Flexible decision-making engages generalizable spiraling dynamical motifs in prefrontal cortex

Summary: The prefrontal cortex (PFC) is critical for flexible decision-making, and dynamical motifs have been identified in PFC population activity that could support certain aspects of this process (e.g., context-dependent line attractors for sensory input integration[1]). PFC is also known to encode a plethora of decision-relevant variables, including task progression, action, and value[2,3]. It remains an open question whether there are dynamical motifs that track all these variables to inform decisions. Further, it is unclear how outcome feedback alters neural representations of these variables and impacts the evolution of network dynamics. Here we addressed these unknowns by examining population dynamics and single neuron activity in the medial PFC (mPFC) as rats made foraging decisions in an environment where probabilistic rewards were distributed over multiple foraging patches. On each trial, rats chose to stay within their current patch or switch to alternative patches based on their estimates of the relative value between the local and distant locations. We found regular structure in mPFC population activity that resembled a high-dimensional 'spiral' where trial progression drove rotational dynamics that were displaced by value. Unrewarded trials shifted the system in a direction associated with a patch-switch decision, but when the rat switched to a new patch, the system was reset in the opposite, patch-stay direction even before an outcome was experienced in that patch. This spiral-like motif was expressed across different patches and for different actions, with each spiral located in a different region of the neural state space, preserving information about places and trial types. These features arose from individual neurons selective for path progression, action, and/or patch identity with multiplicative modulation by value. Our results demonstrate that the geometry of prefrontal dynamics and associated single neuron representations are well-suited to track task-relevant variables and support adaptive decisions across a variety of conditions.

Additional detail: The maze consists of three bifurcating corridors leading to six reward sites (two sites belonging to the same corridor form a 'foraging patch'; **Fig. 1A**)[4]. Rewards were dispensed at each site probabilistically and consecutive visits to the same port were never rewarded. The reward probability at each site depleted following each visit to that site, repleting only when rats went to a distant patch. A hidden Markov model (HMM) fit to rat behavior allowed us to infer subjective choice values and capture value-based foraging decisions: rats (n=4) overall visited patches of higher value more frequently and were more likely to leave a patch as the reward became depleted (**Fig. 1B**). Logistic regressions predicting Switch/Stay choices revealed significant effects of both the Stay value (the value of the current patch) and the Switch value (the maximal value of alternative patches) (**Fig. 1C**). We thus set out to understand the neural basis of the relative Switch/Stay value (hereafter called 'value').

As rats traveled between reward sites, the activity profiles of individual mPFC neurons showed two notable features: 1) temporal patterns that were generalized across journeys sharing common task structures – e.g., path progression and action sequence – regardless of the specific location in the physical space; 2) value modulation that appeared to be a multiplicative gain on the task-structure modulation, resulting in different neurons showing modulation at different levels of cross-condition generality (**Fig. 2**). To understand how neurons implement task-relevant computations beyond the heterogeneity of their individual responses, we looked for dynamical motifs in

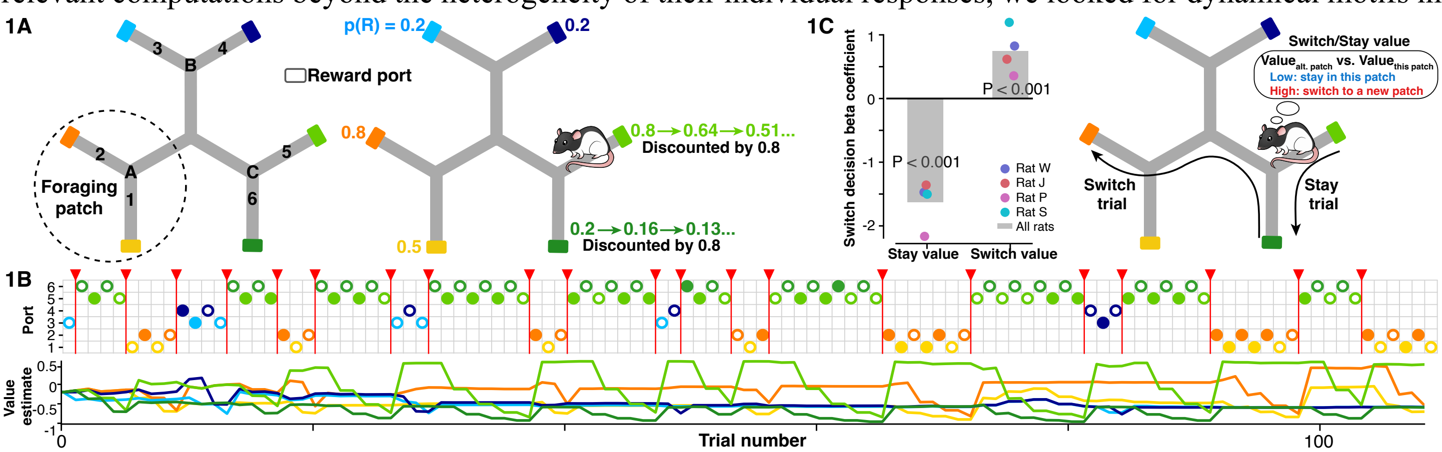


Fig.1 Task design and behavior. A, Left: maze diagram. Right: an example set of reward probabilities, labeled as p(R). If rats only exploit two reward sites in one patch, their p(R) will be discounted by 0.8 after each visit. Rats need to visit a different patch for the reward to replete. **B,** Top: example sequence of choices from the first 109 trials of one representative behavioral session. Filled circles represent rewarded trials and empty circles represent unrewarded trials, color-coded by site identity. Red triangles indicate a patch switch. Bottom: the value of each reward site estimated by the HMM based on the rat's choice and reward history in the same session. **C,** The estimated values of staying in the current patch and switching to an alternative patch both significantly impact rats' choices.

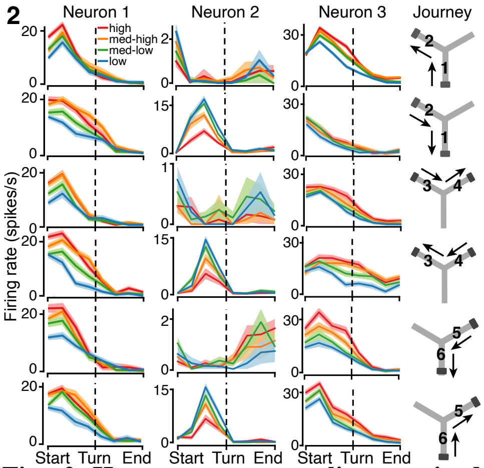
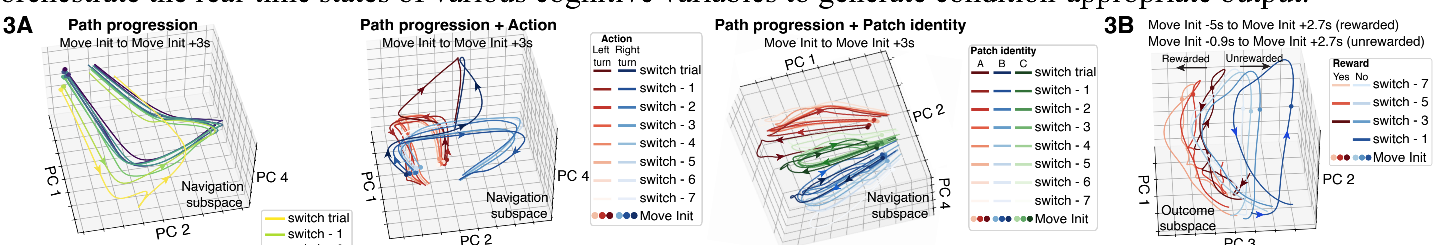


Fig. 2 Heterogeneous, nonlinear mixed code for task structure and value in example mPFC neurons. Neuron 1 shows consistent task-structure and value coding across all 6 types of journeys; neuron 2 shows generalized structure and value coding for all right-turn (action) journeys; neuron 3 shows similar temporal firing patterns and value modulation for journeys in the same patch (patch 3/4 or patch 5/6). Color reflects value.

the population activity, i.e., recurring activity patterns that can support specific computations. Thus, we applied PCA to mPFC population activity during Stay trials leading up to Switch decisions for navigation and outcome periods separately. Neural trajectories in both PC subspaces not only depicted path progression, action sequence (left turn vs. right turn), and patch identity, but were aligned in an order reflecting the trial-by-trial Switch/Stay value updates (**Fig. 3A**). Decision-predictive value codes were implemented by a spirally shaped dynamical motif that was shifted in opposite directions in the state space by reward outcomes (**Fig. 3B**) and reset to the starting point when rats entered a new patch (**Fig. 3A** bottom left). The spiral motif of value tracking during path progression was employed across all journeys, journeys sharing action sequence, and journeys sharing patch identity (**Fig. 3A**). We verified this multi-level abstraction of coding by quantifying the 'cross-condition generalization performance (CCGP)' of task-structure and value representations, a metric assessing the ability of a decoder to generalize to held-out task conditions (**Fig. 3C**)[5]. Finally, to directly extract the Switch/Stay value from neural dynamics, we built cross-validated value decoders at multiple levels of abstraction (**Fig. 4A**). The value readout was consistent with patterns in the PC subspaces (**Fig. 4B**). <u>The dynamical motif shared in specialized and generalized computations in this study provides biological evidence for theories of cognitive flexibility that propose a hierarchy of elementary processes reconfigured across similar conditions[6,7].</u> It also suggests a computational strategy that allows a dynamical system to track and orchestrate the real-time states of various cognitive variables to generate condition-appropriate output.



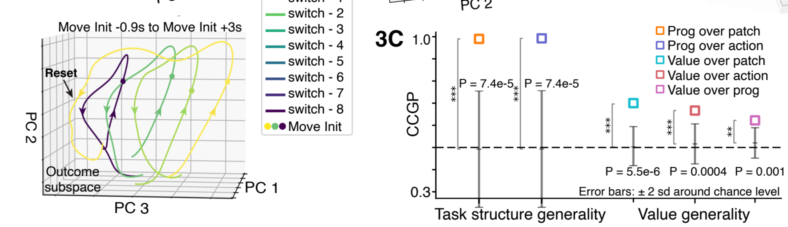


Fig.3 Dynamical motifs of the neural population activity. A, Neural trajectories in the PC subspace constructed from navigation (top panels) period reflect path progression, action sequence, and patch identity. Neural trajectories in the outcome-period PC subspace (bottom-left panel) also reflect these variables, with path-progression trajectories shown as a representative example of the motif. The arrangement and spiral geometry of these trajectories reflect the gradually increasing Switch/Stay value across trials leading up to a switch decision. **B,** Reward outcomes drive the neural dynamics in opposite directions. Neural trajectories in the outcome-period PC subspace are shown. **A, B,** Filled circles: movement initiation ('Move Init'); Colored arrows: the direction of neural trajectory evolution. **C,** GGCP = 1 means full generalization and CCGP = 0 means no generalization. Here, CCGP is close to 1 for path-progression coding across patch and action conditions. CCGP is lower but significantly greater than chance for value across task-structure conditions, suggesting coexistence of condition-specific and generalized value coding in the same neural population. **A-C,** Results from one representative rat.

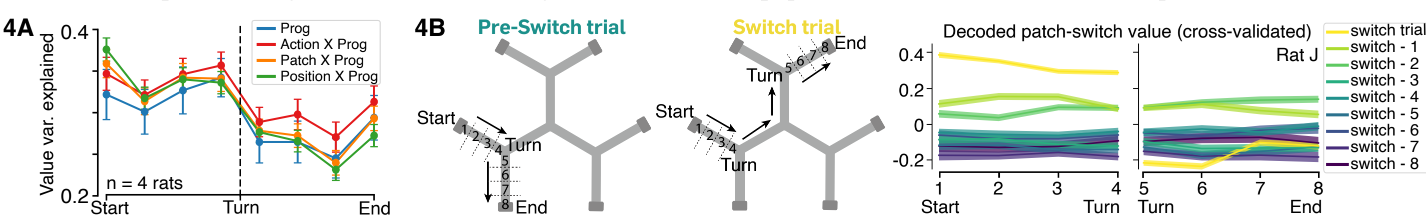


Fig.4 Decoding relative Switch/Stay value. A, Decoding performance at different levels of generality (highest generality: progression; lowest generality: position by progression). All decoders explain 20-40% of the variance in the Switch/Stay value, with the action-based decoder overall performing the best. **B,** Decoded value reveals its gradual ramp-up on trials leading up to a switch (violet to yellow) before rats express decisions (bins 1-4). The right panel shows readout from the action-based value decoder as a representative example.

References: [1] Mante & Sussillo et al., *Nature*, 2013. [2] Enel et al., *eLife*, 2020. [3] Guidera et al., *bioRxiv*, 2024. [4] Comrie et al., *bioRxiv*, 2024. [5] Bernadi et al., *Cell*, 2020. [6] Cole et al., *Nat Neurosci*, 2013. [7] Driscoll et al., *Nat Neurosci*, 2024.