Efficient learning through abstracted generative replay

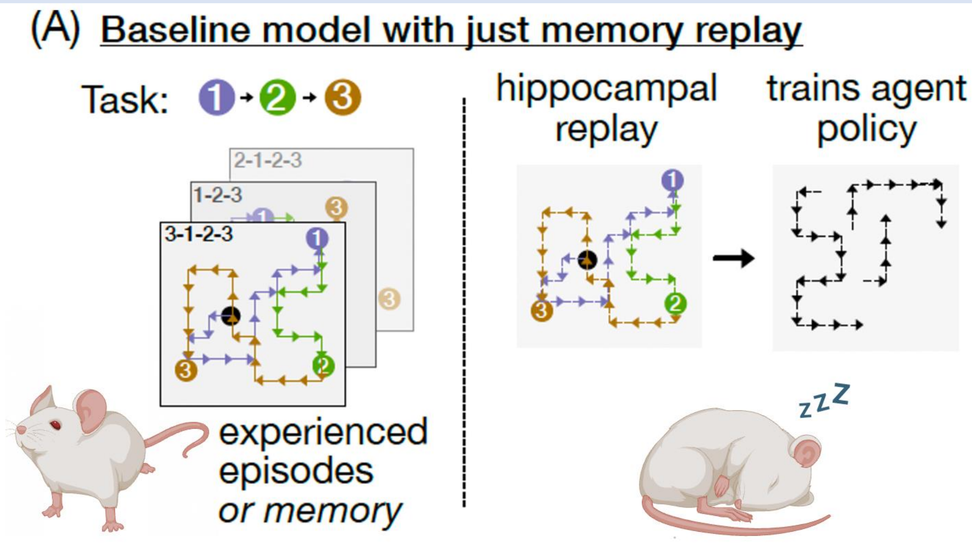
Eleanor Spens*[1], Diksha Gupta*[2], Emily Lewis [2], Andrea Castegnaro [2], Neil Burgess [2], Tom Mrsic-Flogel [2], Timothy Behrens[1,2]

[1]University of Oxford, [2]University College London

Learning efficiently from limited experience requires the brain to simulate new experiences and reuse existing abstractions.

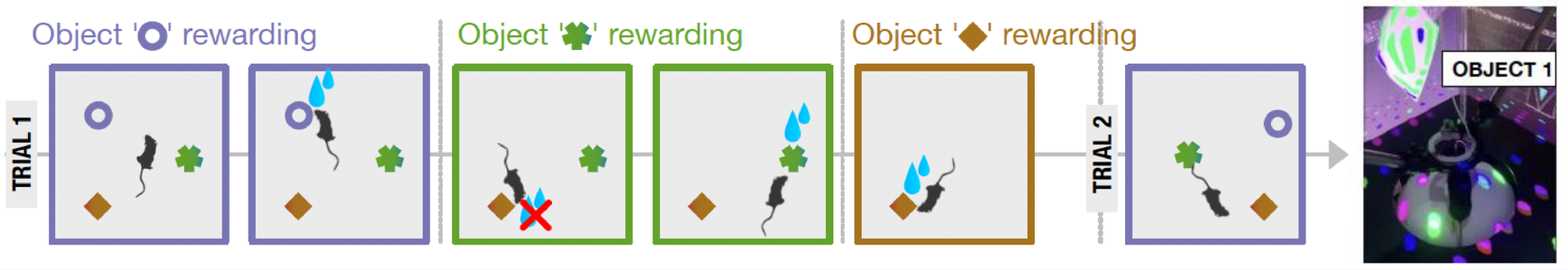
In reinforcement learning, **options** provide action chunks that simplify learning [1] while **generative replay** enables offline policy learning [2].

We propose that the brain combines these mechanisms to learn efficiently and compress memories.



TL;DR: An agent with i) options and ii) a world model can learn efficiently from limited experience with limited storage capacity, predicting how memory encoding and replay is abstracted over time.

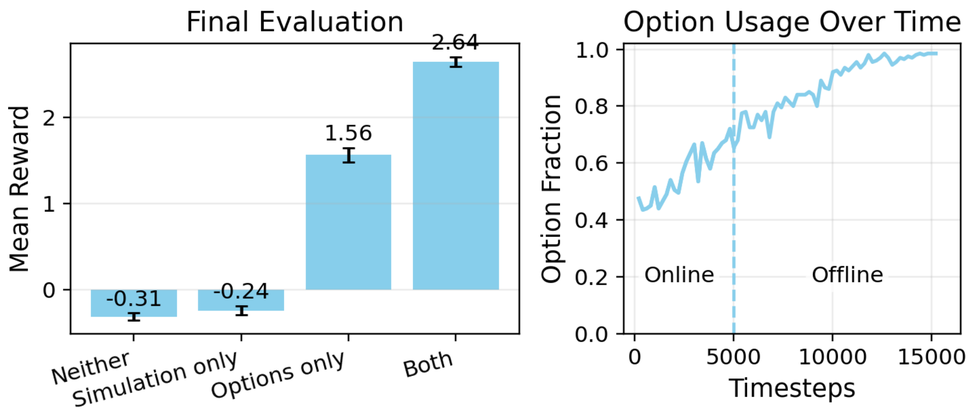
Task: Gupta et al. developed a mouse VR [3] task in which three objects must be visited in order, but their positions change every trial.



Simulations:

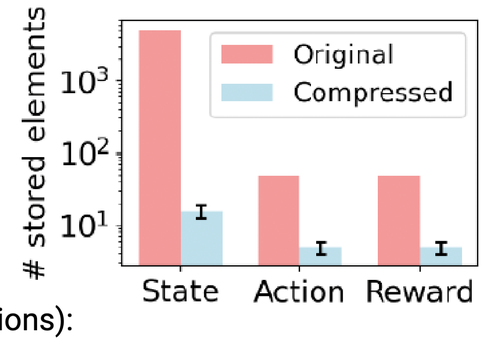
- Agents compared on task implemented in RL environment
- **Options** (e.g. "go to object X") learned from prior experience are reusable chunks that enable rapid learning of new policies
- A **world model** predicts outcomes of actions and options
- Best: agent with options encodes memories → update world model with memory replay → learn offline within world model

Over time generative replay involves simulated option sequences rather than atomic actions

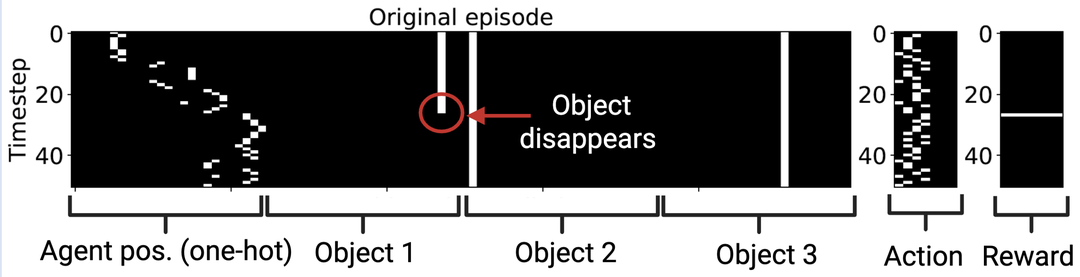


Different offline processes optimal depending on circumstances; [4] predicts *what* events it is optimal to replay *when*.

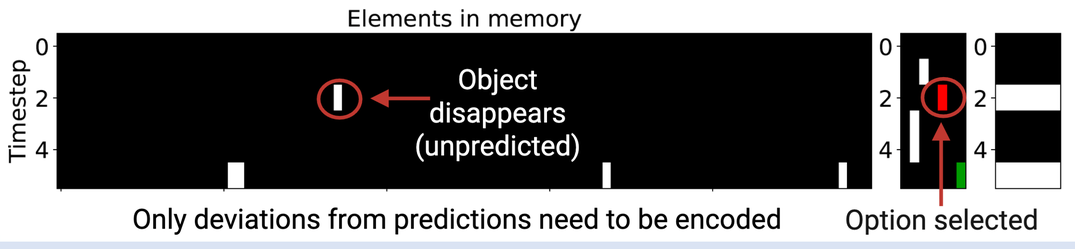
Experiences are **compressed in memory** by using a world model, with only unpredictable observations stored.



Example random policy episode (no options):



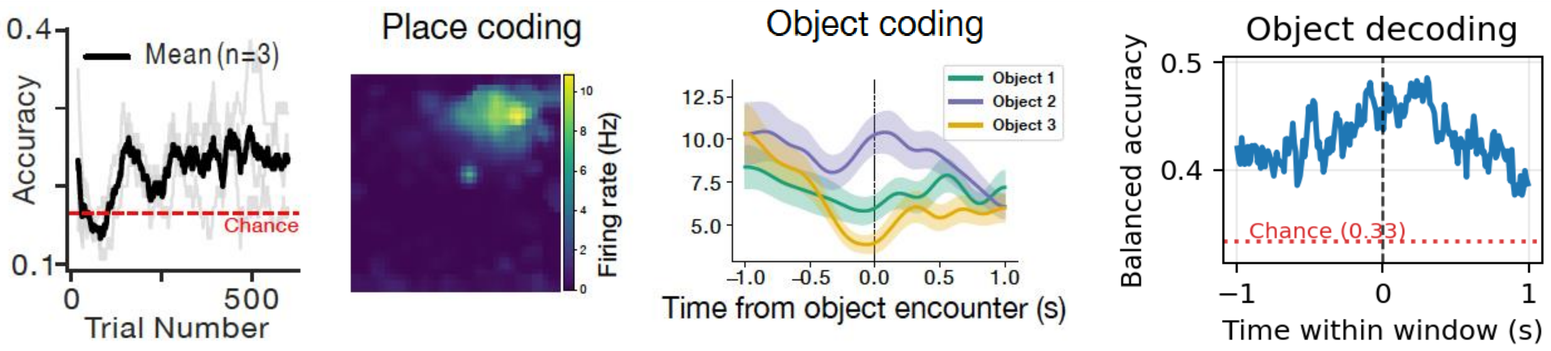
Options and WM reduce storage required:



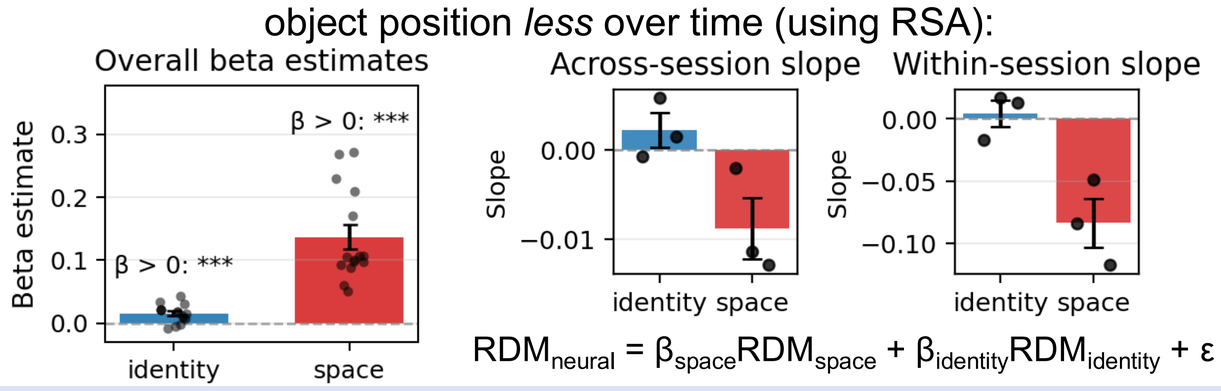
Preliminary data: Mice learn the Gupta et al. task. Hippocampal neurons recorded with Neuropixels [5] encode object identity as well as position.

Predictions:

- Replay becomes more abstract and option-based during learning.
- Object identity gets more decodable from replay while object position is abstracted away.
- Different kinds of replay are associated with training world model vs. simulating with that world model.



Hippocampal data at collisions reflects object identity *more* and object position *less* over time (using RSA):



$RDM_{neural} = \beta_{space} RDM_{space} + \beta_{identity} RDM_{identity} + \varepsilon$

References: [1] Sutton et al. (1999); [2] Sutton (1991); [3] Chen et al. (2018); [4] Spens et al. (2025); [5] Steinmetz et al. (2021).