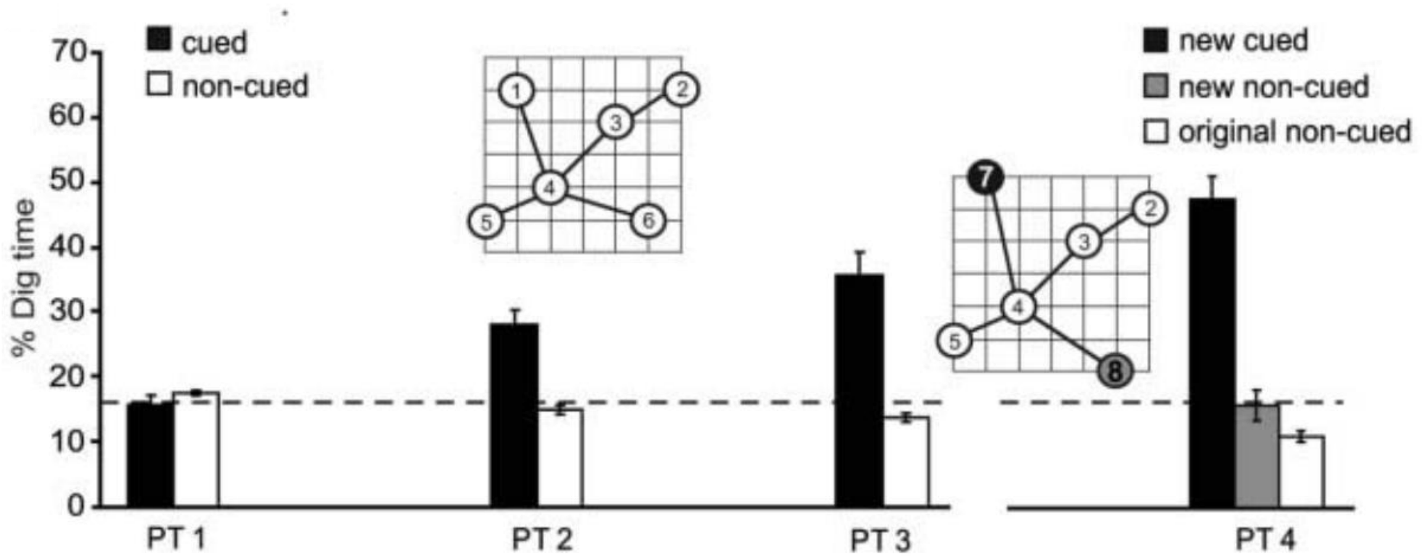
One-shot learning of paired association navigation with biologically plausible schemas

M Ganesh Kumar[1,2], Cheston Tan[2], Camilo Libedinsky[1], Shih-Cheng Yen[1], Andrew Tan[1]
National University of Singapore[1], Agency for Science, Technology and Research[2]
Corresponding E-mail: m_ganeshkumar@u.nus.edu

NUS Graduate School
Integrative Sciences & Engineering Programme
Centre for Frontier AI Research
CFAR

Please scan the QR code to
-- Access the preprint
-- Leave your comments
-- Connect

Summary



Biological phenomenon [1]:
-- Rats gradually learned to associate six flavor cues (given only at the start of the trial) to six corresponding target locations in a square arena over 20 sessions (PT1 to PT3).
-- On the 21st session, the rats rapidly associated two new flavor-location pairs (7 & 8) and demonstrated one-shot learning in the following probe session (PT4).

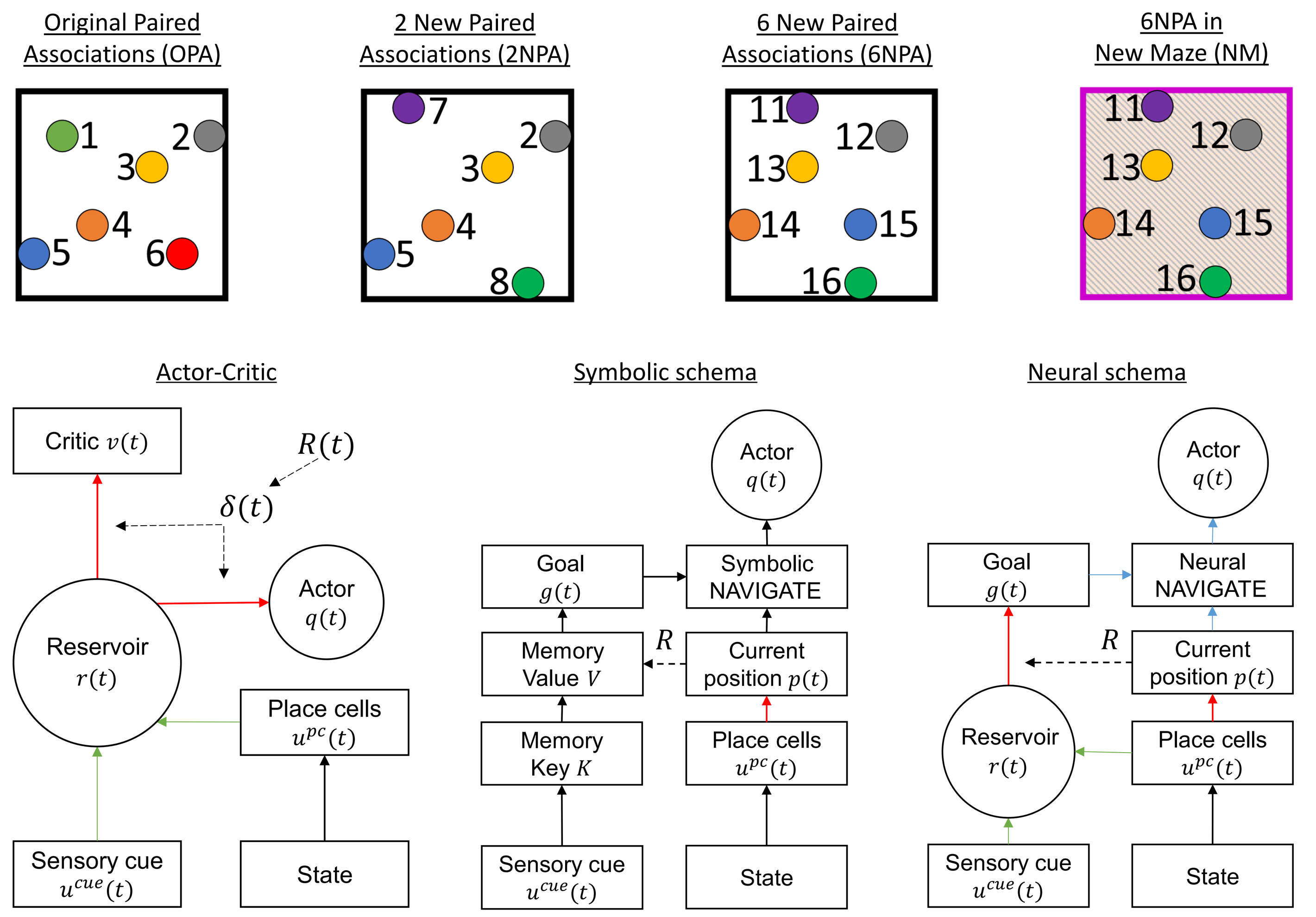
Hypothesis: Rats gradually learned a schema [2] which facilitated the subsequent one-shot learning of new paired associations.

Question: How do schemas described at the computational level correspond to the neural implementation level to explain one-shot learning of paired associations?

Methodology: We developed a reinforcement learning agent that
1) Gradually learns to path integrate using Temporal Difference error modulated local learning rule [3], the agent's estimated coordinates are used to perform vector-based navigation.
2) Rapidly learns flavor-location associations after one trial using reward modulated Hebbian learning rule [4], to recall the flavor associated goal coordinates in the next session.
3) Gradually learns to gate working memory input using Temporal Difference error modulated Hebbian learning rule [5] to attend to task relevant cue while ignoring distractor cues.

Findings: Our biologically plausible agent gradually learns the original paired associations task to demonstrate one-shot learning on novel pairs, even when distractor cues are presented.

Navigation arena & Agents



References

[1] Tse et al. (2007) Schemas and Memory Consolidation. *Science*.
[2] Rumelhart (1980) Schemata: The building blocks of cognition. *Theoretical Issues in Reading Comprehension*.
[3] Foster et al. (2000) A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus*.
[4] Hoerzer et al. (2014) Emergence of complex computational structures from chaotic neural networks through reward-modulated Hebbian learning. *Cerebral Cortex*.
[5] O'Reilly & Frank (2006) Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*.

Connecting cognitive schemas to neural implementation

LEARN METRIC REPRESENTATION schema

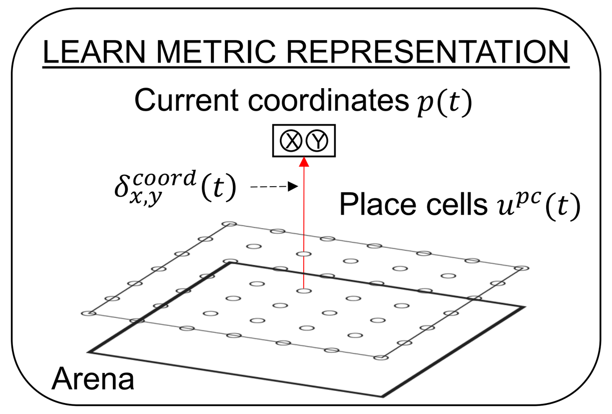
Computation
Use self-motion information and place cell activity to learn a continuous X–Y coordinates–based metric representation to estimate current coordinates.

Neural implementation
Place cells are used as input to the X and Y coordinate cells as outputs. The synapses from place cells to coordinate cells are gradually learned using a path integration derived temporal difference error

$$\delta_i^{coord}(t) = p_i(t) - p_i(t-\Delta t) - \hat{a}_i(t)$$

Which modulates the local learning rule using presynaptic activity-based eligibility traces

$$\Delta W_{ij}^{coord}(t) \propto e_j(t)\cdot \delta_i^{coord}(t)$$



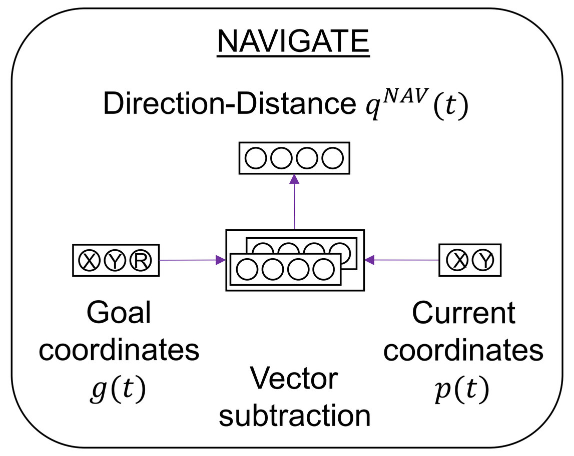
NAVIGATE schema

Computation
Take in the agent's current coordinates, recalled goal coordinates and reward recall value as inputs to output direction of movement.

Neural implementation
Neural network with two nonlinear hidden layers was pre-trained by backpropagation on a dataset with different current coordinate, goal coordinate, reward recall values and action. The direction to move was determined using vector subtraction between the agent's goal and current coordinates.

$$d_{j\in\{x,y\}}(t) = g_{j\in\{x,y\}}(t) - p_j(t)$$

Pre-trained weights were fixed during paired association task learning.



LEARN FLAVOR-LOCATION schema

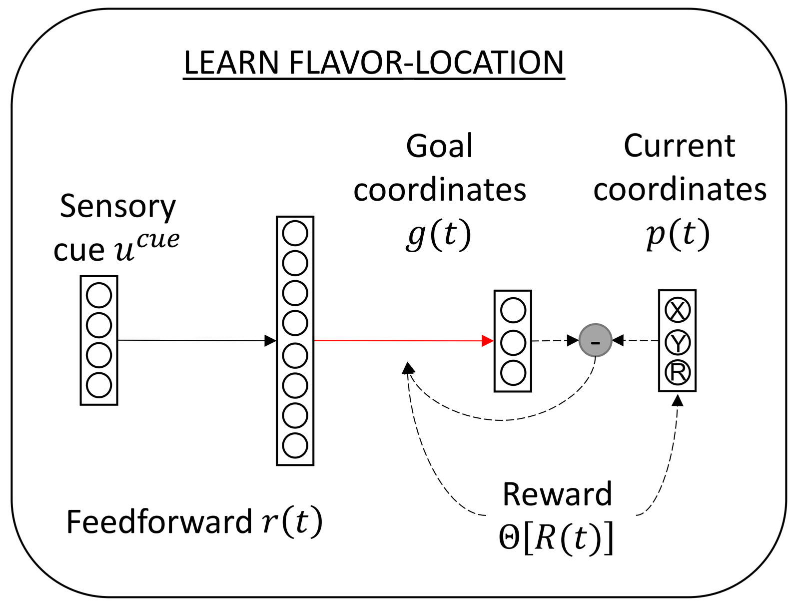
Computation
Associate the flavor cue with the coordinates at which a reward was disbursed after a single trial. Use flavor cues to recall the goal coordinates in the subsequent trial.

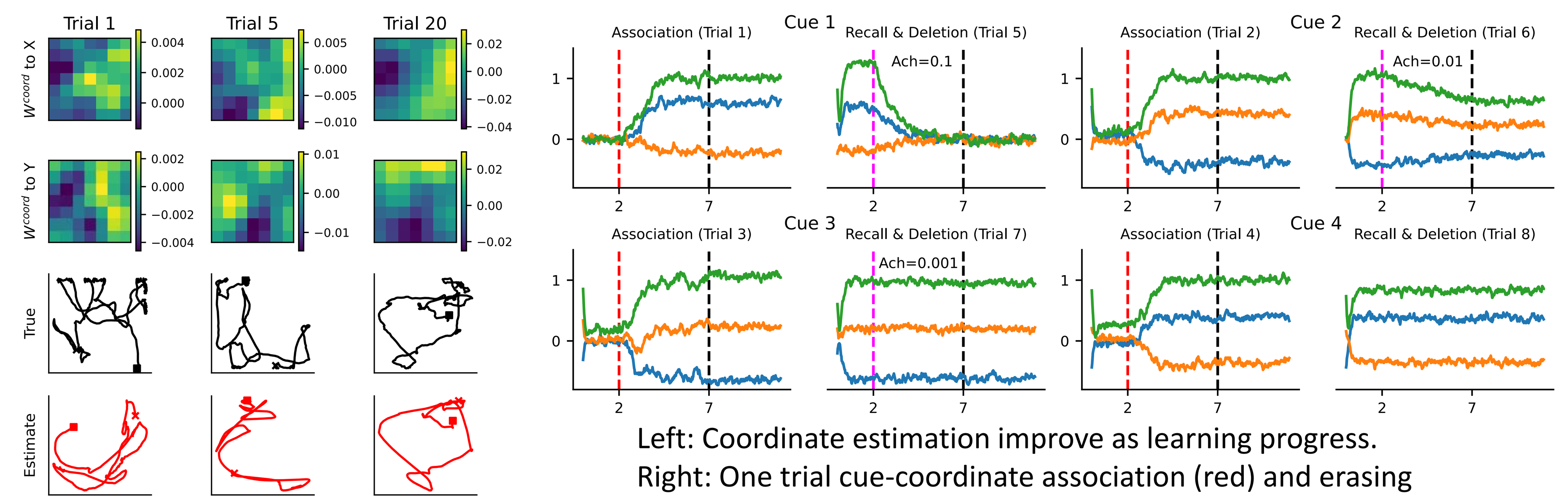
Neural implementation
The cue vector is passed as input to a feedforward layer or reservoir with three output units representing X, Y goal coordinates and reward recall value. Only the output synapses are subject to the reward modulated Exploratory Hebbian rule.

$$\Delta W_{ij}^{goal}(t) \propto r_j(t)\cdot (g_i^{noisy}(t) - g_i(t))\cdot M(t)\cdot \Theta[R(t)]$$

Associations are depressed or erased using acetylcholine modulated Hebbian learning when the flavour cue is given.

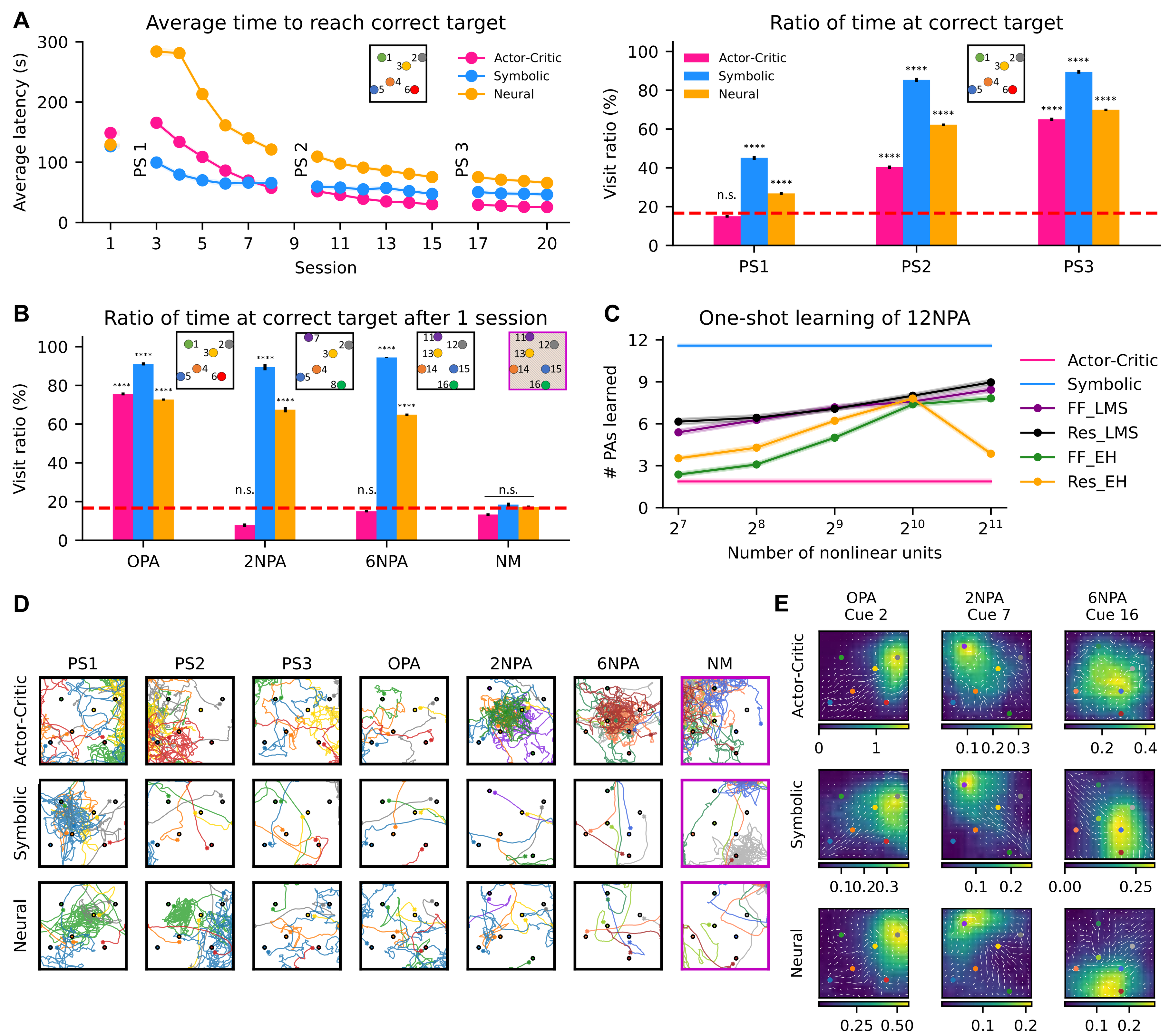
$$\Delta W_{ij}^{goal}(t) \propto r_j(t)\cdot g_i(t)\cdot -\Omega_{Ach}$$





Left: Coordinate estimation improve as learning progress.
Right: One trial cue-coordinate association (red) and erasing (magenta) using biologically plausible learning rules.

Gradual then one-shot learning of paired associations (PAs)



Gradually learning to gate working memory generalizes to new PAs