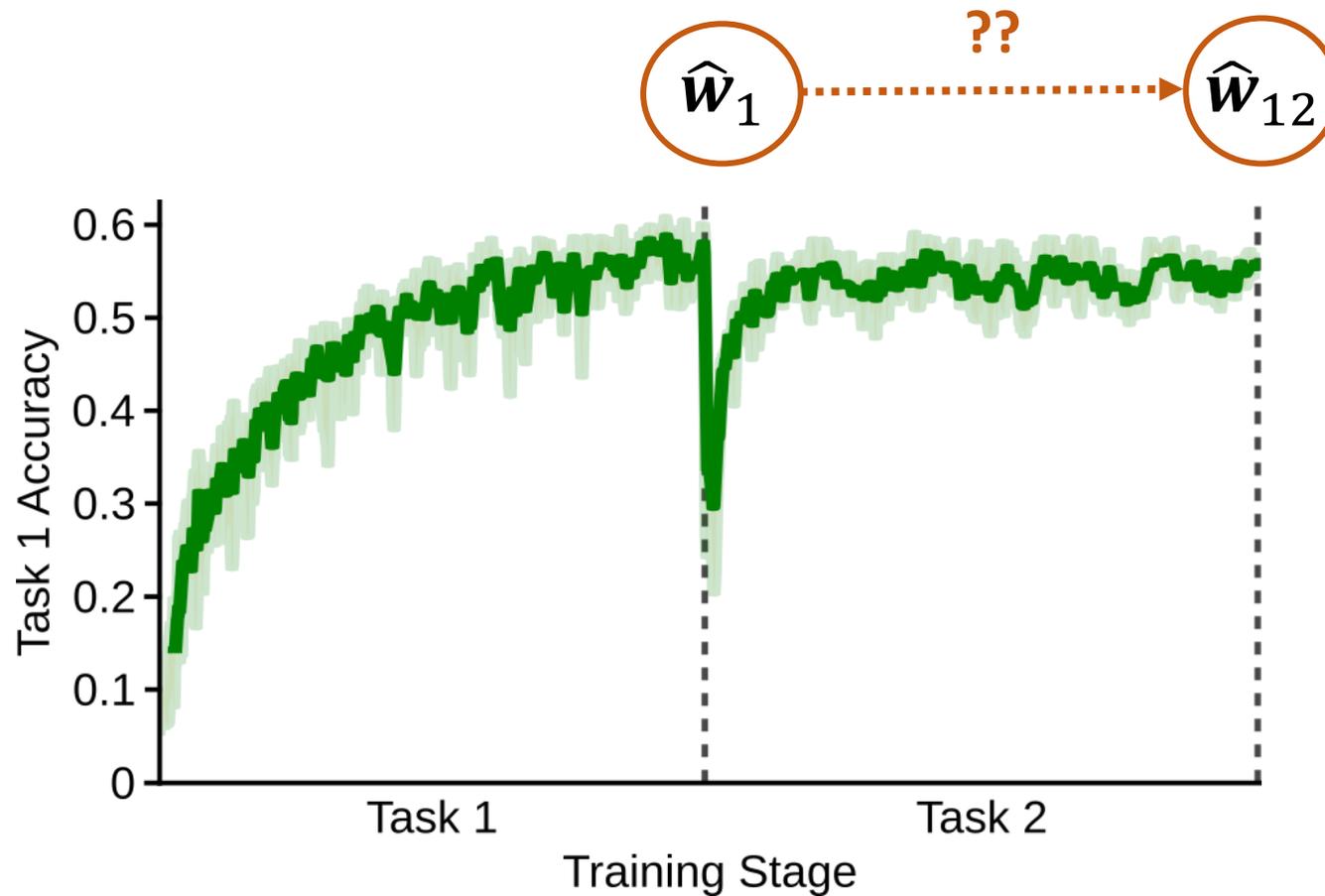
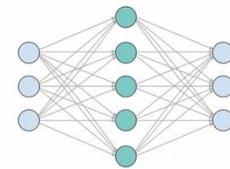
The stability gap occurs even with 'full' replay



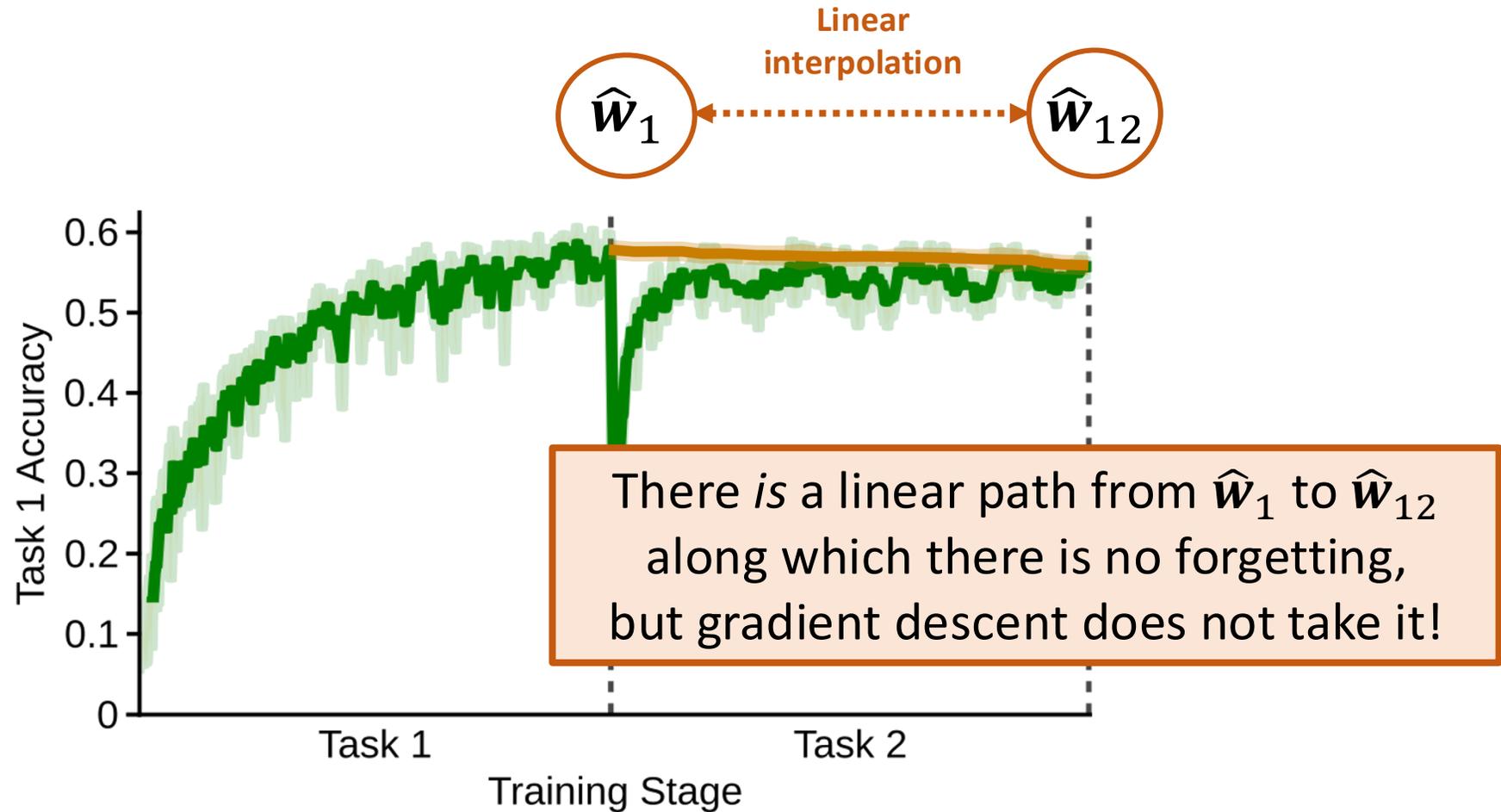
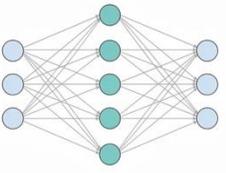
Timm Hess



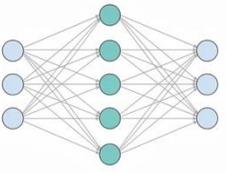
Can the stability gap be avoided at all?



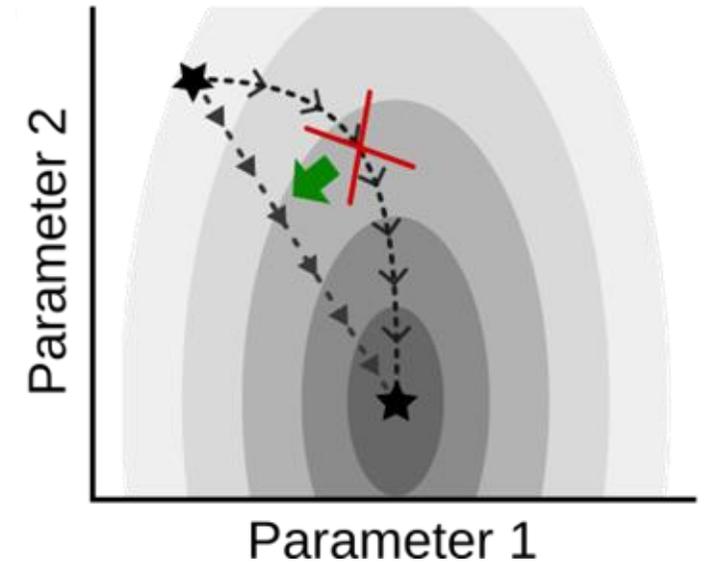
Can the stability gap be avoided at all? → Yes!



Continual learning needs a new direction

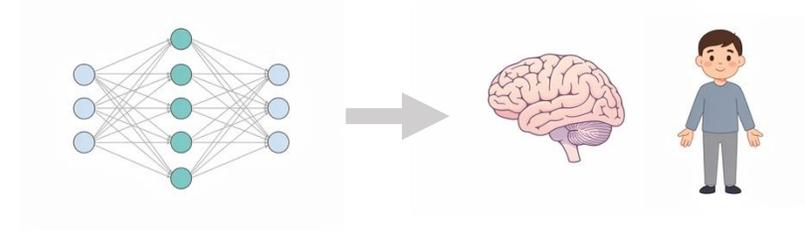


- To overcome the stability gap, changes must be made to *how* optimization is done
- Highlights **optimization-based continual learning** as a promising research direction



Temporary forgetting in humans?

- Do humans suffer from transient forgetting upon learning something new?



Temporary forgetting in humans?

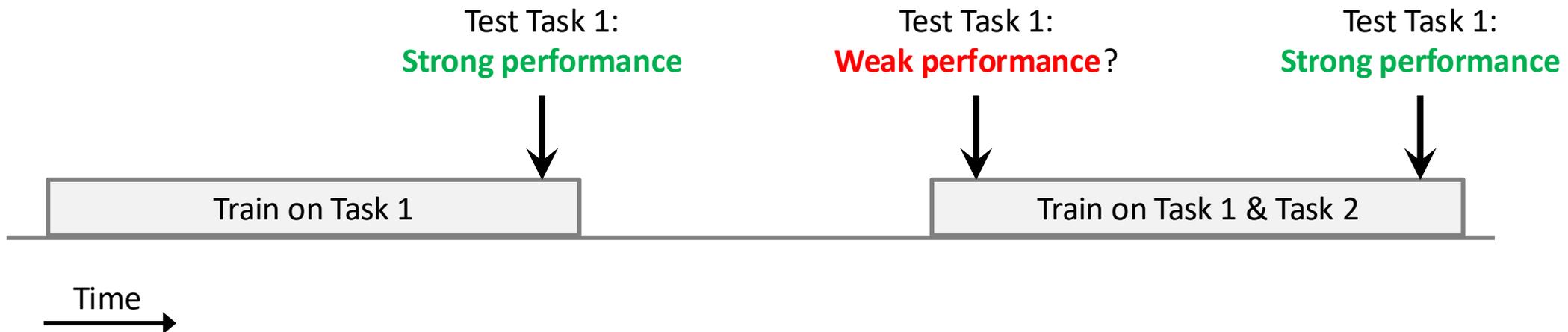


university of
groningen



Milan Van Maldegem

- Do humans suffer from transient forgetting upon learning something new?



Summary

- In the mouse brain, replay stabilizes memory-representing cell assemblies
- In artificial neural networks, replaying few samples can be enough because “not forgetting” is easier than “learning”
- Intriguingly, in artificial neural networks, replay consistently suffers from temporary forgetting when a new task starts
- Do humans suffer from similar temporary forgetting?

