Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2023 April 24.

Published in final edited form as:

Cell Rep. 2023 March 28; 42(3): 112246. doi:10.1016/j.celrep.2023.112246.

Geometric transformation of cognitive maps for generalization across hippocampal-prefrontal circuits

Wenbo Tang[1,2,3,*], **Justin D. Shin**[1,3], **Shantanu P. Jadhav**[1,4,*]

[1]Neuroscience Program, Department of Psychology, and Volen National Center for Complex Systems, Brandeis University, Waltham, MA 02453, USA

[2]Present address: Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853, USA

[3]These authors contributed equally

[4]Lead contact

SUMMARY

The ability to abstract information to guide decisions during navigation across changing environments is essential for adaptation and requires the integrity of the hippocampal-prefrontal circuitry. The hippocampus encodes navigational information in a cognitive map, but it remains unclear how cognitive maps are transformed across hippocampal-prefrontal circuits to support abstraction and generalization. Here, we simultaneously record hippocampal-prefrontal ensembles as rats generalize navigational rules across distinct environments. We find that, whereas hippocampal representational maps maintain specificity of separate environments, prefrontal maps generalize across environments. Furthermore, while both maps are structured within a neural manifold of population activity, they have distinct representational geometries. Prefrontal geometry enables abstraction of rule-informative variables, a representational format that generalizes to novel conditions of existing variable classes. Hippocampal geometry lacks such abstraction. Together, these findings elucidate how cognitive maps are structured into distinct geometric representations to support abstraction and generalization while maintaining memory specificity.

In brief

Tang et al. show that, as rats generalize rules across environments, cognitive maps are constructed on neural manifolds in the hippocampus and prefrontal cortex with distinct geometries

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

[*]Correspondence: wbtang@brandeis.edu (W.T.), shantanu@brandeis.edu (S.P.J.).

AUTHOR CONTRIBUTIONS

W.T. and S.P.J. designed the study. J.D.S. collected the data. W.T. conceptualized and performed the analysis. W.T. and S.P.J. interpreted data. W.T. visualized data. W.T. and S.P.J. wrote the paper. S.P.J. supervised the project and obtained funding.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

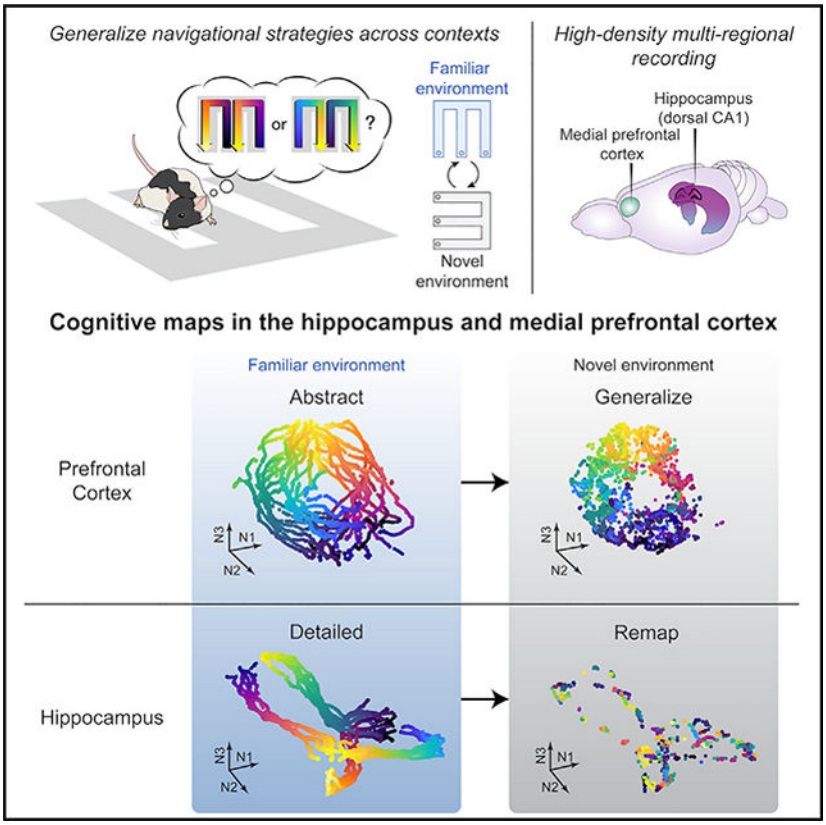
SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112246.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

for complementary functions. Prefrontal geometry enables abstraction of task knowledge generalizable across contexts, whereas hippocampal manifolds remap to maintain memory specificity.

Graphical Abstract



INTRODUCTION

Memory generalization, the ability to abstract knowledge from prior experiences, is critical for adaptive behavior in novel situations. In the spatial memory domain, generalization manifests in the application of learned navigational rules across environments with distinct sensory cues. Animals can thus rapidly adapt in new contexts without having to learn *de novo*. Considerable research in humans and animals has established that the hippocampus is essential for memory of contextual details[1,2] and that generalization across contexts requires the medial prefrontal cortex (PFC).[3–7] Memory generalization thus depends on distributed circuits in the hippocampus and PFC; however, little is known about the underlying neural representations.

Memory generalization can be facilitated by a systematic organization of past episodes into a map-like representation, a “cognitive map,”[8] and a promising neural substrate for spatial cognitive maps is provided by hippocampal place cells.[9] Place cells form context-specific spatial maps by responding to particular locations in an environment and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

remapping to distinct place codes in a new environment.[9–12] How these spatial cognitive maps are transformed across hippocampal-prefrontal representations to enable abstraction and generalization remains largely unknown. Recently, several studies have reported spatially tuned neurons in rodent PFC that strikingly resemble hippocampal place cells.[13–23] Furthermore, similar to the hippocampus,[24–26] spatial locations are conjunctively encoded with other task-relevant variables in the PFC,[17,27–30] termed "mixed selectivity."[31–33] These parallel representations with shared task information raise a fundamental question about how these representations can also subserve the distinct functions in memory generalization ascribed to the hippocampus and PFC.

An emerging theoretical framework suggests that circuit function is determined not only by representational content but also by representational geometry, the structural arrangement of task-relevant variables in a space of neural population states, called a manifold.[26,34–36] A recent study demonstrated the existence of a geometric neural manifold in the hippocampus that represents task knowledge in the current environment during navigation.[26] However, it is unclear whether the PFC preserves or reorganizes the hippocampal manifold structure and whether and how these representational geometries contribute to generalization across environments. To address these questions, we simultaneously recorded hippocampal-PFC neural ensembles as rats generalized memory-guided decision-making in a hippocampal-PFC-dependent navigation task across two distinct environments. This approach allows for a quantitative comparison of the intrinsic structures of hippocampal-PFC representations, as well as their dynamics, during generalization of navigational behavior.

RESULTS

Rats generalize learned rules across contexts

We trained five rats to perform a spatial memory task known to require hippocampal-prefrontal interactions[37–39] and to switch between familiar and novel environments. In this continuous alternation W-maze task (Figure 1A, left), rats had to return to the center (C) from either the right (R) or left (L) side of a W-track (inbound trajectory left or right [IN_L or IN_R]) and then correctly choose the opposite side (outbound trajectory right or left [OUT_R or OUT_L]) to be rewarded at the end of each trajectory. Therefore, the critical memory demand of this task is to learn and remember the correct IN-OUT sequences (IN_R-OUT_L and IN_L-OUT_R), termed "task sequences."

Animals initially had no prior W-track experience and first learned the task in one environment with at least eight 20-min W-track sessions interleaved with rest sessions in a sleep box (*de novo* learning; final performance, 84.3% ± 7.7%, mean ± SD; STAR Methods). On a subsequent day, they were introduced to a novel W-track in a visually distinct room for one session (denoted as N; Figure S1A), then returned to the familiar environment (F) where they had learned the task *de novo* on previous days, and finally reexposed to the novel environment (N′; Figure 1A, right). Compared with *de novo* learning, all animals acquired the inbound and outbound task rules faster in the novel environment (Figures 1B, 1C, and S1B). To determine whether rats had indeed generalized learned rules, instead of confusing the novel environment with the familiar one, we examined the performance of another set of well-learned rats (n = 3), which were exposed to the familiar

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

environment first and then switched to the novel environment within a day, for an opposite order of exposures as F-N-F′-N′ (STAR Methods). We found that for both animal groups, the performance in the familiar environment was significantly higher than that in the novel one regardless of the order of experience (Figures 1C and S1C), suggesting that animals were able to remember and distinguish the two environments. Together, these data suggest that rats learned to generalize spatial rules across environments.

PFC spatial map generalizes, while CA1 completely remaps, across contexts

Since rats demonstrated transfer of learned knowledge from one environment to another, we leveraged this behavior to assess neural representations that support memory generalization. We recorded simultaneously from the PFC and the dorsal CA1 region of the hippocampus using high-density arrays with 32–64 tetrodes (Figure S2; STAR Methods). Cells were stably tracked across environments (mean ± SD = 30.5 ± 9.8 CA1 and 37.4 ± 19.3 PFC cells per session per animal; Figure S2). In agreement with previous hippocampal findings,[9–12] CA1 place cells exhibited extensive remapping in response to environment switches (Figures 2A–2C and S3). Notably, the distributions of CA1 spatial-representation similarity between two environments were not different from distributions obtained by shuffling cell identities in the same environment (Figure S3A), suggesting that CA1 place cells provide statistically independent ensemble activity in separate environments (i.e., global remapping).[11,12] Furthermore, the previous CA1 representation was reinstated upon reexposure to the same environment (Figures 2A–2C and S3). Therefore, the hippocampus maintains unique maps of individual environments, even when the animals perform the same spatial memory task.

In striking contrast with hippocampal global remapping, we found that the PFC spatial representation was stable across environments (Figures 2A and 2B). While there was representational drift of PFC cells even within the same environment[16] (Figure 2B), the PFC spatial-representation similarity between two different environments was just as large as with repeated testing in the same environment (Figures 2C–2E, S3A, and S3B), suggesting that novel and familiar environments shared the same prefrontal map. Furthermore, there was greater assimilation of PFC representations across environments with experience, while CA1 place cells maintained their independent representations (Figures S3C and S3D). Of note, the persistence of environment-specific representations in the hippocampus suggests that the animals did not ignore environmental cues when generalizing task knowledge, as it has been shown that inattentiveness to environmental cues changes spatial coding and reduces hippocampal place-field stability.[40,41]

Since the environments shared both the same rules (i.e., remembering correct task sequences) and the same reward configuration, we examined the degree of generalization attributed to coding of reward expectation and task sequences. Previous studies have shown that firing rates of some hippocampal[42–44] and prefrontal[45,46] neurons ramp up as the animal moves toward the reward, representing relative proximity of goals and thus reflecting reward expectation.[47–49] Consistent with previous reports, we identified ramping neurons in both CA1 and PFC (Figures S4A and S4B; STAR Methods). We found that PFC ramping cells generalized reward representations, but the CA1 representation remapped

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

across environments (Figures S4A and S4B). In addition, the generalization for reward ramping cells was seen to a similar degree to other cells within regions (Figure S4C). Finally, in both regions, the similarity of population activity across different environments at reward wells was the same as other locations (Figure S4D). Therefore, the generalization of PFC maps was not simply attributed to reward expectation.

To determine how fast ensemble codes shift upon environment switch, we next examined trial-to-trial dynamics of CA1 and PFC representations. We used Bayesian decoding to apply spatial representations of CA1 and PFC cells from the novel environment to reconstruct locations in the familiar and novel environments (STAR Methods). Confirming the remapping results, the error between actual and reconstructed positions in the familiar environment was significantly higher than that in the novel environment and no better than chance for CA1, whereas PFC decoding was well above chance and similar across environments (Figures 2B and 2E). Interestingly, the decoding error rapidly increased during the first trial after environment switching for CA1 but remained stable for PFC (Figure 2F; trial −1 vs. 1: $p = 0.038$ and > 0.99 for CA1 and PFC, respectively, Kruskal-Wallis test with Dunn's post hoc). Together, these data suggest that when animals performed the same memory task in different environments, unique CA1 representations appeared immediately after introduction to a novel environment, contributing to separation of spatial contexts in memory, whereas PFC representations carried over from one environment to the other, allowing generalization.

Distinct geometries of CA1 and PFC representations

To further investigate how CA1 and PFC neurons encode task knowledge, we examined their single-cell responses to different task variables. We found that PFC and CA1 single-cell codes had distinct generalization properties for multiple task variables beyond spatial locations. First, many PFC cells exhibited similar activity profiles at the same trajectory phase (defined by the distance from the starting point) across different trajectory types (Figure S5), termed "path equivalence."[50,51] Such path equivalence is relatively rare in CA1[52,53] (Figure S5). Furthermore, beyond simple "path equivalence" comprising one generalized representation of all trajectories as previously suggested,[50,52,53] we found that the firing patterns of PFC neurons specifically distinguished the two IN-OUT task sequences critical for task performance (IN_R-OUT_L vs. IN_L-OUT_R in Figure 1A; "task-sequence selectivity") and thus generalized the responses across the two IN and OUT trajectories covering distinct spatial locations within a task sequence (Figures 3B, 3C, and S5F). In contrast, most CA1 cells showed consistent firing for trajectories with similar spatial coverage (IN_R-OUT_R or IN_L-OUT_L; Figures 3A and 3C). As a result, the selectivity indices of IN (IN_L vs. IN_R) and OUT (OUT_L vs. OUT_R) trajectories exhibited a positive correlation for CA1 and an opposite, negative correlation for PFC (Figure 3D; STAR Methods). We noted, however, that while the two trajectories within a task sequence shared similar motor patterns (e.g., performing an L/R turn), the response of PFC cells was not a pure encoding of actions because spatial positions associated with similar actions can still be accurately decoded from PFC activity[17,22] (Figures 2B and 2E). Therefore, the PFC representation did not strictly encode spatial or motor information, per se,[29] but variable conjunctions relevant

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

to the task rule. Finally, consistent with spatial remapping results, these properties were stable across environments for PFC neurons but not CA1 (Figures 3E and S5E).

The examination of single-cell encoding profiles revealed extensive heterogeneity, with a gradient in selectivity and generalization properties in each region (Figures 2 and 3). To further gain insights into how CA1 and PFC circuits engaged individual neurons into functional ensembles to encode task knowledge, we characterized the structure of population activity by extracting neural manifolds using uniform manifold approximation and projection (UMAP; STAR Methods). This unsupervised algorithm revealed that the population activity of PFC cells resides on a ring-shaped manifold, making it distinct from the CA1 neural manifold with a double "V" shape that topologically resembles the physical spatial structure of the W-maze (Figures 3F and 3G, left).

To validate the manifold approach, we examined if the extracted geometric structure accurately reflected the diverse tuning properties of individual cells. If so, four key predictions should be true. First, locations along different trajectories should have distinct population neural states in the manifolds to support episodic encoding of locations in both CA1 and PFC, with the location-selective firing of individual neurons differentiating the context of specific trajectories (Figures 2B and 3E).[17,22,27,28,52,54] Figures 3F and 3G demonstrate that neural states of single passes of the same trajectory exhibited similar dynamics, yet maintained separation between those belonging to different trajectories, enabling trajectory-specific representations of locations. Second, the task-sequence selectivity of PFC cells (Figure 3D) predicts that the trajectories in a correct task sequence (InSeq; IN_R-OUT_L or IN_L-OUT_R) should reside in closer proximity than those out of the sequence (OutSeq; IN_R-OUT_R or IN_L-OUT_L). Indeed, the two InSeq trajectory pairs lie close together on opposite sides of the ring of the PFC manifold (Figure 3G). When measuring the distance between neural trajectories (STAR Methods), we found that the distance between two InSeq trajectories was significantly smaller than two OutSeq ones in the PFC manifold, whereas OutSeq pairs with same spatial coverages had smaller distances than InSeq pairs in the CA1 manifold (Figure 3H). Third, the same trajectory phase across the four different trajectory types should be geometrically aligned in the PFC manifold to encode "path equivalence." Indeed, for PFC, the same phases of different trajectories (e.g., the start points) were aligned, regardless of their physical distances (Figure 3G, left). As a result, the distances between PFC neural states for the same trajectory phases were significantly smaller than those for the same spatial locations, while the opposite is true for CA1 (Figure 3I). Fourth, the generalization of PFC spatial representations (Figure 2) predicts that the manifold structure should be preserved across environments. To test this, we embedded the population activity from the familiar environment into the manifold extracted from the activity in the novel environment (Figures 3F and 3G, right; STAR Methods). We found that the geometric structure of the PFC manifold was largely preserved across environments, whereas the distortion of the CA1 manifold was similar to that from random shuffles (Figure 3J). In all the cases above, the distance relationship of neural states, which determines the representational geometry, was preserved between the low-dimensional manifold space and the original high-dimensional neural space (Figure S6). Together, our data reveal that neural ensembles in CA1 and PFC simultaneously represent multiple task variables with distinct geometries that are consistent with single-neuron selectivity. More

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

importantly, the demonstration of population-level neural manifolds uncovers an internal structure that is not readily apparent from individual neuronal responses, elucidating how two seemingly opposing functions of the single-neuron codes, i.e., maintenance of episodic detail (e.g., location and trajectory specificity) and generalization across experiences, can operate in tandem at the level of neural populations.

Distinct abstraction in CA1 and PFC

Generalization is posited to arise from abstraction of commonalities across experiences.[3,32,35] Previous studies have operationally defined abstraction using the generalization performance of neural decoders across task conditions not used for training (cross-condition generalization performance [CCGP]), which requires a specific representational geometry.[35,55] Unlike traditional cross-validation decoders, trained on a subset of samples from all conditions and tested on held-out samples from the same conditions, CCGP evaluates a decoder's ability to reuse its existing trained structure to separate similar but novel inputs, allowing generalization.[35]

To assess if CA1 and PFC manifolds emerged as an abstraction of task variables, we further characterized their representational geometries using CCGP. To simplify decoding, we considered an early and a late phase for each trajectory in the neural manifolds (Figures 3F and 3G, left), resulting in 8 different clusters for each region (Figures 4A and 4B; STAR Methods). We then tested the CCGP for the three different dichotomies introduced in the previous section (i.e., early vs. late trajectory phase, task sequence IN_R-OUT_L vs. IN_L-OUT_R, and novel vs. familiar environment; Figure 4C, top to bottom). For example, to classify early and late trajectory phases (Figure 4C, top), a linear decoder was trained on 3 out of all 4 trajectory types and then tested on the one trajectory type not used for training. Similarly, we tested the CCGP of the task sequence by holding out a pair of clusters with the same trajectory phase but belonging to different task sequences; for the CCGP of spatial contexts, we tested the decoder trained by neural activity in the novel environment with the activity in the familiar environment (STAR Methods). We found that the CCGP for all three task variables showed high accuracy for PFC but was at or below chance for CA1 with a trial-label permutation test (Figures 4C and 4D), as well as against a geometric random model (STAR Methods; Figures S7A and S7B).[35] PFC representation thus has a geometric structure that enables a decoder to classify neural responses to these task variables in novel conditions (i.e., an abstract format). Notably, if we measured the amount of information decodable using the traditional cross-validation decoder for all 35 possible dichotomies (dividing the 8 clusters into 2 groups of 4 conditions),[35] with each dichotomy corresponding to a task variable that could be decoded, we found that most of the dichotomies were accurately separated in both CA1 and PFC, suggesting that both regions contain a similar amount of decodable variables (Figures S7C and S7D). Together, these data reveal that although CA1 and PFC representations shared similar information content of the current task, such content was formatted with distinct levels of abstraction in CA1 and PFC.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

DISCUSSION

Here, we have identified neural manifold signatures of memory generalization in the rat hippocampal-prefrontal circuitry using a spatial navigation task across environments with distinct sensory cues. Animals acquired the task significantly faster in the novel environment than during *de novo* learning (Figures 1 and S1) and thus exhibited behavior reflecting generalization of remembered task knowledge. During memory generalization, we found that the hippocampus remapped and conjunctively linked the task structure with current spatial context, whereas PFC did not remap but organized task knowledge in an abstract format, which is generalizable to novel spatial contexts (Figure 2) and to variable conjunctions relevant to the task rule (Figures 3 and 4).

We note that the generalizable PFC map observed here is broadly consistent with previous reports of its task-rule representation, where PFC neurons respond similarly to different sensory stimuli associated with the same behavioral outcome.[56–58] Our navigational task densely samples the space of task variables, rather than using a few pre-defined task states as in the previous studies,[56–58] and thus reveals the organized cognitive map instantiated by a neural manifold in PFC. Further, the distinct geometry of the prefrontal cognitive map provides a striking contrast to the exemplar cognitive map in the hippocampus for representing episodic details.[9] Finally, our results provide neurophysiological support to the prevailing theoretical framework regarding the complementary contributions of the hippocampus and PFC to memory generalization.[5–7] Of note, the complementary representations observed here have the capacity to retrieve learned task knowledge in new spatial contexts without losing the separability of different contexts. Proposed models with similar complementary architectures can optimize the trade-off between separability and generalizability in neural computations and thus achieve enhanced flexibility and reduced interference in artificial agents.[7,25,33,59]

The use of the W-maze task is motivated by causal evidence that an intact hippocampal-prefrontal network is necessary for optimally learning and performing the task.[37–39,60] In contrast, recent experiments using simple exploratory tasks, such as linear-track running and random foraging, showed that PFC remapped, reminiscent of the hippocampus.[16,19,30,61] Notably, this type of behavior without explicit memory requirements does not require the hippocampal-prefrontal network.[62,63] Together, these results demonstrate that while PFC responds to contextual information, the generalizability of the PFC cognitive map is task dependent. This dependency is further supported by previous evidence that PFC neurons changed their response to similar behaviors across two different tasks in the same environment.[29,64,65] Therefore, shared visual cues and motor patterns alone are insufficient to robustly evoke a generalizable representation of task space in PFC. This idea echoes previous studies, which found that cognitive demands and behavioral engagement can profoundly change hippocampal and prefrontal spatial representations.[29,40,41,66–68] Given that the hippocampus provides a major input to PFC[69,70] and that hippocampal lesions impair the PFC spatial code,[61] we postulate that the spatial responses in PFC may be passively invoked by hippocampal input when not actively engaged in the behavior, such as during spontaneous locomotion, whereas specific task requirements may necessitate further neural computations of generalizable task knowledge in PFC.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

In contrast to the distinct representational formats, we found that the representational content of current task information is largely shared across the hippocampus and PFC, consistent with previous reports of various response types found in both regions, including spatially tuned cells (Figure 2),[14,16,22,28] trajectory-dependent cells (Figures 2 and 3),[17,22,27,28,52,54] and path-equivalent cells (Figure S5).50–53 Our data also support recent studies demonstrating highly distributed encoding of task variables across diverse brain areas involved in flexible behavior,[71–74] posing the challenge of reconciling highly distributed encoding with specialized functions in distinct brain circuits. Here, we bridged this gap by revealing an important relationship between coding properties of individual neurons and population geometry in the hippocampal-prefrontal circuitry (Figures 3 and S6) and demonstrating how distinct representational geometries can be compatible with hippocampal and PFC specialized functions. Notably, the PFC representation achieved generalization across many task variables (Figure 4) yet did not discard their information (Figure S7). This coding mechanism can confer flexibility for readout of task variables (Figure S7), in agreement with studies in other species, including humans and non-human primates,[35,36,55] suggesting that it may be a conserved computational strategy.

Limitations of the study

The present study did not explore the relationship between the dynamics of geometric properties of neural manifolds (e.g., CCGP) and behavioral improvement due to the limited numbers of different trial types and neurons recorded. Recent developments in optical imaging and high-density electrophysiological recordings offer opportunities to address this question, and our study provides an ideal foundation to characterize the geometric structure of cognitive maps in such data. Also, the effective dimensionality of CA1 and PFC population activity remains to be determined. One common strategy to measure dimensionality is to apply linear dimensionality reduction methods, such as principal-component analysis (PCA), to neural population activity.[75,76] While such “embedding dimensionality” carries important information about manifold structures, it does not correspond to the “intrinsic dimensionality” (i.e., the minimal number of latent variables needed to parametrize the manifold).[77] Using a non-linear dimensionality reduction method, a recent study has demonstrated that the hippocampal manifold embedded in 4 dimensions of the non-linear space needed a much larger number of dimensions when applying PCA (~29 dimensions).[26] Therefore, accurately estimating the dimensions of CA1 and PFC manifolds remains a challenge. Finally, our results open questions not addressed here regarding the circuit mechanisms that govern transformation of hippocampal-PFC representations underlying memory generalization. The hippocampal-PFC interaction involves a broad network via multiple direct and indirect pathways,[69,70] and their functional connectivity is modulated by multiple network patterns, including theta oscillations and sharp-wave ripples.[18,78] Unraveling the roles of these pathways and network patterns in memory generalization is an important avenue for future studies.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

STAR★METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Shantanu P. Jadhav (shantanu@brandeis.edu).

Materials availability—This study did not generate new unique reagents.

Data and code availability

- Raw electrophysiological and rat behavioral data have been deposited on servers at Brandeis University, and are available upon request by contacting the lead contact.
- Custom code in MATLAB (R2017a) used for analyses in this study is available on GitHub at http://github.com/JadhavLab/GeometricTransformation.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Eight adult male Long-Evans rats (450–550 g, 4–6 months) were used in this study. Animals were individually housed and kept on a on a 12-hr regular light/dark cycle. All procedures were approved by the Institutional Animal Care and Use Committee at the Brandeis University and conformed to US National Institutes of Health guidelines.

METHOD DETAILS

Experimental design—After habituation to daily handling and resting in a high-walled, opaque sleep box (~30 × 30 cm), animals were food deprived to 85–90% of their *ad libitum* weight, and pre-trained to run on a linear track (~1-m long) for rewards (sweetened evaporated milk).[17,22] For experiments on the W-maze (~80 × 80 cm with ~7-cm wide tracks), animals were food-deprived and naïve to the W-mazes and the novel experimental room (Figure S1A). Animals were rewarded for performing a hippocampal and PFC-dependent continuous alternation task (Figure 1A, left)[37–39]: returning to the center well after visits to either side well (left or right well; inbound, IN_L or IN_R), and choosing the opposite side well from the previously visited side when starting from the center (outbound, OUT_R or OUT_L). Animals learned the W-track alternation rule by trial-and-error. Rewards were automatically delivered in the three reward wells (left well: L; right well: R; center well: C), triggered by crossing of an infrared beam by the animal's nose. Each W-maze session lasted 15–20 mins, and was interleaved with rest sessions in the sleep box placed near the W-maze (Figures 1A and S1A). The task performance was estimated using a state-space model (Figures 1 and S1).[39,84]

We collected behavioral and electrophysiological data from five rats. For these animals, *de novo* learning of the W-maze task occurred in one room (Figure S1A) with 8 W-maze sessions within a single day.[17,22] Three of the five rats were trained for 2 additional days

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in the same environment (i.e., familiar environment). On a subsequent day after the *de novo* learning, rats were first trained to perform the W-maze task in a novel room for one session (denoted as N). In addition to a different size and number of prominent visual cues, the novel room incorporated large cue cards with distinct patterns to increase the discriminability from the familiar room (Figure S1A). After this novel session, rats ran one session in the familiar environment (F), followed by another session in the novel environment (N'; Figure 1A).

To ensure that animals can remember and discriminate the two environments, we analyzed the behavioral data of another three rats from a previous experiment[23,85] (Figure 1C). These animals were well-trained for the W-maze task in one environment for 5 days. Subsequently, for two further days, they ran one session in the familiar environment first, and then one session in the novel environment each day, and thus experienced a test sequence as F-N-F'-N'.

Recording procedures—After linear-track pretraining, animals were implanted with a multi-tetrode drive containing 32–64 independently moveable tetrodes targeting dorsal hippocampal region CA1 (−3.6 mm AP and 2.2 mm ML) and medial PFC (+3.0 mm AP and 0.7 mm ML), both in the right hemisphere (for 32-tetrode drives) or bilaterally (for 64-tetrode drives). On the days following surgery, hippocampal tetrodes were gradually advanced to the desired depths with characteristic EEG patterns (sharp wave polarity, theta modulation) and neural firing patterns.[17,22,23,85] One tetrode in corpus callosum served as the hippocampal reference, and another tetrode in overlying cortical regions with no spiking signal served as the prefrontal reference. The reference tetrodes reported voltage relative to a ground (GND) screw installed in skull overlying cerebellum. The W-maze experiments started at ~21 days after implantation, and electrodes were not moved at least 4 hours before and during the experiments. Following the conclusion of the experiments, micro-lesions were made through each electrode tip to mark recording locations[17,22] (Figures S2A–S2F). PFC recordings were primarily in PreLimbic cortex with a smaller number in rostral Anterior Cingulate Cortex (Figures S2D–S2F).

Data were collected using a SpikeGadgets data acquisition system (SpikeGadgets LLC). Spike data were sampled at 30 kHz and bandpass filtered between 600 Hz and 6 kHz. LFPs were sampled at 1.5 kHz and bandpass filtered between 0.5 Hz and 400 Hz. The animal's position and running speed were recorded with an overhead color CCD camera (30 fps) and tracked by color LEDs affixed to the headstage. Single units were identified offline using a manual clustering method (MatClust, M. P. Karlsson).[17,22] Only well isolated neurons with stable spiking waveforms (Figures S2G–S2I), and at least 100 spikes in a given session were included in analyses. Putative interneurons were identified and excluded based on spike width and firing rate criterion.[17,22,23,85]

QUANTIFICATION AND STATISTICAL ANALYSIS

Spatial fields and linearization—Spatial fields were calculated only during locomotor periods (> 5 cm/s; all SWR times excluded) at positions with sufficient occupancy (> 20 ms). 2D occupancy-normalized rate maps (Figure 2A) was constructed using spike counts

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and occupancies with 2-cm spatial bins, and smoothed with a 2D Gaussian (σ = 8 cm). To construct the 1D linearized rate maps on the 4 trajectory types (IN_R, OUT_L, IN_L, OUT_R), animal's linear positions were first estimated by projecting its actual 2D positions onto pre-defined idealized paths along the track, and further classified as belonging to one of the four trajectory types.[17,22] The linearized rate maps were then calculated with 2-cm spatial bins of the linear positions, and smoothened with a 1D Gaussian (s = σ cm; Figure 2B). A peak rate ≥ 3 Hz across linear spatial bins was required for a cell to be considered as a spatially-tuned cell.

Quantification of single-cell remapping—Single-cell activity on the W-maze was compared across different environments. Rate remapping (i.e., change in firing rate) was quantified using "rate overlap", which was calculated by dividing the mean firing rate in the less active environment to the mean rate in the more active one[11] (Figure S3A, top). Global remapping (i.e., change in firing location) was quantified using "spatial similarity", as Pearson correlation (r) between linearized rate maps of two environments (Figure S3A, bottom). These values were also compared to the values expected from independent distributions, by randomly pairing the activity of a cell in the familiar environment to the activity of a different cell in the novel environment (i.e., cell ID shuffles).

Population vector similarity—To measure remapping in population activity, a population vector (PV) was constructed as the activity vector of all spatially-tuned cells in a certain linear spatial bin. The PV similarity was then defined as the Pearson correlation (r) between the PVs across all bins in two environments (Figures 2C, S3B, and S3C).

Ramping response identification—The ramping response was identified using a previous established method[49] (Figure S4). Session-averaged rate maps were considered as having ramping characteristics toward reward if they have a significantly linear regression coefficient (Pearson's, $r > 0.5$ and $p < 0.05$) over traversed positions to reward. We only included cells with a peak rate ≥ 3 Hz and located within 20 cm around the reward location.

Bayesian decoding—To evaluate neural representations at the ensemble level, Bayesian decoding was implemented.[17,22] A memoryless Bayesian decoder was built to estimate the probability of animals' position given the observed spikes (Bayesian reconstruction; or posterior probability matrix):

$$P(X, Tr \mid \mathbf{spikes}) = \frac{P(\mathbf{spikes} \mid X, Tr)P(X, Tr)}{P(\mathbf{spikes})}$$

where X is the set of all linear positions on the track for different trajectory types (i.e., Tr, $Tr \in \{IN_R, OUT_L, IN_L, OUT_R\}$), and we assumed a uniform prior probability over X and Tr. Assuming that all N cells fired independently and followed a Poisson process:

$$P(\mathbf{spikes} \mid X, Tr) = \prod_{i=1}^{N} \frac{(\tau f_i(X, Tr))^{spikes_i} e^{-\tau f_i(X, Tr)}}{spikes_i!}$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

where τ is the duration of the time window, $f_i(X, Tr)$ is the expected firing rate of the i-th cell as a function of sampled location X and trajectory type Tr, and spikes_i is the number of spikes of the i-th cell in a given time window. Therefore, the posterior probability matrix can be derived as follows:

$$P(X, Tr \mid \textbf{spikes}) = C\left(\prod_{i=1}^{N} f_i(X, Tr)^{spikes_i}\right)e^{-\tau\sum_{i=1}^{N} f_i(X, Tr)}$$

where C is a normalization constant such that $\sum_{k=1}^{4}\sum_{j=1}^{D} P(x_j, tr_k \mid \textbf{spikes} = 1$ (x_j is the j-th spatial bin, D is the total number of spatial bins, and tr_k is the k-th trajectory type).

Specifically, linearized rate maps of neural ensembles in the N' session were used as templates (i.e., $f(X, Tr)$). Spikes were calculated from the population activity in the N or F session binned into 200-ms bins (i.e., τ = 200 ms; moving window with 50 ms overlap), and was restricted to locomotor periods (> 5 cm/s; locations within 15 cm of the reward well were excluded for decoding to prevent contamination from SWR activity). For each time bin, the location with maximum decoded probability was compared to the actual position of the current trajectory that the animal was at in that bin (Figure 2B, bottom). Decoding error was then determined as the linear distance between estimated position and actual position (Figures 2E and 2F), and was compared to distributions from shuffling cell identities of rate maps in the N′ session (Figure 2E).

Trajectory selective index—To measure the trajectory selectivity of single cells, a trajectory selectivity index (SI) was calculated (Figures 3D and 3E) by comparing the mean firing rates on the Left- (or L-) vs. Right- (or R-) side trajectories for outbound (OUT, OUT_L vs. OUT_R) and inbound (IN, IN_L vs. IN_R), respectively:

$$SI = \frac{FR_L - FR_R}{FR_L + FR_R}$$

where FR_L is the mean firing rate on the L-side trajectory, and FR_R is for the R-side trajectory. Only cells that had at least one spatial field (peak rate ≥3 Hz) detected on either the L- or R-side trajectory were considered, and the rate maps in different sessions were analyzed separately.

Path-equivalent coefficient—Path equivalent coefficient of single cells was calculated using a previously established method.[50] We calculated the Pearson correlation of rate maps between two trajectory types. The median of the Pearson correlations across all possible pair combinations was reported as path-equivalent coefficient (R; Figure S5).

Characterization of neural manifold—In each region, all spatially-tuned units that were active in both environments were pooled across animals to build a pseudo-population of neurons (n = 98 CA1 and 171 PFC cells, respectively). For each trajectory type, the minimal number of trials (or passes) across animals starting from the first trial in a given session was used as the trial number. For each population, single-trial linearized rate map of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

each cell was calculated and binned at a resolution of 1 cm. Bins with spikes from less than 5 neurons were discarded. Uniform Manifold Approximation and Projection (UMAP)[80,81] was then run on these *n*-dimensional data to extract low-dimensional neural manifolds (Figures 3J and 3G). The hyperparameters for UMAP were: 'n_dims' = 3, 'metric' = 'cosine', 'n_neighbours' = 50, and 'min_dist' = 0.6, similar to previous studies.[86,87] To compare the neural manifolds between the novel and familiar environment, the UMAP transformation calculated for the N' session was re-applied to the population activity of the F session, by applying the fitted N' UMAP transformation as the 'template_file' to the 'run_umap' function in the UMAP MATLAB toolbox[81] (Figures 3F and 3G, right).[86]

The geometric relationship of CA1 and PFC representations was further estimated using Euclidean distance in this reduced manifold space (Figures 3H–3J), as well as the original *n*-dimensional state space (Figure S6). In both spaces, distance was computed identically, except for the number of dimensions used. The between-trajectory distance was computed between the single-trial neural trajectories from two different trajectory types (Figures 3H and S6A). For trajectory pairs belong to the same task sequence, the distance between their neural states at the same trajectory phase (defined by the distance from the starting point) was calculated; similarly, for trajectory pairs on the same spatial side of the W-maze, the distance between their neural states at the same spatial location was calculated (Figures 3I and S6B). Finally, the distance between F and N^0 representations was computed as the average distance of all pairwise distances of neural states from the same trajectory type and location in F and N' (Figures 3J and S6C). The chance level was estimated by randomly shuffling the cell identities of population activity in N' (Figures 3J and S6C).

Cross-condition generalization performance—The ability of neural manifolds to generalize to unseen experimental conditions was estimated using cross-condition generalization performance (CCGP).[35] For trajectory phase (Figure 4C, top), we first trained the decoders ($n = 4$) to classify the early and late phase of 3 out of 4 trajectory types (early and late phases defined as 20%-30% and 70%-80% trajectory length from the trajectory start, respectively; different phase values yielded similar results), and tested the trained decoders on the remaining trajectory type not used for training. Similarly, for task sequence (Figure 4C, middle), we first trained the decoders ($n = 4$) to classify the IN_R-OUT_L versus IN_L-OUT_R task sequences by holding out two clusters belonging to the same trajectory phase of two different OutSeq trajectories, and tested on the held-out conditions. Finally, these two types of decoders ($n = 8$) trained with neural activity in the N' session were tested with neural activity in the F session (Figure 4C, bottom).

The CCGP was further compared with traditional cross-validation decoders. For cross-validation decoders, we considered all 35 dichotomies that divide the 8 clusters in Figures 4A and 4B into 2 groups of 4 conditions.[35] For each dichotomy, the binary decoder was built using a 4-fold cross-validation, of which data samples of all conditions were split into 4 groups with one group as the test data (Figures S7C and S7D). For all the binary decoders, we used a C-Support Vector Machine (SVM) with a linear kernel through the libsvm library.[79] Significance was estimated using permutation of trial labels ($n = 1{,}000$ times).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

We also tested the CCGP against a geometric random model that destroys the geometrical structure but preserves within-condition noise structure in the data (Figures S7A and S7B), as in previous studies.[35,55] To achieve this, the noise clouds (the trial-by-trial firing rates of each condition) were rotated by permuting the axes of the firing rate space, and we repeated this procedure for each cluster separately (Figures S7A, left). We then computed the CCGP as described above on these moved clusters to obtain a null model CCGP value (Figures S7A, right), and repeated this 100 times to get a null distribution (Figures S7B).

Statistical analysis—Data analysis was performed using custom routines in MATLAB (MathWorks) and GraphPad Prism 9 (GraphPad Software). We used nonparametric and two-tailed tests for statistical comparisons throughout the paper, unless otherwise noted. We used ANOVA for multiple comparisons of Gaussian distributions, followed by a Tukey's test, when appropriate. For non-Gaussian distributions of multiple groups, we used Kruskal-Wallis test with *post hoc* analysis performed using a Dunn's test. $p < 0.05$ was considered the cutoff for statistical significance. Boxplots show median, 75th (box), and 90th (whiskers) percentile, unless indicated otherwise. All heat maps were obtained using perceptually uniform colormaps from MatPlotLib.[83,82] No statistical methods were used to pre-determine sample sizes, but our sample sizes are similar to those generally employed in the field.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank members of the Jadhav laboratory for feedback on earlier versions of this manuscript. This work was supported by NIH grant R01 MH112661 to S.P.J. Schematics in the graphical abstract are adapted with permission from SciDraw.io (CC-BY; DOIs: https://doi.org/10.5281/zenodo.3925923; https://doi.org/10.5281/zenodo.3926011).

REFERENCES

1. Squire LR (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. Psychol. Rev. 99, 195–231. 10.1037/0033-295X.99.3.582. [PubMed: 1594723]
2. Eichenbaum H, and Cohen NJ (2004). From Conditioning to Conscious Recollection: Memory Systems of the Brain (Oxford University Press).
3. Schlichting ML, and Preston AR (2015). Memory integration: neural mechanisms and implications for behavior. Curr. Opin. Behav. Sci. 1, 1–8. 10.1016/j.cobeha.2014.07.005. [PubMed: 25750931]
4. Tse D, Langston RF, Kakeyama M, Bethus I, Spooner PA, Wood ER, Witter MP, and Morris RGM (2007). Schemas and memory consolidation. Science 316, 76–82. 10.1126/science.1135935. [PubMed: 17412951]
5. Xu W, and Südhof TC (2013). A neural circuit for memory specificity and generalization. Science 339, 1290–1295. 10.1126/science.1229534. [PubMed: 23493706]
6. Kumaran D, Hassabis D, and McClelland JL (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. Trends Cognit. Sci. 20, 512–534. 10.1016/j.tics.2016.05.004. [PubMed: 27315762]
7. McClelland JL, McNaughton BL, and O'Reilly RC (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. 102, 419–457. 10.1037/0033-295X.102.3.419. [PubMed: 7624455]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

8. Tolman EC (1948). Cognitive maps in rats and men. Psychol. Rev. 55, 189–208. 10.1037/h0061626. [PubMed: 18870876]
9. O'Keefe J, and Nadel L (1978). The hippocampus as a Cognitive Map (Oxford university press).
10. Muller RU, and Kubie JL (1987). The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J. Neurosci. 7, 1951–1968. 10.1523/JNEUROSCI.07-07-01951.1987. [PubMed: 3612226]
11. Leutgeb S, Leutgeb JK, Treves A, Moser MB, and Moser EI (2004). Distinct ensemble codes in hippocampal areas CA3 and CA1. Science 305, 1295–1298. 10.1126/science.1100265. [PubMed: 15272123]
12. Alme CB, Miao C, Jezek K, Treves A, Moser EI, and Moser MB (2014). Place cells in the hippocampus: eleven maps for eleven rooms. Proc. Natl. Acad. Sci. USA 111, 18428–18435. 10.1073/pnas.1421056111. [PubMed: 25489089]
13. Mashhoori A, Hashemnia S, McNaughton BL, Euston DR, and Gruber AJ (2018). Rat anterior cingulate cortex recalls features of remote reward locations after disfavoured reinforcements. Elife 7, e29793. 10.7554/eLife.29793. [PubMed: 29664400]
14. Zielinski MC, Shin JD, and Jadhav SP (2019). Coherent coding of spatial position mediated by theta oscillations in the Hippocampus and prefrontal cortex. J. Neurosci. 39, 4550–4565. 10.1523/JNEUROSCI.0106-19.2019. [PubMed: 30940717]
15. Hasz BM, and Redish AD (2020). Spatial encoding in dorsomedial prefrontal cortex and hippocampus is related during deliberation. Hippocampus 30, 1194–1208. 10.1002/hipo.23250. [PubMed: 32809246]
16. Sauer JF, Folschweiller S, and Bartos M (2022). Topographically organized representation of space and context in the medial prefrontal cortex. Proc. Natl. Acad. Sci. USA 119, e2117300119. 10.1073/pnas.2117300119. [PubMed: 35121665]
17. Tang W, Shin JD, and Jadhav SP (2021). Multiple time-scales of decision-making in the hippocampus and prefrontal cortex. Elife 10, e66227. 10.7554/eLife.66227. [PubMed: 33683201]
18. Tang W, and Jadhav SP (2022). Multiple-timescale representations of space: linking memory to navigation. Annu. Rev. Neurosci. 45, 1–21. 10.1146/annurev-neuro-111020-084824. [PubMed: 34936810]
19. Bota A, Goto A, Tsukamoto S, Schmidt A, Wolf F, Luchetti A, Nakai J, Hirase H, and Hayashi Y (2021). Shared and unique properties of place cells in anterior cingulate cortex and hippocampus. Preprint at bioRxiv. 10.1101/2021.03.29.437441.
20. Remondes M, and Wilson MA (2015). Slow-gamma rhythms coordinate cingulate cortical responses to hippocampal sharp-wave ripples during wakefulness. Cell Rep. 13, 1327–1335. 10.1016/j.celrep.2015.10.005. [PubMed: 26549454]
21. Hok V, Save E, Lenck-Santini PP, and Poucet B (2005). Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. Proc. Natl. Acad. Sci. USA 102, 4602–4607. 10.1073/pnas.0407332102. [PubMed: 15761059]
22. Shin JD, Tang W, and Jadhav SP (2019). Dynamics of awake hippocampal-prefrontal replay for spatial learning and memory-guided decision making. Neuron 104, 1110–1125.e7. 10.1016/j.neuron.2019.09.012. [PubMed: 31677957]
23. Jadhav SP, Rothschild G, Roumis DK, and Frank LM (2016). Coordinated excitation and inhibition of prefrontal ensembles during awake hippocampal sharp-wave ripple events. Neuron 90, 113–127. 10.1016/j.neuron.2016.02.010. [PubMed: 26971950]
24. Manns JR, and Eichenbaum H (2006). Evolution of declarative memory. Hippocampus 16, 795–808. 10.1002/hipo.20205. [PubMed: 16881079]
25. Whittington JCR, Muller TH, Mark S, Chen G, Barry C, Burgess N, and Behrens TEJ (2020). The Tolman-Eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. Cell 183, 1249–1263.e23. 10.1016/j.cell.2020.10.024. [PubMed: 33181068]
26. Nieh EH, Schottdorf M, Freeman NW, Low RJ, Lewallen S, Koay SA, Pinto L, Gauthier JL, Brody CD, and Tank DW (2021). Geometry of abstract learned knowledge in the hippocampus. Nature 595, 80–84. 10.1038/s41586-021-03652-7. [PubMed: 34135512]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

27. Ito HT, Zhang SJ, Witter MP, Moser EI, and Moser MB (2015). A prefrontal-thalamo-hippocampal circuit for goal-directed spatial navigation. Nature 522, 50–55. 10.1038/nature14396. [PubMed: 26017312]
28. Fujisawa S, Amarasingham A, Harrison MT, and Buzsáki G (2008). Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. Nat. Neurosci. 11, 823–833. 10.1038/nn.2134. [PubMed: 18516033]
29. Jung MW, Qin Y, McNaughton BL, and Barnes CA (1998). Firing characteristics of deep layer neurons in prefrontal cortex in rats performing spatial working memory tasks. Cerebr. Cortex 8, 437–450. 10.1093/cercor/8.5.437.
30. Hyman JM, Ma L, Balaguer-Ballester E, Durstewitz D, and Seamans JK (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. Proc. Natl. Acad. Sci. USA 109, 5086–5091. 10.1073/pnas.1114415109. [PubMed: 22421138]
31. Duncan J (2001). An adaptive coding model of neural function in prefrontal cortex. Nat. Rev. Neurosci. 2, 820–829. 10.1038/35097575. [PubMed: 11715058]
32. Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, and Fusi S (2013). The importance of mixed selectivity in complex cognitive tasks. Nature 497, 585–590. 10.1038/nature12160. [PubMed: 23685452]
33. Fusi S, Miller EK, and Rigotti M (2016). Why neurons mix: high dimensionality for higher cognition. Curr. Opin. Neurobiol. 37, 66–74. 10.1016/j.conb.2016.01.010. [PubMed: 26851755]
34. Chung S, and Abbott LF (2021). Neural population geometry: an approach for understanding biological and artificial neural networks. Curr. Opin. Neurobiol. 70, 137–144. 10.1016/j.conb.2021.10.010. [PubMed: 34801787]
35. Bernardi S, Benna MK, Rigotti M, Munuera J, Fusi S, and Salzman CD (2020). The geometry of abstraction in the Hippocampus and prefrontal cortex. Cell 183, 954–967.e21. 10.1016/j.cell.2020.09.031. [PubMed: 33058757]
36. Fu Z, Beam D, Chung JM, Reed CM, Mamelak AN, Adolphs R, and Rutishauser U (2022). The geometry of domain-general performance monitoring in the human medial frontal cortex. Science 376, eabm9922. 10.1126/science.abm9922. [PubMed: 35511978]
37. Kim SM, and Frank LM (2009). Hippocampal lesions impair rapid learning of a continuous spatial alternation task. PLoS One 4, e5494. 10.1371/journal.pone.0005494. [PubMed: 19424438]
38. Maharjan DM, Dai YY, Glantz EH, and Jadhav SP (2018). Disruption of dorsal hippocampal - prefrontal interactions using chemogenetic inactivation impairs spatial learning. Neurobiol. Learn. Mem. 155, 351–360. 10.1016/j.nlm.2018.08.023. [PubMed: 30179661]
39. Jadhav SP, Kemere C, German PW, and Frank LM (2012). Awake hippocampal sharp-wave ripples support spatial memory. Science 336, 1454–1458. 10.1126/science.1217230. [PubMed: 22555434]
40. Kentros CG, Agnihotri NT, Streater S, Hawkins RD, and Kandel ER (2004). Increased attention to spatial context increases both place field stability and spatial memory. Neuron 42, 283–295. 10.1016/s0896-6273(04)00192-8. [PubMed: 15091343]
41. Fenton AA, Lytton WW, Barry JM, Lenck-Santini PP, Zinyuk LE, Kubı´k S, Bures J, Poucet B, Muller RU, and Olypher AV (2010). Attention-like modulation of hippocampus place cell discharge. J. Neurosci. 30, 4613–4625. 10.1523/JNEUROSCI.5576-09.2010. [PubMed: 20357112]
42. Jarzebowski P, Hay YA, Grewe BF, and Paulsen O (2022). Different encoding of reward location in dorsal and intermediate hippocampus. Curr. Biol. 32, 834–841.e5. 10.1016/j.cub.2021.12.024. [PubMed: 35016008]
43. Hok V, Lenck-Santini PP, Roux S, Save E, Muller RU, and Poucet B (2007). Goal-related activity in hippocampal place cells. J. Neurosci. 27, 472–482. [PubMed: 17234580]
44. Duvelle E´ ., Grieves RM, Hok V, Poucet B, Arleo A, Jeffery KJ, and Save E (2019). Insensitivity of place cells to the value of spatial goals in a two-choice flexible navigation task. J. Neurosci. 39, 2522–2541. 10.1523/JNEUROSCI.1578-18.2018. [PubMed: 30696727]
45. Ma L, Hyman JM, Phillips AG, and Seamans JK (2014). Tracking progress toward a goal in corticostriatal ensembles. J. Neurosci. 34, 2244–2253. 10.1523/JNEUROSCI.3834-13.2014. [PubMed: 24501363]
46. Shidara M, and Richmond BJ (2002). Anterior cingulate: single neuronal signals related to degree of reward expectancy. Science 296, 1709–1711. 10.1126/science.1069504. [PubMed: 12040201]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

47. Hikosaka O, Sakamoto M, and Usui S (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. J. Neurophysiol. 61, 814–832. 10.1152/jn.1989.61.4.814. [PubMed: 2723722]
48. Sosa M, and Giocomo LM (2021). Navigating for reward. Nat. Rev. Neurosci. 22, 472–487. 10.1038/s41583-021-00479-z. [PubMed: 34230644]
49. Howe MW, Tierney PL, Sandberg SG, Phillips PEM, and Graybiel AM (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature 500, 575–579. 10.1038/nature12475. [PubMed: 23913271]
50. Yu JY, Liu DF, Loback A, Grossrubatscher I, and Frank LM (2018). Specific hippocampal representations are linked to generalized cortical representations in memory. Nat. Commun. 9, 2209. 10.1038/s41467-018-04498-w. [PubMed: 29880860]
51. Rubin A, Sheintuch L, Brande-Eilat N, Pinchasof O, Rechavi Y, Geva N, and Ziv Y (2019). Revealing neural correlates of behavior without behavioral measurements. Nat. Commun. 10, 4745. 10.1038/s41467-019-12724-2. [PubMed: 31628322]
52. Frank LM, Brown EN, and Wilson M (2000). Trajectory encoding in the hippocampus and entorhinal cortex. Neuron 27, 169–178. 10.1016/s0896-6273(00)00018-0. [PubMed: 10939340]
53. Singer AC, Karlsson MP, Nathe AR, Carr MF, and Frank LM (2010). Experience-dependent development of coordinated hippocampal spatial activity representing the similarity of related locations. J. Neurosci. 30, 11586–11604. 10.1523/JNEUROSCI.0926-10.2010. [PubMed: 20810880]
54. Wood ER, Dudchenko PA, Robitsek RJ, and Eichenbaum H (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. Neuron 27, 623–633. 10.1016/s0896-6273(00)00071-4. [PubMed: 11055443]
55. Minxha J, Adolphs R, Fusi S, Mamelak AN, and Rutishauser U (2020). Flexible recruitment of memory-based choice representations by the human medial frontal cortex. Science 368, eaba3313. 10.1126/science.aba3313. [PubMed: 32586990]
56. Samborska V, Butler JL, Walton ME, Behrens TEJ, and Akam T (2022). Complementary task representations in hippocampus and prefrontal cortex for generalizing the structure of problems. Nat. Neurosci. 25, 1314–1326. 10.1038/s41593-022-01149-8. [PubMed: 36171429]
57. Reinert S, Hübener M, Bonhoeffer T, and Goltstein PM (2021). Mouse prefrontal cortex represents learned rules for categorization. Nature 593, 411–417. 10.1038/s41586-021-03452-z. [PubMed: 33883745]
58. Wallis JD, Anderson KC, and Miller EK (2001). Single neurons in prefrontal cortex encode abstract rules. Nature 411, 953–956. 10.1038/35082081. [PubMed: 11418860]
59. Badre D, Bhandari A, Keglovits H, and Kikumoto A (2021). The dimensionality of neural representations for control. Curr. Opin. Behav. Sci. 38, 20–28. 10.1016/j.cobeha.2020.07.002. [PubMed: 32864401]
60. Fernández-Ruiz A, Oliva A, Fermino de Oliveira E, Rocha-Almeida F, Tingley D, and Buzsáki G (2019). Long-duration hippocampal sharp wave ripples improve memory. Science 364, 1082–1086. 10.1126/science.aax0758. [PubMed: 31197012]
61. Esteves IM, Chang H, Neumann AR, Sun J, Mohajerani MH, and McNaughton BL (2021). Spatial information encoding across multiple neocortical regions depends on an intact Hippocampus. J. Neurosci. 41, 307–319. 10.1523/JNEUROSCI.1788-20.2020. [PubMed: 33203745]
62. Eichenbaum H(2017). The role of the hippocampus in navigation ismemory. J. Neurophysiol. 117, 1785–1796. 10.1152/jn.00005.2017. [PubMed: 28148640]
63. Preston AR, and Eichenbaum H (2013). Interplay of hippocampus and prefrontal cortex in memory. Curr. Biol. 23, R764–R773. 10.1016/j.cub.2013.05.041. [PubMed: 24028960]
64. Rich EL, and Shapiro M (2009). Rat prefrontal cortical neurons selectively code strategy switches. J. Neurosci. 29, 7208–7219. 10.1523/JNEUROSCI.6068-08.2009. [PubMed: 19494143]
65. Karlsson MP, Tervo DGR, and Karpova AY (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. Science 338, 135–139. 10.1126/science.1226518. [PubMed: 23042898]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

66. Aoki Y, Igata H, Ikegaya Y, and Sasaki T (2019). The integration of goal-directed signals onto spatial maps of hippocampal place cells. Cell Rep. 27, 1516–1527.e5. 10.1016/j.celrep.2019.04.002. [PubMed: 31042477]
67. Pettit NL, Yuan XC, and Harvey CD (2022). Hippocampal place codes are gated by behavioral engagement. Nat. Neurosci. 25, 561–566. 10.1038/s41593-022-01050-4. [PubMed: 35449355]
68. Markus EJ, Qin YL, Leonard B, Skaggs WE, McNaughton BL, and Barnes CA (1995). Interactions between location and task affect the spatial and directional firing of hippocampal neurons. J. Neurosci. 15, 7079–7094. 10.1523/jneurosci.15-11-07079.1995. [PubMed: 7472463]
69. Eichenbaum H (2017). Prefrontal-hippocampal interactions in episodic memory. Nat. Rev. Neurosci. 18, 547–558. 10.1038/nrn.2017.74. [PubMed: 28655882]
70. Robinson JC, and Brandon MP (2021). Skipping ahead: a circuit for representing the past, present, and future. Elife 10, e68795. 10.7554/eLife.68795. [PubMed: 34647521]
71. Stefanini F, Kushnir L, Jimenez JC, Jennings JH, Woods NI, Stuber GD, Kheirbek MA, Hen R, and Fusi S (2020). A distributed neural code in the dentate gyrus and in CA1. Neuron 107, 703–716.e4. 10.1016/j.neuron.2020.05.022. [PubMed: 32521223]
72. Tseng SY, Chettih SN, Arlt C, Barroso-Luque R, and Harvey CD (2022). Shared and specialized coding across posterior cortical areas for dynamic navigation decisions. Neuron 110, 2484–2502.e16. 10.1016/j.neuron.2022.05.012. [PubMed: 35679861]
73. Steinmetz NA, Zatka-Haas P, Carandini M, and Harris KD (2019). Distributed coding of choice, action and engagement across the mouse brain. Nature 576, 266–273. 10.1038/s41586-019-1787-x. [PubMed: 31776518]
74. Allen WE, Kauvar IV, Chen MZ, Richman EB, Yang SJ, Chan K, Gradinaru V, Deverman BE, Luo L, and Deisseroth K (2017). Global representations of goal-directed behavior in distinct cell types of mouse neocortex. Neuron 94, 891–907.e6. 10.1016/j.neuron.2017.04.017. [PubMed: 28521139]
75. Gao P, and Ganguli S (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. Curr. Opin. Neurobiol. 32, 148–155. 10.1016/j.conb.2015.04.003. [PubMed: 25932978]
76. Cunningham JP, and Yu BM (2014). Dimensionality reduction for large-scale neural recordings. Nat. Neurosci. 17, 1500–1509. 10.1038/nn.3776. [PubMed: 25151264]
77. Jazayeri M, and Ostojic S (2021). Interpreting neural computations by examining intrinsic and embedding dimensionality of neural activity. Curr. Opin. Neurobiol. 70, 113–120. 10.1016/j.conb.2021.08.002. [PubMed: 34537579]
78. Shin JD, and Jadhav SP (2016). Multiple modes of hippocampal-prefrontal interactions in memory-guided behavior. Curr. Opin. Neurobiol. 40, 161–169. 10.1016/j.conb.2016.07.015. [PubMed: 27543753]
79. Chang C-C, and Lin C-J (2011). ACM Trans. Intell. Syst. Technol. 2, 1–27. 10.1145/1961189.1961199.
80. McInnes L, Healy J, and Melville J (2018). Umap: uniform manifold approximation and projection for dimension reduction. Preprint at arXiv. 10.48550/arXiv.1802.03426.
81. Meehan C, Ebrahimian J, Moore W, and Meehan S (2022). Uniform Manifold Approximation and Projection (UMAP) (MATLAB Central File Exchange). https://www.mathworks.com/matlabcentral/fileexchange/71902.
82. Cobeldick S (2022). MatPlotLib Perceptually Uniform Colormaps (MATLAB Central File Exchange). https://www.mathworks.com/matlabcentral/fileexchange/62729-matplotlib-perceptually-uniform-colormaps.
83. Nuñez JR, Anderton CR, and Renslow RS (2018). Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data. PLoS One 13, e0199239. 10.1371/journal.pone.0199239. [PubMed: 30067751]
84. Smith AC, Frank LM, Wirth S, Yanike M, Hu D, Kubota Y, Graybiel AM, Suzuki WA, and Brown EN (2004). Dynamic analysis of learning in behavioral experiments. J. Neurosci. 24, 447–461. 10.1523/JNEUROSCI.2908-03.2004. [PubMed: 14724243]
85. Tang W, Shin JD, Frank LM, and Jadhav SP (2017). Hippocampal-prefrontal reactivation during learning is stronger in awake compared with sleep states. J. Neurosci. 37, 11789–11805. 10.1523/JNEUROSCI.2291-17.2017. [PubMed: 29089440]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

86. Gardner RJ, Hermansen E, Pachitariu M, Burak Y, Baas NA, Dunn BA, Moser MB, and Moser EI (2022). Toroidal topology of population activity in grid cells. Nature 602, 123–128. 10.1038/s41586-021-04268-7. [PubMed: 35022611]

87. Campbell MG, Attinger A, Ocko SA, Ganguli S, and Giocomo LM (2021). Distance-tuned neurons drive specialized path integration calculations in medial entorhinal cortex. Cell Rep. 36, 109669. 10.1016/j.celrep.2021.109669. [PubMed: 34496249]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- Rats generalize knowledge of a hippocampal-PFC-dependent task to novel environments
- As rules transfer across contexts, CA1 remaps, but PFC representations generalize
- CA1 and PFC neural manifolds encode spatial cognitive maps with distinct geometries
- PFC manifolds organize task knowledge in an abstract, generalizable format

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

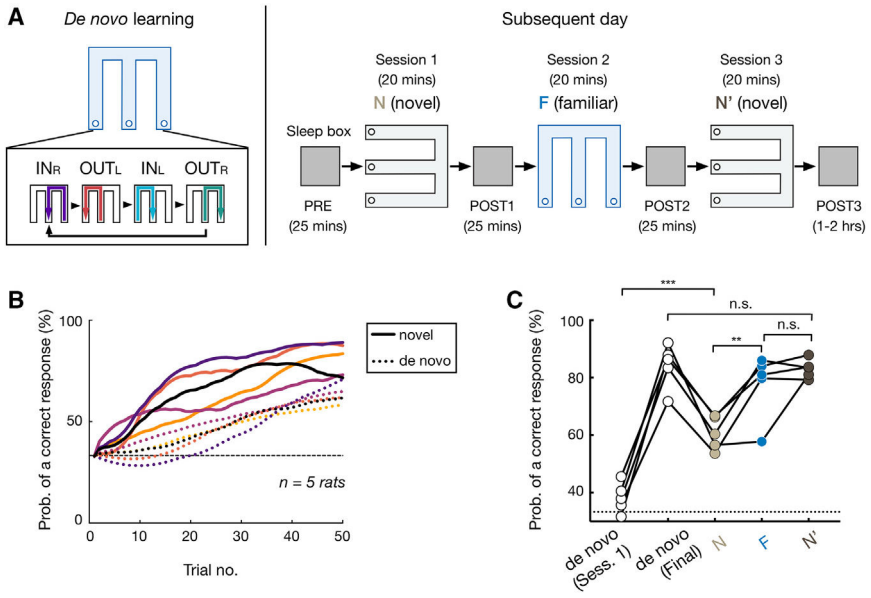


Figure 1. Prior knowledge facilitates learning in novel environment

(A) Experimental design.

(B and C) After *de novo* learning, rats (n = 5) acquired the task more rapidly in the novel environment (novel learning).

(B) Proportion correct for the first 50 trials during *de novo* vs. novel learning. Each pair of dashed (*de novo* learning) and solid (novel learning) lines represents an animal (color coded).

(C) Proportion correct against session number (data points for individual animals; repeated-measures ANOVA with Tukey's post hoc, $F = 40.48$, n.s., $p > 0.75$, ***p = 0.0002, **p = 0.0045). De novo (session 1), the first W-track session of *de novo* learning; de novo (final), final performance of *de novo* learning before N-F-N′ sessions. Horizontal dashed line: chance level.

See also Figure S1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

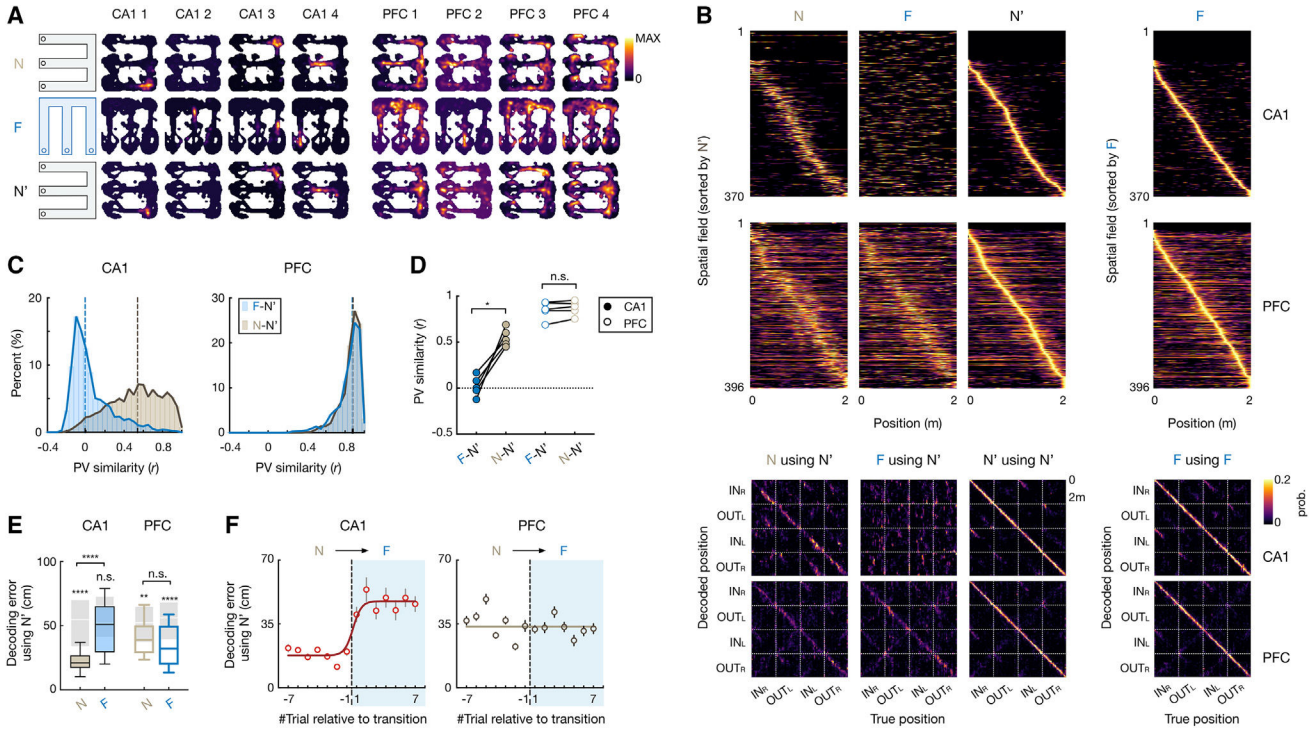


Figure 2. Hippocampus shows global remapping, whereas prefrontal spatial representations generalize across environments

(A) 2D rate maps of 4 CA1 and 4 PFC cells simultaneously recorded across environments (N, F, and N′).

(B) Top two rows: normalized rate maps on 2-m-long linearized trajectories of all spatial-tuned neurons tracked over environments, sorted by peak activity in N′ (left 3 columns) and in F (right). Bottom two rows: corresponding confusion matrices between true and decoded positions (with leave-one-out cross-validation; STAR Methods), where diagonal elements represent decoding accuracy.

(C and D) Population vector (PV) similarity across novel (N-N′; light gray) and familiar-novel (F-N′; blue) environments. Dashed vertical line on histogram: median. Each circle in (D) is for an animal (repeated-measures ANOVA with Tukey's post hoc, $F = 83.24$, n.s., $p = 0.41$, *$p = 0.01$).

(E) Decoding error using spatial maps from N′. Kruskal-Wallis test with Dunn's post hoc comparing N and F, n.s., $p > 0.99$, ****$p < 0.0001$. p values above each box were computed by comparing with its cell-ID shuffles (shaded boxes; n.s., $p = 0.14$, **$p = 0.0023$, ****$p < 0.0001$, Wilcoxon paired test). Box plots show median, 75th (box), and 90th (whiskers) percentile.

(F) Trial-by-trial decoding error during environment switch (N to F; median ± SEMs; line, sigmoid fit). Trial −1 vs. 1: p(CA1) = 0.038, and p(PFC) > 0.99, Kruskal-Wallis test with Dunn's post hoc.

See also Figures S2–S4.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

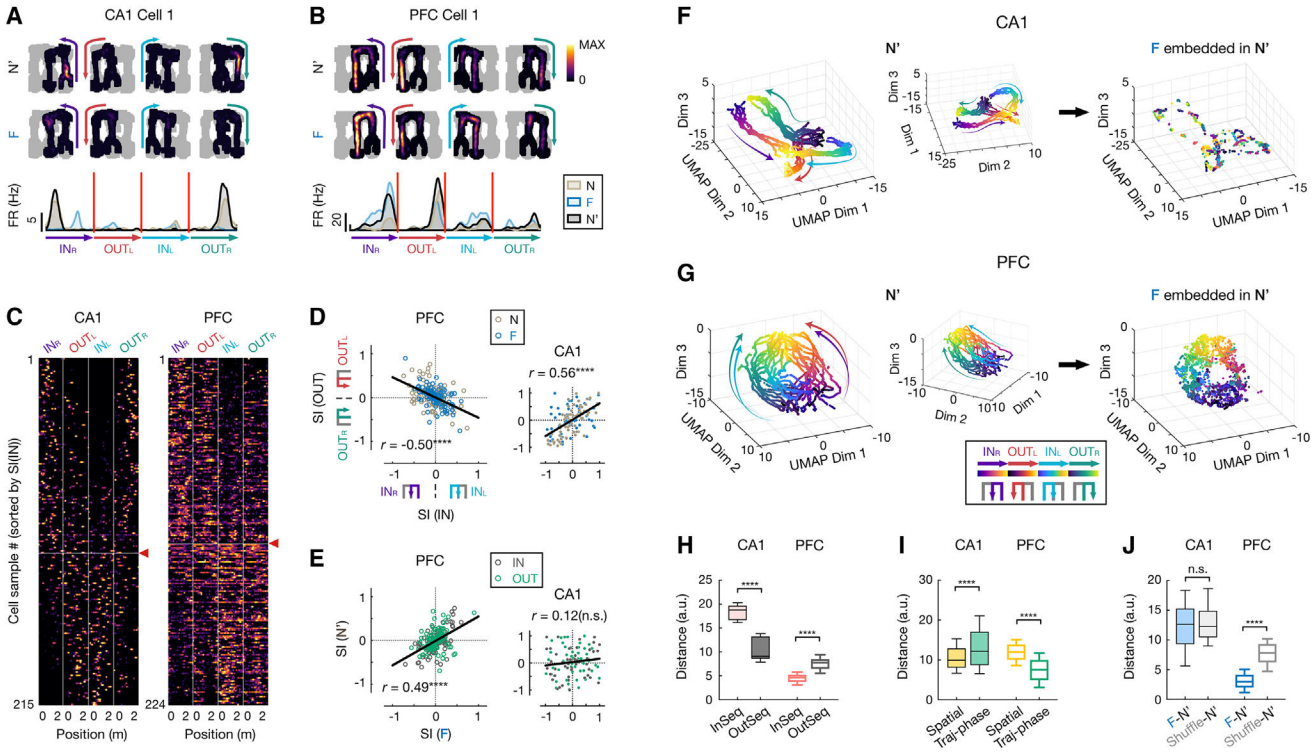


Figure 3. Distinct geometric representations of task space in CA1 and PFC

(A–E) CA1 and PFC neurons show distinct task-sequence selectivity.

(A and B) Firing fields on different trajectories of example (A) CA1 and (B) PFC cells. Top: 2D rate maps on each of 4 trajectories in the rewarded order (IN_R, OUT_L, IN_L, OUT_R, color coded) in N′ and F (all locations visited in a given session in gray). Bottom: corresponding rate maps on linearized trajectories (color-coded arrowheads indicate trajectory start and end).

(C) Linearized rate maps on the 4 trajectories of all spatially tuned neurons, sorted by selectivity index of inbound trajectories (i.e., SI(IN); IN_L vs. IN_R). The vertical lines separate each trajectory type. Horizontal line with a red arrowhead: division between cell samples with SI(IN) < 0 (above; selective to IN_R) vs.> 0 (below; selective to IN_L).

(D) Selectivity index of inbound against outbound trajectories (SI(IN) vs. SI(OUT); PFC, $r = -0.5$, ****p = 1.35e16; CA1, r = 0.56, ****p = 1.39e11; Pearson correlation).

(E) Selectivity index in familiar against novel environment (PFC, $r = 0.49$, ****p = 2.77e–14; CA1, $r = 0.12$, n.s., p = 0.22; Pearson correlation). Dark gray: SI(IN); green: SI(OUT). Each dot is a cell sample, and lines show least-squares linear fits in (D) and (E).

(F and G) 3D unsupervised embeddings of neural manifolds using UMAP in (F) CA1 and (G) PFC. Each dot represents the populational neural state (same 98 and 171 cells for CA1 and PFC, respectively) at one spatial position of a single trajectory pass (color coded by trajectory types and phases). Arrow-headed lines on manifolds schematically illustrate direction of motion along trajectory phases (from start to end). Left and middle: two different views of the manifold from the N′ session. Right: embeddings of population activity from F, guided by its template from N′ (STAR Methods).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(H) Distance between pair of trajectories in manifold. InSeq, pair of trajectories in a correct task sequence (IN_R and OUT_L or IN_L and OUT_R); OutSeq, pair of trajectories out of the sequence (IN_R and OUT_R or IN_L and OUT_L).

(I) Distance between pair of neural states at same physical spatial locations or trajectory phases (traj-phases) in manifold.

(J) Distance between F and N′ neural manifolds. Shuffle: cell IDs of N′ activity randomly shuffled.

Box plots in (H)–(J) show median, 75^{th} (box), and 90^{th} (whiskers) percentile (rank-sum tests, n.s., $p = 0.16$, ****$p < 0.0001$).

See also Figures S5 and S6.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

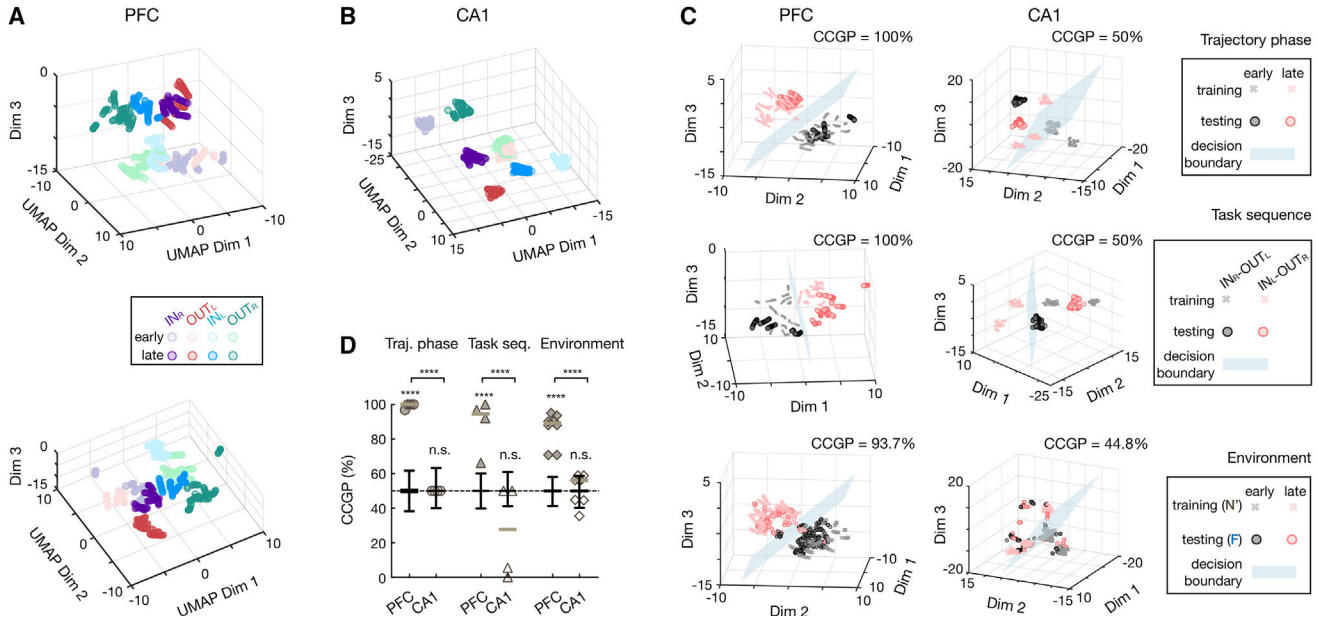


Figure 4. Distinct abstraction in CA1 and PFC representations

(A and B) UMAP embeddings of 8 clusters for early and late trajectory phases of 4 trajectory types (color coded) in (A) PFC and (B) CA1.

(A) Two different views of PFC clusters shown on top and bottom.

(C) CCGP of example decoding of dichotomies in PFC (left) and CA1 (right). From top to bottom: decoding for trajectory phases, task sequences, and environments, respectively (STAR Methods). A linear classification was implemented by a "decision hyperplane" (blue), trained on a subset of conditions in (A) and (B) (training set, crosses), and divided the space into two classes (red or black). CCGP (denoted on top) was then measured as how well this classifier generalize to an unseen set of conditions (testing set, circles; STAR Methods). Note the good performance on training sets but poor generalization on testing sets for CA1 decoding in contrast to high CCGP for PFC.

(D) CCGP for trajectory phases (circles), task sequences (triangles), and environments (diamonds; horizontal lines, median). Early, early trajectory phase; late, late trajectory phase. Note that all dichotomies can be accurately decoded in PFC (****$p < 0.001$) but not in CA1 (n.s., $p > 0.05$, trial-label permutation test). Boxplots are for trial-label shuffles ($n = 1{,}000$ times) and show median, interquartile range (box), and min-max (whiskers). See also Figure S7.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Cresyl Violet	Acros Organics	Cat#: AC229630050
Formaldehyde	Fisher	Cat#: 50–00–0,67561,7732–18–5
Isoflurane	Patterson Veterinary	Cat#: 07–806–3204
Ketamine	Patterson Veterinary	Cat#: 07–803–6637
Xylazine	Patterson Veterinary	Cat#: 07–808–1947
Atropine	Patterson Veterinary	Cat#: 07–869–6061
Bupivacaine	Patterson Veterinary	Cat#: 07–890–4881
Beuthanasia-D	Patterson Veterinary	Cat#: 07–807-3963
Sucrose	Sigma-Aldrich	Cat#: S8501–5KG
Experimental models: Organisms/strains		
Rat: Long Evans	Charles River	Cat#: Crl:LE 006; RRID: RGD_2308852
Software and algorithms		
MATLAB 2017a	Mathworks, MA	RRID: SCR_001622
Trodes	SpikeGadgets	http://www.spikegadgets.com
Matclust	Mattias P. Karlsson	https://www.mathworks.com/matlabcentral/fileexchange/39663-matclust, V1.7
Libsvm	Chang and Lin 2011[79]	https://www.csie.ntu.edu.tw/~cjlin/libsvm/, V3.12
Uniform Manifold Approximation and Projection (UMAP)	McInnes et al., 2018[80]; Meehan et al., 2022[81]	https://www.mathworks.com/matlabcentral/fileexchange/71902-uniform-manifold-approximation-and-projection-umap, V4.2.1
MatPlotLib	Cobeldick, 2022[82]; Nunez et al., 2018[83]	https://www.mathworks.com/matlabcentral/fileexchange/62729-matplotlib-perceptually-uniform-colormaps, V2.1.3
Prism 9	GraphPad Software	RRID: SCR_002798
Other		
128 Channel electrophysiology data acquisition system	SpikeGadgets	http://www.spikegadgets.com
12.7 μm NiCr tetrode wire	Sandvik	Cat#: PX000004