# Internally Organized Abstract Task Maps in the Mouse Medial Frontal Cortex

Mohamady El Gaby, Adam Harris, James Whittington, Mark Walton, Thomas Akam, Timothy Behrens

**Summary**

New tasks are often similar in structure to old ones. Animals that take advantage of such conserved or "abstract" task structures can master new tasks with minimal training. To understand the neural basis of this abstraction, we developed a novel behavioural paradigm for mice and recorded from their medial frontal neurons as they learned. Freely moving mice learned multiple tasks where they had to visit 4 rewarded locations in sequence (ABCD) on a 3x3 spatial maze. Tasks shared the same circular transition structure (… ABCDABCD …) but differed in the locations and geometric arrangement of rewards. As well as improving across tasks, mice inferred that A followed D (i.e. completed the loop) on the very first trial of a new task. This "zero-shot inference" is only possible if animals had learned the abstract structure of the task. Medial frontal cortex (mFC) neurons showed several signatures of internally organized tuning to abstract task-space. Firstly, the majority of state-tuned neurons in the mFC responded to the mouse's "location" in abstract task space, conserving their state tuning across distinct tasks. Secondly, we found robust, task-stage modulated offline replay of activity in task-space during sleep. Thirdly, a minority of state-tuned neurons remapped across tasks. Preliminary evidence suggests that such remapping is quasi-coherent across neurons, consistent with the existence of task-space modules analogous to modules of grid-cells that coherently remap in physical space. These findings point to separable neuronal substrates for internally organised representations of task structure that may guide abstraction in the mammalian brain.

**Additional Details**

When faced with a new task we are rarely completely clueless. We can often draw upon previous experiences of similarly structured tasks to make a "best guess" at how to solve the task at hand[1,2]. How are we able to do this? Behavioural, neuroimaging and computational studies suggest that humans learn transition statistics that are common amongst multiple tasks[3–6]. Knowledge of the transition structure relating latent task states can be used to predict the next step in a new task. However, the neuronal mechanisms of such abstract task structure knowledge remain unexplored. To achieve such cellular insight, we designed a novel behavioural paradigm for mice. Mice learned multiple tasks, each of which involved finding four reward locations (***a***, ***b***, ***c*** and ***d***) in a simple 3x3 spatial maze (**Figure 1a**). Rewards were available sequentially within a circular task structure (***abcdabcda*** ... *etc*), with reward acquisition at each location defining the beginning of a new task state (A,B,C or D; **Figure 1a**). State A was signified by a brief (2 second) tone upon reward consumption in ***a***. The locations of the rewards, their relative distances and their geometrical relationships changed across tasks (**Figure 1a**). Moreover, distances in task space were unrelated to those in physical space (e.g. ***a*** and ***b*** were on average as close to each other in physical space as ***a*** and ***c***). Mice improved in performance across tasks to the point that, on latter tasks, they displayed a direct sign of abstract structural knowledge: zero shot inference. On the first trial of a new task, having experienced all four rewards only once, animals took the shortest route from ***d*** back to ***a*** more than chance (**Figure 1b**), indicating they had learned the abstract four node loop structure of the task.

To investigate the neuronal mechanisms of this abstraction, we implanted silicon probes chronically into the medial prefrontal cortex (prelimbic area) of freely moving mice from the outset of training (**Figure 1c**). Neural responses were strongly tuned to the relative distance between two goals (**Figure 1d**). Some neurons fired between every pair of goals regardless of state **("Phase cells", Figure 1d)**, in line with previous studies[7]. Crucially, a large proportion of neurons were tuned to task state (**Figure 1d**). Of these neurons, the majority maintained their tuning to the same state across tasks, despite such states being associated with distinct reward locations (**"Generalizing state cells"**; **Figure 1d,e**). Such tuning was unrelated to lower level factors such as the animal's speed, acceleration or location/location type (i.e. corners vs middle). To investigate whether such state tuning of mFC neurons was internally organized, we recorded the activity of the same prefrontal neurons during off-task rest/sleep in a separate spatial context. Generalizing state cells exhibited significant reverse replay in task space despite animals only traversing task space in the forward direction (21% significant events; $Z=3.56$, $P<0.001$, **Figure 1f**). A smaller subset of state-tuned mFC neurons remapped in task space, with their firing fields rotating to a different state across tasks while maintaining their phase tuning (**"Remapping state cells"**, **Figure 1d,e**). This was not always explained by their tuning to physical space. Moreover, we found preliminary evidence that subsets of neurons remap coherently, maintaining their relative angles in task space across distinct tasks (**Figure 1g**). Coherently remapping neuron pairs showed higher cofiring during off-task sleep than incoherent pairs (**Figure 1h**). This is analogous to modules of grid cells that rotate coherently across distinct spatial contexts[8,9] and maintain higher cofiring during sleep[10]. Overall, these findings suggest that mFC neural activity is organized based on the abstract relational structure of the task, either by being anchored to salient cues (i.e. tone-anchored **"Generalizing state cells"**) or relative to each other (**"Coherently Remapping state cells"**). These findings open avenues into understanding the cellular mechanisms and computational principles underlying abstract relational learning, including analogies with other cortical regions mapping physical space.

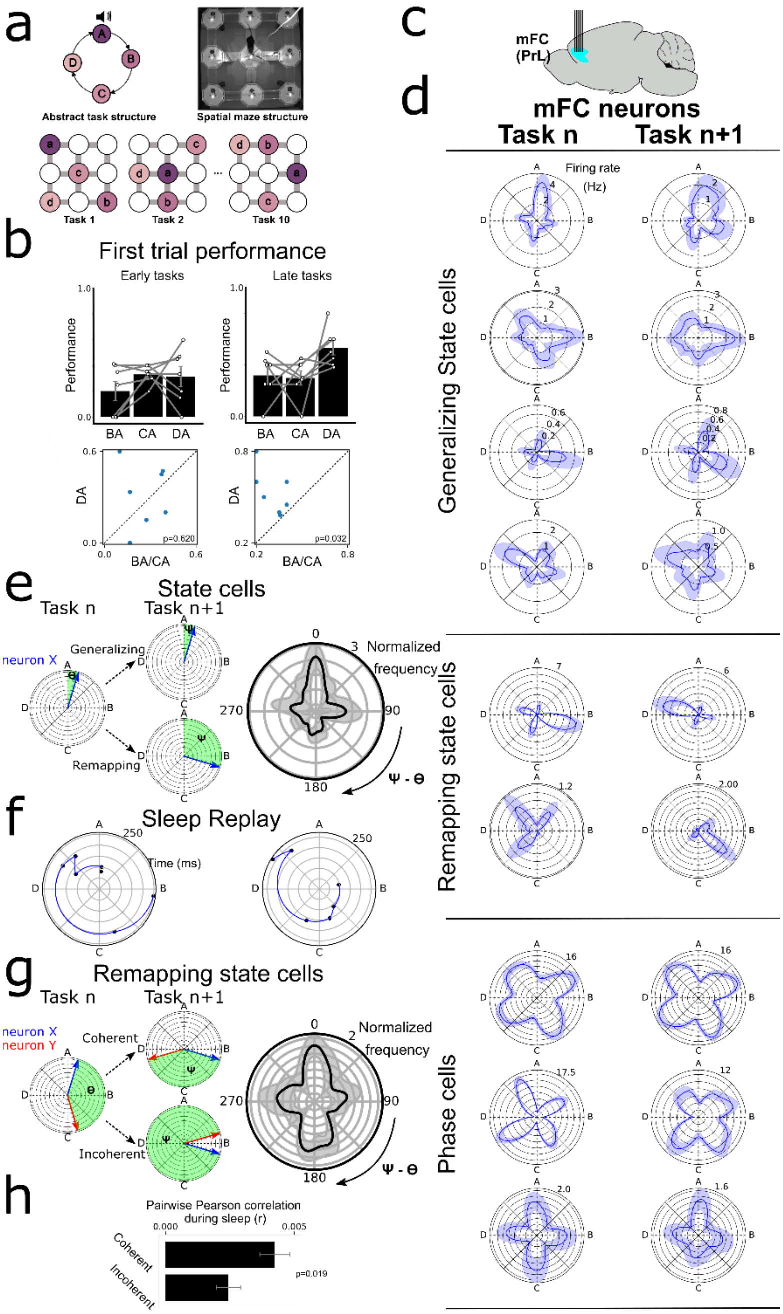


**Figure 1** **a)** All tasks followed a 4 state circular structure while reward locations in the spatial maze change across tasks. **b)** Structure guided zero-shot inference in seven mice. Plots show performance on the first trial of a new task, averaged across the tasks in the first half ("Early tasks") or second half ("Late tasks") of the total tasks completed (40 for 3 mice, 10 for 4 mice). The plots quantify the percentage of times an animal navigates via the shortest route directly from reward location ***d*** to reward location ***a*** on the first trial compared to animals (prematurely) taking the shortest route from location ***c*** or ***b*** back to ***a***. **c)** Chronic silicon probe implants targeting 64 channels to mFC (prelimbic cortex). **d)** Polar plots showing example "generalizing" (top) and "remapping" (middle) state cells as well as state untuned "phase cells" (bottom) recorded from the mFC. Dashed lines represent reward delivery time-points in each state. **e)** Left: schematics of generalizing and remapping state-cells. Right: Smoothed polar histogram of peak angle difference across tasks for state-tuned mFC neurons. Peak at zero indicates generalization. **f)** Polar plots showing example reverse (left) and forward (right) replays in task space during sleep. Time is on the radial axis. **g)** Left: schematic of state cell pairs that remap coherently (i.e. maintain their relative angle) and those that remap incoherently. Right: Smoothed polar histogram showing changes in relative angle between pairs of remapping state cells across tasks. Peak at zero indicates relative angle between neurons remains unchanged across tasks. **h)** Coherently remapping state cell pairs are more coactive (25 ms time-windows) during sleep than incoherently remapping pairs.

1. Lake et al *Behav. Brain Sci.* **40**, e253 (2017).
2. Tenenbaum et al *Science* **331**, 1279-1285 (2011).
3. Wilson et al *Neuron* **81**, 267-279 (2014).
4. Behrens *et al Neuron* **100**, 490-509 (2018).
5. Baram et al *Neuron* **109**, 713-723.e7 (2021).
6. Gardner & Schoenbaum (2020) PsyArXiv
7. Rubin et al Nat. Commun. 10, 4745 (2019).
8. Fyhn et al *Nature* **446**, 190-194 (2007).
9. Stensola et al. Nature 492, 72–78 (2012).
10. Gardner et al *Nat. Neurosci.* **22**, 598-608 (2019).