

Construction of a hippocampal cognitive map depends upon spatial context

Wan-chen Jiang^{1,*}, Shengjin Xu¹ & Joshua T. Dudman^{1,*}

¹ Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147

* correspondence: dudmanj@janelia.hhmi.org; jiangw@janelia.hhmi.org

Abstract

The hippocampus encodes both spatial and non-spatial features of an environment thought to be critical to guide navigational trajectories and associative learning, respectively. These seemingly dichotomous roles have been reconciled in a proposed cognitive map - a representation of environment structure abstracted away from specific behavioral demands. However, the extent to which a cognitive map is independent of behavioral demands remains unclear because direct comparisons across environments with common structure but different spatial/behavioral context is lacking. Here we compare behaviors in mice trained to navigate to a hidden target in a physical arena and manipulate a joystick to a virtual target to collect a delayed reward. Comparison of behaviors with an artificial agent revealed a common algorithmic basis for learned foraging trajectories in both contexts. Imaging CA1 neural activity revealed a similar map-like encoding of active foraging; however, detailed analysis of ensemble activity, optogenetic inactivation, and modeling revealed a context-specific functional dissociation. In a navigational context, CA1 was critical for retrospective evaluation of spatial trajectories, but dispensable for initiation. In a non-navigational context, CA1 activity was critical for initiation and planning foraging trajectories. This work highlights how construction of a cognitive map to facilitate idiosyncratic behavioral demands⁷ is critical for foraging in diverse spatial contexts.

Main text

Recent proposals have argued that the hippocampus produces a cognitive map that provides an internal model of the relational structure of a task as experienced by a subject^{1,4-6,9-11}. However, there are diverse ways in which a cognitive map may be constructed by an animal, a point articulated by Tolman⁷. As an example, our memory of a local grocery store can be in the form of a spatial map-like representations ("At 74th and Broadway") or a representation of a navigational trajectory ("Walk up to the corner and then head south for 6 blocks") or non-navigational, relational representation ("I got this jar of peanut butter there on a recent trip"). Each of these representations could be considered as different forms of cognitive maps; and there is evidence for spatial⁸, navigational^{1,12}, and relational¹³⁻¹⁵ maps, respectively. One can envisage how each of these representations could be preferred depending upon behavioral demands, efficiency, or other constraints. To date it is unclear whether hippocampal representations in navigational and non-navigational contexts are common forms of map-like representations or whether they may be specifically tailored to the unique behavioral demands of given task contexts^{3,7}. A major barrier to addressing this question is that quantitative measurements of neural activity, behavior, and targeted manipulations of hippocampal activity

have not been directly compared in animals trained in distinct task contexts (e.g., navigational and non-navigational) but analogous task structures.

Neural activity in the dorsal hippocampus, which provides rich and flexible representations of recent experience, is critical for learned spatiotemporal associations between locations, stimuli, and outcomes³. Activity in the dorsal CA1 (dCA1) region is organized at multiple spatiotemporal scales. Individual principal cells in dCA1 tend to be active in circumscribed regions of space (“place cells” and “place fields”)² with a broadly distributed propensity for activation that allows for efficient representation of spaces differing by orders of magnitude¹⁶. On a timescale of seconds, ensembles of place cells are organized into brief sequential bouts of activation as an animal actively locomotes through an environment¹⁷. When animals are not actively moving^{18,19} or asleep^{20,21}, large ensembles of dCA1 neurons burst in brief synchronous population events (SPEs) that occur for durations on the order of 150 milliseconds and tend to be associated with sharp wave ripples (SWRs) in the local field potential^{22,23}. Perturbation experiments have implicated each of these components of the dCA1 spatiotemporal representation in aspects of learned behavior^{24–27}. Although SWRs have been associated with functions ranging from memory recall²⁸, to consolidation²⁶, to decision-making²⁹, and are thought to reflect the underlying cognitive map of a task¹¹, there has not been a detailed comparison of these aspects of spatiotemporal representation in CA1 across foraging behaviors in navigational and non-navigational contexts.

To address these questions, we developed two tasks in which the learning problem was the same - initiate foraging trajectories that intersect a hidden target - yet the navigational context differed. First, we described a self-paced foraging task in which freely moving mice ran to an unmarked target location tens of centimeters away (spatial target foraging [STF] task; Fig. 1a; Supplementary video 1). Second, we modified a head-fixed operant task³⁰ in which a head-fixed mouse displaced a spring-loaded joystick to a target area on the order of ten millimeters away from its central resting location (Non-navigational target foraging [NTF] task; Fig. 1b). (Note: here we use the adjective ‘non-navigational’ as it has often been applied to operant tasks, e.g.³¹). In both tasks the delivery of reward was dissociated from movement into the target location that triggered reward. In the STF task, this was accomplished by delivering a reward via a water port at a specific ‘home’ location on a wall ~30 cm away from the target location. In contrast, in the NTF task, this was accomplished by delivering water reward with a 1 second delay after movement to the target area (forelimb movement duration ~ 0.5 second³²).

As the target location shifted, mice reliably adjusted their movements to collect rewards in both tasks (Fig. 1). For example, changes in the target location that triggered delayed reward delivery were primarily accompanied by rapid (<10 trials) reduction in errors made per trial (Fig. 1d,i) as the magnitude of foraging runs was adjusted (Fig. 1e,h). Trial initiation was uncued (‘self-paced’); however, in both tasks there was an unsignaled intertrial interval (ITI) in which no reward could be triggered. Mice tended to elicit movements at consistent intervals that were matched to the refractory ITI (Fig. 1f,j) - suggesting that mice were adapting behavioral timing to the inferred task structure.

The STF and NTF task variants are clearly similar - forage for a hidden target location that triggers reward delivery. Were mice solving the tasks in a similar manner? We noted two key features of behavior that indicate a potential common solution. One, changes in target position resulted in an increase in errors followed by rapid performance improvement with very similar time courses in both variants (Fig. 1k). Two, mice exhibited consistent trajectories that appear primarily scaled to match target location and not random exploration for a specific location (Fig. 1e,h).

A challenge in understanding foraging behavior is that many common models of learning agents differ from behavior either or both with respect to learning rates and to observed navigational trajectories³³. The primary limitation that makes reinforcement learning (RL) agents slow to learn is the 'curse of dimensionality'^{13,34}. In the case of spatial navigation, a typical RL agent uses a random search policy to sample all³⁴ or a sufficient number⁹ of discrete states (locations) that tile space to update the optimal action to take in every state (Extended Data Fig. 1). To explore alternative formulations, we took inspiration from the non-navigational NTF task variant. We have previously found that common RL model formulations fail to provide a good match to behavioral learning of continuous kinematic variables³². However, in that work we described an alternative learning rule ('MeSH') that was successful. MeSH avoids the curse of dimensionality by learning in the low-dimensional space of a generative parameter of movement as opposed to the high dimensional space of all movement parameters. This approach successfully explained the relatively rapid time course of learning during closed-loop optogenetic feedback³², the bidirectional learning in a task very similar to the NTF task here³⁵, and admitted to a plausible biological implementation³². Nonetheless, it was unclear whether or how this approach could be modified to model foraging trajectories, especially navigational trajectories.

Here we develop an artificial agent that can closely match observed properties of foraging trajectories and the time course of behavioral adaptation in both the non-navigational and navigational task contexts (trajectory MeSH Learning, 'tML' for simplicity; Fig. 2; see Methods). The tML agent modifies two key parameters (speed scaling and heading angle) governing generation of foraging trajectories so as to learn trajectories that successfully intersect the target location (Fig. 2a-c); a form of direct policy learning³⁶. Instead of attempting to learn a target location in a high dimensional spatial map (e.g. as in Q learning³⁴), tML learns a low dimensional set of parameters from which optimal trajectories are generated. A tML agent was capable of learning to intercept the target and could rapidly adapt to changes in target position within a few trials (Fig. 2d,e) analogous to our behavioral observations (Fig. 1c,d). Optimizing a forward learning rate and a behavioral variation parameter was capable of producing very good agreement with behavioral data (Fig. 2f) and much better than predictions derived from learning target location *per se* (Extended Data Fig. 1).

The NTF task used targets that were primarily constrained along a single dimension (amplitude) and included bidirectional changes in target distance. The tML model could also accurately simulate the joystick trajectories in the NTF task (Fig. 2g). Again the agent adjusted to both

increases and decreases in target distance within a few trials (Fig. 2h) analogous to the performance of mice trained on the same task (Fig. 1g-i). By optimizing for exploration and learning rate parameters, it was again possible to very closely match average learning behavior around target switches (Fig. 2i). These results revealed that behavioral performance of the NTF and STF tasks could be well described with a single, common RL model indicating that mice could be using the same algorithmic approach to refine foraging trajectories across the navigational and non-navigational contexts.

We next examined neural activity during performance of these tasks using epifluorescence imaging of dCA1 principal neuron activity with genetically-encoded calcium indicators (GCaMP6f) and a head-mounted miniscope^{37,38} (see Methods). Many dCA1 neurons exhibit clear place fields within circumscribed regions of their enclosure. When an animal navigates along a reliable trajectory, this leads to a sequential activation of place cells^{17,39,40} that depends upon the spatial context¹². We observed dCA1 place field activity distributed along the foraging trajectories (Extended Data Fig. 2). The reliable trajectories taken in the STF task (Fig. 1) resulted in reliable, sequentially ordered activity across a population of dCA1 principal neurons (n=5133 ROIs x sessions; Fig. 3a-e) as revealed in an alignment of peri-movement time histograms (PMTH) of activity (Fig. 3f) and apparent in individual trials (Fig. 3g). Activity of individual dCA1 neurons could be stable across days (Extended Data Fig. 3). While not specifically analysed here, dCA1 activity underwent some slow remapping consistent with previous observations in foraging tasks in environments of a similar size⁴¹.

Although the head-fixed NTF task has a similar conceptual structure, and operant tasks with delayed reward (“trace conditioning”) depend upon hippocampal function^{42,43}, less is known about hippocampal representations in non-navigational task contexts⁴⁴. Previous work has demonstrated the presence of ‘time cells’ that show sequential activation over elapsed time in the absence of substantial changes in position (or changing sensory input)⁴⁵⁻⁴⁷; however, in many cases these observations were made during active locomotion (running on stationary wheels^{46,47}). Recent work has also found that activity can tile a cognitive space (not time *per se*) in which a changing sensory stimulus evolves over several seconds¹⁰. Indeed, dCA1 activity is often locked (in a context-dependent manner) to sensory stimuli and this dependence upon changing sensory input is accentuated when head-fixed mice navigate in virtual reality^{22,48-51}.

We observed a number of analogous properties in CA1 ensemble activity during the STF and NTF tasks. For example, in the head-fixed NTF task, the activity of dCA1 cells exhibited a similar heavy-tailed distribution^{16,52} although the calcium-event rate was notably reduced relative to the freely moving task (Fig. 3c,j ; $p \ll 0.001$; ranksum test). The duration and peak rates of active cells were very similar (Fig. 3e,l). We next examined the responses of individual dCA1 neurons aligned to when the forelimb movement triggered the water reward. A cross-validated alignment revealed robust time-locked responses in a subset of cells and sequential activation of hippocampal neuronal ensemble (Fig. 3f,m). Similar to the freely-moving STF task, sequential activity was robustly observed in single trial activity during task performance (Fig. 3g,n). However, stable activation sequences were enriched during active

behavioral phases (e.g. movement of the forelimb) and little reliable activation patterns at the individual cell level were apparent in the intertrial interval⁵³.

To evaluate whether the trial by trial consistency of individual dCA1 neuron responses was related to movement we calculated the reliability of a response (ratio of mean to standard deviation; see Methods) as a function of the instantaneous velocity of the body (freely moving STF task; Fig. 3d) or the limb (head-fixed NTF task; Fig. 3k). In both cases we found that the reliability of dCA1 responses was significantly correlated with movement (STF task; Fig. 3d; Pearson's $r=0.21$, $p \ll 0.001$; NTF task; Fig. 3k; Pearson's $r=0.41$, $p \ll 0.001$). Although not possible to resolve in calcium imaging due to low temporal resolution, theta frequency modulation of forebrain local field potential was associated with performance of the NTF task in previously reported dorsal striatal recordings³⁰ (Extended Data Fig. 4) and has been associated with related non-navigational behavior in hippocampus⁵⁴. This suggests that forelimb movements reflect an active state distributed across multiple forebrain circuits⁵⁵ similar to locomotion in freely-moving animals⁵⁶.

As noted above, sequential activations of dCA1 ensembles are apparent as sequential patterns of activity observed on individual trials in both tasks (Fig. 3g,n). Previous work in freely moving navigation tasks has established that these activation sequences are correlated to the specific trajectories through space. There is substantial evidence that in navigational foraging behaviors dCA1 place cells tile the environment and thereby allow for an accurate encoding of allocentric position along a trajectory. However, it is unclear whether the sequential patterns observed during head-fixed forelimb movements encode 2D limb position. To address this question, we trained continuous-time linear decoders to examine whether trajectories could be reliably decoded in both task contexts (see Methods). We found that foraging run trajectories could be much better decoded than forelimb movements from dCA1 (Fig. 4a-c) exemplified in activation of neurons with place fields along the run trajectory during navigational foraging (Extended Data Fig. 2).

We can envision learning about different target locations spanning a range of possibilities between two extremes. On the one hand, if the trajectories themselves are encoded then the intercepted target location can be inferred - a spatial map. On the other hand, it is also possible to encode a heuristic corresponding to the target identity (e.g. target A vs. target B) - a relational or episodic-like map. The decoding above suggests a weak (but significant in a subset of sessions) metric relationship between dCA1 activity and the changing amplitude of forelimb movement trajectories. This implies that dCA1 activity sequences may encode task information in the non-navigational context, but not an encoding of behavioral trajectories *per se*⁵³. For example, a partially unique ensemble could be specifically recruited for trials to each target location through remapping⁵⁷ even though trial by trial fluctuations in the ensemble activity are independent of changes in trajectory.

We next examined how the ensemble patterns of activity were related to changes in target position. We first asked whether dCA1 ensemble activity could be used to differentiate target

locations by training ensemble classifiers (see Methods) to decode the location from dCA1 ensemble activity. We found that boosted tree ensemble classifiers (but not simpler decoders) achieve robust performance - correctly decoding the target location for $98 \pm 2\%$ s.d. and $77 \pm 10\%$ of trials on cross validation in the freely moving (STF) and head-fixed (NTF) tasks, respectively (Fig. 4d).

The above data confirm that the sequential recruitment of dCA1 neurons during active behavior is characteristic in navigational and non-navigational task contexts, which is consistent with previous observations in non-navigational tasks^{10,14} or tasks with non-navigational epochs^{45,46}. At the same time, dCA1 neurons recruited during active behavior provided more information encoding navigational trajectories than non-navigational forelimb movement trajectories (Fig. 4b). In contrast, we found little evidence for reliable, sequential recruitment of active dCA1 neurons in the absence of ongoing, active behavior (Fig. 3) similar to recent results in the context of conditioning with long trace intervals⁵³. This raised the question about whether there was additional information in CA1 activity during periods of behavioral inactivity.

Activity in dCA1 during inactivity has often been associated with the recall of information²⁸ and appears to be critical for learning associations between non-contiguous stimuli^{42,58}. Alert, but inactive, periods in rodents are characterized by the presence of SWRs, especially during consumption^{23,29}. Several lines of evidence suggest that SPEs observed in calcium imaging experiments correspond to SWR associated population events observed in electrophysiological recordings^{22,59}. SWR events are correlated with prospective planning of future foraging trajectories⁴⁰ and replay of experience in rodent studies^{18,25}. Perturbation of SWR offline or during task execution can produce learning deficits in navigational foraging tasks in rodents²⁴⁻²⁷. The probability of observing SWR events is correlated with recall of remembered (non-navigational) items in human recordings^{11,28}; however, to date it is unclear whether such events can be observed in animals performing non-navigational tasks.

Here we define SPEs as near-simultaneous activation of roughly 10% of the imaged population (within ~200 ms, see Methods). Using a conservative detection approach we observed SPEs at a rate of roughly 1 Hz throughout all of our imaging datasets (Fig. 5a,f; Ended Data Fig. 5; see Methods for details) consistent with previous observations using imaging^{22,59} or inferred replay events without single unit resolution¹¹. In both tasks, and analogous to previous observations with imaging⁵⁹, we observed multiple clusters of dCA1 ensembles active during SPEs. On average we observed similar numbers of SPE clusters per session (silhouette method): 3.6 ± 0.98 and 3.7 ± 1.92 (s.d.) in the STF and NTF tasks, respectively.

SPEs occurred largely during the intertrial intervals in both tasks (Fig. 5b,d,g,i); however, there was a clear difference in the timing of SPEs across spatial contexts. In navigational tasks SWR-associated replay events detected with electrophysiological recordings have generally been observed during reward consumption at the end of a trial¹⁸. Similarly, in the STF foraging task we observed SPEs primarily at the termination of a foraging run as the mouse returned to the reward collection area (Fig. 5b,c). In contrast, in the non-navigational NTF task SPEs were

primarily observed near the end of the intertrial interval just *prior* to initiation of a trial (Fig. 5g,h). The probability of observing SPEs significantly correlated with the quality of task performance in the non-navigational task but not the STF (Fig. 5e,j). We noted that the duration of intertrial intervals were much shorter in this STF task than often used when studying SWRs in navigational tasks (~2 seconds vs. ~30 seconds); perhaps as a result there was not sufficient time to observe multiple SPEs at their standard low (<1Hz) rate^{18,19,25,60}. The relative timing of SPEs depending upon task context suggest that in the STF foraging task they might be related to evaluation of successful performance (reward delivery) whereas in the NTF task they appear prospective and are predictive of performance⁶¹. Consistent with a potential evaluative role of SPEs in the STF task, the probability of observing an SPE was enhanced on correct trials that led to reward delivery compared with incorrect trials (0.77 ± 0.18 vs 0.49 ± 0.08 (s.d.); $p \ll 0.001$; ranksum test).

While previous work using lesions and pharmacological inactivation have demonstrated key roles for dCA1 in non-navigational tasks^{53,62–65}, it remains unclear whether and how representations during active performance (place cell activation) and/or during resting periods (associated with SPEs) of the task are individually or collectively critical for behavioral performance. We observed a clear dissociation of SPE timing across task variants suggestive that dCA1 may make distinct functional contributions to behavior depending upon context. Thus, we next sought to compare the contributions of dCA1 activity at multiple phases of task performance. We took advantage of a mouse line that expresses the optogenetic activator channelrhodopsin-2 (ChR2) in inhibitory neurons (using VGAT-ChR2-EYFP mice; see Methods) previously demonstrated to suppress activity of principal cells in dCA1 at modest light intensities that mitigate against effects on underlying areas^{66,67}. We then designed perturbations targeting both intertrial intervals (associated with SPEs) and following self-initiation of a trial during execution of movement trajectories (associated with place cell activation) on random subsets (~30%) of trials in separate sessions.

We used real time tracking of mouse position in the STF task to trigger inactivation at distinct trial epochs, initially targeting the period of occupancy in the reward collection area when SPEs were observed (Fig. 6a1). We found that such inactivation had no effect on general performance variables such as the time to initiate the next trial nor performance success. Our analyses above indicated that dCA1 (SPEs in particular) could be consistent with an ‘evaluative’ function. In the context of the tML agent model an evaluative role would be a role in updating the parameters of the future trajectory. We next examined the two key parameters of future trajectories - initial heading and amplitude - and found that there were indeed systematic changes in the variance in the initial heading and total distance on trials immediately after inactivation (Fig. 6a1). For comparison, we carried out inactivation of dCA1 when mice intercepted the target area on their foraging run to assess whether trajectories could be affected during their execution or whether encoding of the target location per se could be disrupted (Fig. 6a2). We found no effect on any metric of performance to be affected by this perturbation (Fig. 6a2). This insensitivity to inactivation in the navigational task is consistent with prior work in which inactivation of dCA1

during SWRs had no gross effect on navigational trajectories^{24–26} and in some cases no deficit even in navigational decisions²⁵.

As noted above, we observed distinct SPE timing in the freely-moving STF task compared to the head-fixed NTF task. The reliable timing of SPEs prior to trial initiation in the NTF task suggest they could play a role in prospective planning or initiation of forelimb trajectories. To test this possibility, we next examined the effect of optogenetic suppression of dCA1 activity in the intertrial period overlapping the time when mice reliably initiated a trial (Fig. 1j) in the NTF task (Fig. 6b1,b2). On trials with optogenetic inactivation mice dramatically reduced the probability of initiating a joystick movement. Upon release of inactivation, mice reliably initiated a joystick movement and successfully intersected the target. This bias against initiation could last for seconds with constant inactivation (Fig. 6b2). Despite a profound suppression of trial initiation, it would seem unlikely from known anatomy that dCA1 is obligate for movement of the forelimb - although interactions with lateral septum can modulate locomotion speed⁶⁸. We also considered whether the modest amount of light from the light flash could alter initiation or become a cue, however, visual light flashes in control experiments did not have any effect on trial initiation (Fig. 6b3). During sustained inactivation of dCA1 we observed a reduction, but not a complete block, of voluntary initiation of a trial (Fig. 6b2). Finally, we examined movements that were made during inactivation and found no systematic difference in peak velocity of movement (Fig. 6c, and Extended Data Fig. 6). Thus, dCA1 activity does not appear to be obligate for ongoing movement unlike in the case of inactivating an obligate structure for movement^{69,70}, but rather would appear to be a component of the circuit mechanisms involved in initiation.

dCA1 exhibits qualitatively similar representations of active foraging behavior in two distinct task contexts. However, direct quantitative comparison of the encoding of spatial information, timing of SPEs, and effects of optogenetic inactivation point to context-dependent, dissociable functional roles for dCA1. The tasks used here require a change in trajectory after a target shift that is primarily an adjustment in trajectory magnitude. This implies a challenging problem for a spatial map representation: the sequence of place cell activity is highly overlapping across target locations. Hippocampal representations, specifically in CA1, are thought to allow for conjunctive coding of a reward location (*i.e.*, spatial position \cap reward probability^{71–73}). The reward is in a constant location in our tasks and thus it is also not clear how a pure reward representation⁷³ (*i.e.* at the collection port) would provide a useful representation for learning about the target in the context of these foraging tasks. Moreover, unlike tasks that deliver reward at a variable target location(s)^{71–73}, we did not find clear evidence for an enrichment of place fields nor enhanced decoding resolution⁷¹ in the vicinity of the target location. dCA1 inactivation during foraging trajectories had no clear effect on performance (Fig. 6a2). Next, we turned back to the tML agent simulations to provide insight into how perturbation of CA1 activity could alter NTF and STF task performance.

A critical computation for the tML update rule is to calculate the difference between the parameterization of the current foraging trajectory and the average parameters of the policy. For

navigational foraging trajectories in the absence of clear spatial cues or landmarks (as in the STF task) trajectory information is thought to be ascertained by path integration^{1,74}. Information about trajectory length is thus available at the end of a trajectory at the time SPEs are primarily observed (Fig. 5a-d). This likewise implies that the primary update in the tML model might have to be computed at the end of a trajectory (Fig. 7a-e). We note a property of scaled navigational trajectories apparent from behavior: due to the elliptical shape both the departure and return heading angle scale with amplitude (Fig. 1, Fig. 2). We next asked whether SPE ensembles could provide information about the scaling of navigational trajectories in the STF task. To capture the variance of individual SPEs we projected the population vector (integrated activity per cell over 250ms around each SPE event) onto its principal components (PC^{SPE}). We then compared these loadings to a performance measure on each trial (return trajectory heading vector). All 20 sessions had a significant correlation between PC^{SPE} and return trajectory heading (ρ^2 : 0.38 ± 0.17 (s.d.); p-values: $6.8e-4 \pm 3e-3$) indicating that SPE ensembles in the STF task potentially represent a key variable required by the tML model. We reasoned that inactivation of CA1 at the time of port return could then be thought of as preventing an update in the tML model. To assess whether this could explain the behavioral effects of optogenetic inactivation we simulated the STF task in which on ~25% of trials the tML model update was blocked and the agent defaulted to a baseline policy (see Methods). Such simulations produced behavioral effects that were very similar to observed data (Fig. 7e).

In contrast to a navigational trajectory, for a forelimb movement the amplitude of the movement is represented prior to and at initiation of movement execution^{32,75}. This suggests that the computation of updates can be present prior to movement and putatively maintained via an eligibility trace until reward feedback is delivered³². Previous work has indicated that information about the current movement amplitude is present in motor cortex and striatum in an NTF task⁷⁶, however, a number of questions remain about how such information can be compared to the current policy parameter. We again considered the possibility that the update computation could depend upon dCA1 (Fig. 7f-h). This would suggest that information about the current policy could be reflected in SPEs elicited around movement initiation. We again assessed whether variance in the ensembles recruited during pre-movement SPEs might encode putative policy variables - the current target block and an online estimate of the current speed scaling parameter. In both cases we observed significant correlations in all mice (N=5) and in 100% and 76% of individual sessions, respectively (Fig. 7i). These results suggested a crucial contribution of CA1 SPEs to an implementation of the tML update computation.

Finally, the optogenetic inactivation data from the NTF task (Fig. 6b) argue that initiation of forelimb trajectories is facilitated by SPEs in dCA1. This can be incorporated into the tML model in which dCA1 SPEs interact with and facilitate an initiation process (Fig. 7f). We produced simulations of such a model that again provided a good match to the observed effects of dCA1 inactivation during normal movement initiation (Fig. 7j).

Discussion

The neural representations - a putative cognitive map - that underlie key aspects of a volitional foraging behavior remain incompletely understood; however, many lines of evidence indicate that the CA1 region of hippocampus plays a critical role in both navigational and non-navigational contexts. One of the most robust features that determines dCA1 activity is an animal's allocentric position in space ². The performance of a task, depending upon its demands, can enhance ⁷¹⁻⁷³ or reduce ^{49,77} the relative contribution of spatial position to hippocampal and entorhinal activity ⁷¹. In addition, population activity of dCA1 also represents non-navigational information, such as task events and elapsed intervals ^{10,14,44,45,78,79}. In non-navigational contexts, dCA1 activity has been described as an "internally generated" sequential recruitment of neuronal activity over suprasecond durations of active behavior ^{10,22,45,46}. Thus, in both navigational and non-navigational contexts sequential activation patterns can be thought of as reflecting a more abstract sequence of experience ^{9,80} that could provide a cognitive map of a given environment or task. The hippocampal circuit sequentially activates a sparse ensemble pattern during active behavior ⁸¹ and 'reactivates' these ensemble patterns as a synchronous burst of activity during periods of inactivity ²³. There are diverse proposals about the potential contributions of hippocampal reactivation to spatial memory, social behavior, foraging, decision making and/or reinforcement learning ^{9,19,23,29,40,82,83}. One critical outstanding question was whether reactivation of hippocampal activity is a common computational element in all contexts or whether it plays context-dependent, dissociable roles.

Here we developed two foraging tasks that are conceptually similar, but differ in the spatial context. In both the STF and NTF tasks, mice foraged for target locations that elicit reward either by navigating through a spatial environment or moving a joystick to different positions, respectively. By imaging hundreds of CA1 neurons in mice trained to perform both tasks, we discovered a clear dissociation: SPEs occurred time locked to trial initiation in the non-navigational NTF task and time locked to successful trial completion in the navigational STF task - even in the same mice trained on each task. This dissociation was consistent with a dissociation in the effect of dCA1 inactivation at the time of SPEs. These data provide some of the first causal evidence for the proposed role of SWRs (here indexed by SPEs ²²) in immediate (this trial or next trial) use for updating behavior ^{11,28,29}. This complements the well established, necessary role of SWR reactivation in learning and consolidation over tens of minutes or hours in spatial tasks ²⁴⁻²⁶. Our data replicate the observation that disruption of dCA1 during SWRs does not grossly disrupt performance of navigational trajectories^{24,25}. While we did not find strong evidence for SPEs that prospectively encoded future trajectories as in other navigational foraging tasks where rewards were located at variable positions in the environment ⁴⁰, we did observe evaluative replay events that reflected previously taken trajectory and thus could be relevant for credit assignment ¹⁹. Inactivation of CA1 during these 'evaluative' replay events caused changes in the upcoming foraging trajectories - reduced change in heading and increased amplitude - that is consistent with an impaired updating process as (generally) predicted from RL models of the role of dCA1 proposed here and previously^{82,84}.

What might explain the dissociation in dCA1 function across spatial contexts? One possibility is the different circuits that underlie the initiation of locomotion and the initiation of goal-directed

forelimb movements. The initiation of a navigational trajectory involves an orienting response and triggering of locomotion initiated in a circuit thought to include (at least) superior colliculus and mesencephalic locomotor areas^{85–87}. As might be expected *a priori* from mutually exclusive actions, skilled forelimb movements are thought to be initiated via distinct cortical and subcortical areas including the premotor and primary motor cortical regions and reticular nuclei^{69,70,88,89}. dCA1 projects directly to frontal (premotor) cortical areas in rodents^{90,91} and may thus be a critical node in the thalamocortical circuit dynamics⁹² underlying the initiation of goal-directed forelimb movements. In the primary motor cortex, an area upon which the NTF task depends⁷⁰, activity sufficient for decoding the current forelimb trajectory is present during movement initiation and execution⁷⁰. Higher motor areas and frontal cortical regions are thought to be critical for initiation of skilled forelimb movements⁷⁵ and their projections to dorsal hippocampus may also be critical for generation of SPEs⁹³. In contrast to the spatial STF task, dCA1 seems to only play a relatively minor role in representing the trajectory of forelimb movements. There are also common (although topographically distinct) circuits such as basal ganglia that are critical for scaling both forelimb movement trajectories^{30,32} and locomotor trajectories⁹⁴. This is reminiscent of the complementary representations^{95–97} and partially shared function⁹⁸ of striatum and hippocampus in other tasks and may help to explain the common learning rates observed despite the largely distinct circuits involved.

Our results highlighted the fact that understanding the contribution of dCA1 to given tasks requires assessing not only the qualitative similarity of neural activity, but also detailed activity patterns in the behavioral context⁹⁹; specifically, how an animal is ‘constructing’ its cognitive map to solve given foraging problems⁷. We articulate a computational model of an RL agent that is novel in its approach to learning foraging problems. The core idea of the tML model is that an agent can learn in the low-dimensional space of parameters governing a foraging trajectory rather than (exclusively) learning in the high dimensional state space of the environment. This perspective helps explain the smooth, closed trajectories that are readily produced by foraging animals and may provide a useful perspective on the representation of foraging trajectories in diverse species. For example, honeybees return to the hive and convey information about foraging targets via the waggle dance^{100,101}. The waggle dance is thought to represent locations in the form of a heading angle and distance, perhaps analogous to the heading angle and amplitude representation used in the tML model. Follower bees then are thought to turn this observed representation of location into a closed foraging trajectory that intercepts the target food source^{102,103}. Here we articulate how a learning rule with a plausible biological implementation³² can effectively learn to parameterize trajectories and lead to efficient foraging and rapid learning of new locations. This provides an alternative computational perspective that may complement map-based models and could be important given the systematic differences between existing model predictions and observed navigational trajectories in multiple vertebrate species³³. Our tasks are conceptually similar to a central place foraging problem in the absence of strong spatial landmark cues thus making the use of remembered trajectories essential. In a richer environment or in a distinct context animals will also presumably utilize more explicitly spatial cognitive maps. In the future we propose it is critical to incorporate learned foraging

trajectories with other well known components of navigation such as visual landmarks, local search, and path integration¹⁰⁴.

Acknowledgements

The authors declare no conflicts of interest nor relevant financial interest. The authors thank Albert Lee, Isabel Muzzio, Nelson Spruston, Laura Grima, and all members of the Dudman lab for comments on previous versions of the manuscript. Albert Lee and his lab initially provided miniscopes for imaging and expertise in surgical procedures through an HHMI equipment grant. The work in this manuscript was supported by funding from the Howard Hughes Medical Institute (HHMI). J.T.D. is a Senior Group Leader at Janelia Research Campus.

Author contributions

Study conceptualization and design: WJ, JTD. Behavior experiment and imaging data collection: WJ. Data preprocessing: WJ, SX. Data analysis: WJ, JTD, SX. Computational models: JTD. Primary writing: WJ, JTD.

Data availability

The imaging data used in this manuscript will be made available at janelia.figshare.com upon publication or can be obtained via reasonable request to the authors at www.dudmanlab.org. Code for running all simulations in paper will be made available at <https://github.com/DudLab> and/or at <http://dudmanlab.org> upon publication.

Abbreviations

Dorsal CA1 (dCA1); Synchronous population events (SPEs).

References

1. McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I. & Moser, M.-B. Path integration and the neural basis of the 'cognitive map'. *Nat. Rev. Neurosci.* **7**, 663–678 (2006).
2. O'Keefe, J. A review of the hippocampal place cells. *Prog. Neurobiol.* **13**, 419–439 (1979).
3. Eichenbaum, H. Hippocampus: cognitive processes and neural representations that underlie declarative memory. *Neuron* **44**, 109–120 (2004).
4. Epstein, R. A., Patai, E. Z., Julian, J. B. & Spiers, H. J. The cognitive map in humans: spatial navigation and beyond. *Nat. Neurosci.* **20**, 1504–1513 (2017).
5. Behrens, T. E. J. *et al.* What Is a Cognitive Map? Organizing Knowledge for Flexible Behavior. *Neuron* **100**, 490–509 (2018).
6. Stachenfeld, K. L., Botvinick, M. & Gershman, S. J. Design Principles of the Hippocampal Cognitive Map. in *Advances in Neural Information Processing Systems 27* (eds. Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N. D. & Weinberger, K. Q.) 2528–2536 (Curran Associates, Inc., 2014).
7. Tolman, E. C. Cognitive maps in rats and men. *Psychol. Rev.* **55**, 189–208 (1948).
8. O'keefe, J. & Nadel, L. *The hippocampus as a cognitive map*. (Oxford: Clarendon Press, 1978).
9. Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive map. *Nat. Neurosci.* **20**, 1643–1653 (2017).
10. Aronov, D., Nevers, R. & Tank, D. W. Mapping of a non-spatial dimension by the hippocampal-entorhinal circuit. *Nature* **543**, 719–722 (2017).
11. Liu, Y., Dolan, R. J., Kurth-Nelson, Z. & Behrens, T. E. J. Human Replay Spontaneously Reorganizes Experience. *Cell* vol. 178 640–652.e14 (2019).
12. Dudchenko, P. A. & Wood, E. R. Splitter Cells: Hippocampal Place Cells Whose Firing Is Modulated by Where the Animal Is Going or Where It Has Been. in *Space, Time and Memory in the Hippocampal Formation* (eds. Derdikman, D. & Knierim, J. J.) 253–272 (Springer Vienna, 2014).
13. Benna, M. K., Rigotti, M., Munuera, J., Fusi, S. & Salzman, C. D. The geometry of abstraction in hippocampus and pre-frontal cortex. *bioRxiv* (2018).
14. Wood, E. R., Dudchenko, P. A., Robitsek, R. J. & Eichenbaum, H. Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron* **27**, 623–633 (2000).
15. Eichenbaum, H. & Cohen, N. J. Can we reconcile the declarative memory and spatial navigation views on

- hippocampal function? *Neuron* **83**, 764–770 (2014).
16. Rich, P. D., Liaw, H.-P. & Lee, A. K. Place cells. Large environments reveal the statistical structure governing hippocampal representations. *Science* **345**, 814–817 (2014).
 17. Dragoi, G. & Buzsáki, G. Temporal encoding of place sequences by hippocampal cell assemblies. *Neuron* **50**, 145–157 (2006).
 18. Carr, M. F., Jadhav, S. P. & Frank, L. M. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nat. Neurosci.* **14**, 147–153 (2011).
 19. Foster, D. J. & Wilson, M. A. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* **440**, 680–683 (2006).
 20. Louie, K. & Wilson, M. A. Temporally structured replay of awake hippocampal ensemble activity during rapid eye movement sleep. *Neuron* **29**, 145–156 (2001).
 21. Lee, A. K. & Wilson, M. A. Memory of sequential experience in the hippocampus during slow wave sleep. *Neuron* **36**, 1183–1194 (2002).
 22. Villette, V., Malvache, A., Tressard, T., Dupuy, N. & Cossart, R. Internally Recurring Hippocampal Sequences as a Population Template of Spatiotemporal Information. *Neuron* **88**, 357–366 (2015).
 23. Buzsáki, G. Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus* **25**, 1073–1188 (2015).
 24. Girardeau, G., Benchenane, K., Wiener, S. I., Buzsáki, G. & Zugaro, M. B. Selective suppression of hippocampal ripples impairs spatial memory. *Nat. Neurosci.* **12**, 1222 (2009).
 25. Jadhav, S. P., Kemere, C., German, P. W. & Frank, L. M. Awake hippocampal sharp-wave ripples support spatial memory. *Science* **336**, 1454–1458 (2012).
 26. van de Ven, G. M., Trouche, S., McNamara, C. G., Allen, K. & Dupret, D. Hippocampal Offline Reactivation Consolidates Recently Formed Cell Assembly Patterns during Sharp Wave-Ripples. *Neuron* **92**, 968–974 (2016).
 27. Ego-Stengel, V. & Wilson, M. A. Disruption of ripple-associated hippocampal activity during rest impairs spatial learning in the rat. *Hippocampus* **20**, 1–10 (2010).
 28. Vaz, A. P., Inati, S. K., Brunel, N. & Zaghoul, K. A. Coupled ripple oscillations between the medial temporal lobe and neocortex retrieve human memory. *Science* **363**, 975–978 (2019).
 29. Joo, H. R. & Frank, L. M. The hippocampal sharp wave-ripple in memory retrieval for immediate use and consolidation. *Nat. Rev. Neurosci.* **19**, 744–757 (2018).

30. Panigrahi, B. *et al.* Dopamine Is Required for the Neural Representation and Control of Movement Vigor. *Cell* **162**, 1418–1430 (2015).
31. Bradfield, L. A., Leung, B. K., Boldt, S., Liang, S. & Balleine, B. W. Goal-directed actions transiently depend on dorsal hippocampus. *Nat. Neurosci.* **23**, 1194–1197 (2020).
32. Yttri, E. A. & Dudman, J. T. Opponent and bidirectional control of movement velocity in the basal ganglia. *Nature* **533**, 402–406 (2016).
33. de Cothi, W. *et al.* Predictive Maps in Rats and Humans for Spatial Navigation. *Cold Spring Harbor Laboratory* 2020.09.26.314815 (2020) doi:10.1101/2020.09.26.314815.
34. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction*. (MIT Press, 1998).
35. Yttri, E. A. & Dudman, J. T. A proposed circuit computation in basal ganglia: History-dependent gain. *Mov. Disord.* (2018).
36. Kober, J., Bagnell, J. A. & Peters, J. Reinforcement learning in robotics: A survey. *The International Journal of* (2013).
37. Ghosh, K. K. *et al.* Miniaturized integration of a fluorescence microscope. *Nat. Methods* **8**, 871–878 (2011).
38. Chen, T.-W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
39. Gupta, A. S., van der Meer, M. A. A., Touretzky, D. S. & Redish, A. D. Segmentation of spatial experience by hippocampal θ sequences. *Nat. Neurosci.* **15**, 1032–1039 (2012).
40. Pfeiffer, B. E. & Foster, D. J. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature* **497**, 74–79 (2013).
41. Ziv, Y. *et al.* Long-term dynamics of CA1 hippocampal place codes. *Nat. Neurosci.* **16**, 264–266 (2013).
42. Thompson, R. F. *et al.* Hippocampal substrate of classical conditioning. *Physiological Psychology* vol. 8 262–279 (1980).
43. Moustafa, A. A. *et al.* Why trace and delay conditioning are sometimes (but not always) hippocampal dependent: a computational model. *Brain Res.* **1493**, 48–67 (2013).
44. Eichenbaum, H. What Versus Where: Non-spatial Aspects of Memory Representation by the Hippocampus. *Curr. Top. Behav. Neurosci.* **37**, 101–117 (2018).
45. MacDonald, C. J., Lepage, K. Q., Eden, U. T. & Eichenbaum, H. Hippocampal ‘Time Cells’ Bridge the Gap in Memory for Discontiguous Events. *Neuron* **71**, 737–749 (2011).
46. Pastalkova, E., Itskov, V., Amarasingham, A. & Buzsáki, G. Internally generated cell assembly sequences in the

- rat hippocampus. *Science* **321**, 1322–1327 (2008).
47. Kraus, B. J., Robinson, R. J., White, J. A., Eichenbaum, H. & Hasselmo, M. E. Hippocampal 'Time Cells': Time versus Path Integration. *Neuron* vol. 78 1090–1101 (2013).
 48. Ravassard, P. *et al.* Multisensory control of hippocampal spatiotemporal selectivity. *Science* **340**, 1342–1346 (2013).
 49. Acharya, L., Aghajan, Z. M., Vuong, C., Moore, J. J. & Mehta, M. R. Causal Influence of Visual Cues on Hippocampal Directional Selectivity. *Cell* **164**, 197–207 (2016).
 50. Bittner, K. C. *et al.* Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. *Nat. Neurosci.* **18**, 1133–1142 (2015).
 51. Zhao, X., Wang, Y., Spruston, N. & Magee, J. C. Synaptic mechanisms of context-dependent sensory responses in the hippocampus. doi:10.1101/624262.
 52. Buzsáki, G. & Mizuseki, K. The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* **15**, 264–278 (2014).
 53. Ahmed, M. S. *et al.* Hippocampal Network Reorganization Underlies the Formation of a Temporal Association Memory. *Neuron* (2020) doi:10.1016/j.neuron.2020.04.013.
 54. Whishaw, I. Q. & Vanderwolf, C. H. Hippocampal EEG and behavior: Change in amplitude and frequency of RSA (Theta rhythm) associated with spontaneous and learned movement patterns in rats and cats. *Behav. Biol.* **8**, 461–484 (1973).
 55. Lalla, L., Orozco, P. E. R., Jurado-Parras, M.-T., Brovelli, A. & Robbe, D. Local or not local: investigating the nature of striatal theta oscillations in behaving rats. *eneuro* **4**, (2017).
 56. Buzsáki, G. Theta oscillations in the hippocampus. *Neuron* (2002).
 57. Colgin, L. L., Moser, E. I. & Moser, M.-B. Understanding memory through hippocampal remapping. *Trends Neurosci.* **31**, 469–477 (2008).
 58. Squire, L. R. Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem.* **82**, 171–177 (2004).
 59. Malvache, A., Reichinnek, S., Villette, V., Haimerl, C. & Cossart, R. Awake hippocampal reactivations project onto orthogonal neuronal assemblies. *Science* **353**, 1280–1283 (2016).
 60. Roux, L., Hu, B., Eichler, R., Stark, E. & Buzsáki, G. Sharp wave ripples during learning stabilize the hippocampal spatial map. *Nat. Neurosci.* **20**, 845–853 (2017).

61. MacDonald, C. J., Fortin, N. J., Sakata, S. & Meck, W. H. Retrospective and Prospective Views on the Role of the Hippocampus in Interval Timing and Memory for Elapsed Time. *Timing & Time Perception* **2**, 51–61 (2014).
62. Rogers, J. L., Hunsaker, M. R. & Kesner, R. P. Effects of ventral and dorsal CA1 subregional lesions on trace fear conditioning. *Neurobiol. Learn. Mem.* **86**, 72–81 (2006).
63. Hall, S. R. & Hagstrom, E. C. Hippocampal lesions impair trace conditioning. *PsycEXTRA Dataset* (1969) doi:10.1037/e666672011-192.
64. Tseng, W., Guan, R., Disterhoft, J. F. & Weiss, C. Trace eyeblink conditioning is hippocampally dependent in mice. *Hippocampus* vol. 14 58–65 (2004).
65. Kim, J. J., Clark, R. E. & Thompson, R. F. Hippocampectomy impairs the memory of recently, but not remotely, acquired trace eyeblink conditioned responses. *Behav. Neurosci.* **109**, 195–203 (1995).
66. Arenkiel, B. R. *et al.* In vivo light-induced activation of neural circuitry in transgenic mice expressing channelrhodopsin-2. *Neuron* **54**, 205–218 (2007).
67. Babl, S. S., Rummell, B. P. & Sigurdsson, T. The Spatial Extent of Optogenetic Silencing in Transgenic Mice Expressing Channelrhodopsin in Inhibitory Interneurons. *Cell Rep.* **29**, 1381–1395.e4 (2019).
68. Bender, F. *et al.* Theta oscillations regulate the speed of locomotion via a hippocampus to lateral septum pathway. *Nature Communications* vol. 6 (2015).
69. Guo, J.-Z. *et al.* Cortex commands the performance of skilled movement. *Elife* **4**, e10774 (2015).
70. Park, J., Phillips, J. W., Martin, K. A., Hantman, A. W. & Dudman, J. T. Flexible routing of motor control signals through neocortical projection neuron classes. *bioRxiv* 772517 (2019) doi:10.1101/772517.
71. Butler, W. N., Hardcastle, K. & Giocomo, L. M. Remembered reward locations restructure entorhinal spatial maps. *Science* vol. 363 1447–1452 (2019).
72. Lee, J. S., Briguglio, J., Romani, S. & Lee, A. K. The statistical structure of the hippocampal code for space as a function of time, context, and value. *bioRxiv* (2019).
73. Gauthier, J. L. & Tank, D. W. A Dedicated Population for Reward Coding in the Hippocampus. *Neuron* **99**, 179–193.e7 (2018).
74. Webb, B. & Wystrach, A. Neural mechanisms of insect navigation. *Curr Opin Insect Sci* **15**, 27–39 (2016).
75. Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* **36**, 337–359 (2013).
76. Park, J., Coddington, L. T. & Dudman, J. T. Basal Ganglia Circuits for Action Specification. *Annu. Rev. Neurosci.*

(2020) doi:10.1146/annurev-neuro-070918-050452.

77. Muzzio, I. A. *et al.* Attention enhances the retrieval and stability of visuospatial and olfactory representations in the dorsal hippocampus. *PLoS Biol.* **7**, e1000140 (2009).
78. McEchron, M. D., Tseng, W. & Disterhoft, J. F. Single neurons in CA1 hippocampus encode trace interval duration during trace heart rate (fear) conditioning in rabbit. *J. Neurosci.* **23**, 1535–1547 (2003).
79. McEchron, M. D. & Disterhoft, J. F. Sequence of single neuron changes in CA1 hippocampus of rabbits during acquisition of trace eyeblink conditioned responses. *J. Neurophysiol.* **78**, 1030–1044 (1997).
80. Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B. & Botvinick, M. M. Neural representations of events arise from temporal community structure. *Nat. Neurosci.* **16**, 486–492 (2013).
81. Buzsáki, G. & Tingley, D. Space and Time: The Hippocampus as a Sequence Generator. *Trends Cogn. Sci.* **22**, 853–869 (2018).
82. Mattar, M. G. & Daw, N. D. Prioritized memory access explains planning and hippocampal replay. *Nat. Neurosci.* **21**, 1609–1617 (2018).
83. Oliva, A., Fernández-Ruiz, A., Leroy, F. & Siegelbaum, S. A. Hippocampal CA2 sharp-wave ripples reactivate and promote social memory. *Nature* (2020) doi:10.1038/s41586-020-2758-y.
84. Foster, D. J., Morris, R. G. & Dayan, P. A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus* **10**, 1–16 (2000).
85. Heindorf, M., Arber, S. & Keller, G. B. Mouse Motor Cortex Coordinates the Behavioral Response to Unpredicted Sensory Feedback. *Neuron* **101**, 1202 (2019).
86. Kiehn, O. & Dougherty, K. Locomotion: Circuits and Physiology. *Neuroscience in the 21st Century* 1337–1365 (2016) doi:10.1007/978-1-4939-3474-4_42.
87. Vale, R. *et al.* A cortico-collicular circuit for accurate orientation to shelter during escape. *Cold Spring Harbor Laboratory* 2020.05.26.117598 (2020) doi:10.1101/2020.05.26.117598.
88. Churchland, M. M., Santhanam, G. & Shenoy, K. V. Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. *J. Neurophysiol.* **96**, 3130–3146 (2006).
89. Esposito, M. S., Capelli, P. & Arber, S. Brainstem nucleus MdV mediates skilled forelimb motor tasks. *Nature* **508**, 351–356 (2014).
90. Swanson, L. W. A direct projection from Ammon's horn to prefrontal cortex in the rat. *Brain Res.* **217**, 150–154 (1981).

91. Jay, T. M. & Witter, M. P. Distribution of hippocampal CA1 and subicular efferents in the prefrontal cortex of the rat studied by means of anterograde transport of Phaseolus vulgaris-leucoagglutinin. *J. Comp. Neurol.* **313**, 574–586 (1991).
92. Sauerbrei, B. A. *et al.* Cortical pattern generation during dexterous movement is input-driven. *Nature* **577**, 386–391 (2020).
93. Rajasethupathy, P. *et al.* Projections from neocortex mediate top-down control of memory retrieval. *Nature* **526**, 653–659 (2015).
94. Rueda-Orozco, P. E. & Robbe, D. The striatum multiplexes contextual and kinematic information to constrain motor habits execution. *Nat. Neurosci.* **18**, 453–460 (2015).
95. van der Meer, M. A. A., Johnson, A., Schmitzer-Torbert, N. C. & Redish, A. D. Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron* **67**, 25–32 (2010).
96. Albouy, G., King, B. R., Maquet, P. & Doyon, J. Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus* vol. 23 985–1004 (2013).
97. Berke, J. D., Breck, J. T. & Eichenbaum, H. Striatal Versus Hippocampal Representations During Win-Stay Maze Performance. *Journal of Neurophysiology* vol. 101 1575–1587 (2009).
98. Ferbinteanu, J. The hippocampus and dorso-lateral striatum link distinct types of memories through time and space, respectively. doi:10.1101/2019.12.16.878470.
99. Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A. & Poeppel, D. Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron* **93**, 480–490 (2017).
100. Menzel, R. The Waggle Dance as an Intended Flight: A Cognitive Perspective. *Insects* **10**, (2019).
101. von Frisch, K. The dance language and orientation of bees. (1967).
102. Riley, J. R., Greggers, U., Smith, A. D., Reynolds, D. R. & Menzel, R. The flight paths of honeybees recruited by the waggle dance. *Nature* **435**, 205–207 (2005).
103. Esch, H. E., Zhang, S., Srinivasan, M. V. & Tautz, J. Honeybee dances communicate distances measured by optic flow. *Nature* vol. 411 581–583 (2001).
104. Sharp, P. E. *The Neural Basis of Navigation: Evidence from Single Cell Recording*. (Springer Science & Business Media, 2012).
105. Cembrowski, M. S., Wang, L., Sugino, K., Shields, B. C. & Spruston, N. Hipposeq: a comprehensive RNA-seq

- database of gene expression in hippocampal principal neurons. *Elife* **5**, e14997 (2016).
106. Osborne, J. E. & Dudman, J. T. RIVETS: a mechanical system for in vivo and in vitro electrophysiology and imaging. *PLoS One* **9**, e89007 (2014).
107. Pnevmatikakis, E. A. *et al.* Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* **89**, 285–299 (2016).
108. Zhou, P. *et al.* Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data. *Elife* **7**, (2018).

Materials & Methods

Male and female mice, typically aged 3-6 months at time of surgery, were used in this study. All procedures were approved by the Janelia Research Campus Institutional Animal Care and Use Committee (IACUC) and were consistent with the standards of the Association for Assessment and Accreditation of Laboratory Animal Care.

Guide cannula implantation

Five male mice (Three GP4.3 mice, two Ai93(TITL-GCaMP6f)-D;ROSA26-ZtTA×Kcnd2-IRES-Cre 3G5 mice) aged 3-6 months at the start were used in this study. The Kcnd2-IRES-Cre 3G5 mice have been generated in house in the Janelia Transgenic Core (<https://www.janelia.org/support-team/gene-targeting-and-transgenics>) based upon evidence for Kcnd2 expression in principal neurons of dorsal CA1¹⁰⁵ and are available upon request (Extended Data Fig. 7). Mice were anesthetized under isoflurane (1.5%-2%) anesthesia. A 1.8 mm-diameter circular craniotomy centered on AP: 1.9mm, ML +1.5 mm was opened with a trephine drill (1.8mm diameter). Dura was removed and the cortex above CA1 was aspirated with a 27 gauge blunt needle followed by a 30 gauge needle as the hippocampus was approached until vertical white fiber tracts were visible. During this procedure, bleeding was controlled by constantly irrigating the exposed tissue with sterile 0.9% saline. Then a guide cannula with a bottom glass window (Diameter (Outer): 1.8mm, Length: ~3.6mm; Part ID: 1050-002191, inscopix) was placed above dorsal CA1. The guide cannula was affixed to the skull with dental cement (Calibra Universal cement), then a head bar¹⁰⁶; details can be obtained from <http://dudmanlab.org/html/rivets.html>) was affixed to the skull with dental cement. At the end of the surgery, the top of the guide cannula was covered by parafilm. A silicone adhesive (Kwik-Sil; World Precision Instruments) was then applied above the parafilm.

Three to four weeks after the guide cannula implantation, awake mice were head-fixed by a head bar holder. An inner cannula lens sleeve (come with the guide cannula, Inner diameter: ~1.0 mm, length: ~4mm) was inserted into the guide cannula first, then a grin lens (1mm diameter, ~4 mm length; Part ID: 1050-002176; Inscopix) was placed into the inner cannula. A baseplate (Part ID:1050-004201; Inscopix) attached to the miniature microscope was positioned above the GRIN lens. The focal plane was adjusted until GCaMP6 fluorescence responses were clearly observed. Then the mice were anaesthetized by isoflurane and the baseplate was affixed to the skull with dental cement.

Optical fiber implantation and optical stimulation

VGAT-ChR2-EYFP (Jackson Labs Stock 014548, VGAT-ChR2-EYFP line 8) mice were used for optical stimulation. A guide cannula was implanted above dorsal CA1 first (same procedure as above imaging window). In NTF task (N = 3 mice), at the start of each session, an optical fiber (200 µm core, 0.53 NA; doric) coupled with 473 nm laser source (Fiberoptics) was placed into the center of the guide cannula (~3 mm depth from the top of the cannula) held by a stereotaxic micromanipulator. After each session, the optical fiber was taken out of the cannula and the top of guide cannula was covered by parafilm. A silicone adhesive (Kwik-Sil; World Precision Instruments) was then applied above the parafilm. In STF task (N = 4 mice), an inner sleeve (~1.0 mm inner diameter, 4.0 mm long) was inserted into the guide cannula first, then an optical fiber (200 µm core, 0.53 NA, 3mm long, doric) was lowered down by a stereotaxic micromanipulator into the inner sleeve, until the ferrule of the optical fiber just touched the inner sleeve. The optical fiber was placed in the center of the inner sleeve. Then the dental cement was used to fix the guide tube, the inner sleeve and the ferrule of optical fiber together.

During the session, optical fiber was coupled to a 473 nm laser source (Fiberoptics) to deliver light onto the dorsal hippocampus through the guide cannula window. 10ms pulses, 25 Hz laser with power measured at the tip of the fiber of 2-3 mW were delivered at different behavior phases with variable time length in 30% of the behavior trials. We chose this intensity to ensure complete suppression of illuminated regions of the hippocampus, while minimizing effects on underlying visual thalamic nuclei⁶⁷.

Behavior: NTF task

Behavioral code was implemented as described previously and run from a microcontroller based system (details can be obtained from <http://dudmanlab.org/html/resources.html>). After surgery, mice were given 5 days of recovery prior to beginning water restriction (1ml water/day). Following 7 days of initial water restriction, they underwent 4-8 weeks training. Mice were head-fixed in a custom made head restraint box using the RIVETS head-fixation apparatus¹⁰⁶. The mouse's front paws rested on a metal bar attached to a spring-loaded joystick, which had unconstrained 2D maneuverability in the horizontal plane as described previously^{30,32}. Mice were trained to displace the joystick to certain target position ranges varying across four different blocks (e.g. 4.5-5.5-5-4.5 mm) to obtain a sweetened water reward delivered 1 s after each threshold crossing. Rewards were followed by a 3.3 s inter-trial interval (ITI) in which no movements would be rewarded. There were up to 160 trials (40 trials per block) per imaging session, with one water reward being available per trial. Forelimb movements were assessed offline to detect individual reaches based on the velocity of joystick movement. *Note*: NTF video data was not recorded in this dataset, but analogous performance data can be found online with a previous publication³².

Behavior: STF task

After surgery, mice were given 5 days of recovery prior to beginning water restriction (1.2ml water/day). Following 7 days of initial water restriction, they underwent 4-8 weeks training. In this self-paced free foraging task, mice were placed in a 75cm x 75cm box. There was a water spout on one wall of the box (we defined an area 20cm x 14cm around the water spout as the reward area). Mice were required to run into an unmarked location (~18cm x 14cm) triggering the reward delivery, then come back to the reward area to consume the water. Next trial starts 2 seconds after the mice enter into the reward area. See Supplementary video 1 for an example set of trials. There were two different blocks with two different unmarked locations, location1(the center is ~34cm away from the reward area), location2 (the center is ~52 cm away from the reward area). As the target location shifted mice were able to reliably adjust their movements to collect rewards in both tasks. There were up to 160 trials (80 trials per block) per imaging session, with one water reward being available per trial.

In the STF task, the mouse's position was recorded via a USB camera mounted below the clear platform of the enclosure. Briefly, a real-time tracking algorithm was developed in which the video frame was converted to black and white, subtracting a blank background without a mouse, blurred, and then a standard OpenCV blob detection algorithm was applied with user customizable threshold settings. The center of the mouse body was calculated at every frame from the center of the detected blob and a running buffer of positions were tracked by custom software written in Processing (www.processing.org) and written to a file. The tracking video was synchronized to the imaging using a TTL signal from a tracking programme to trigger data acquisition on the Inscopix miniature microscope (www.inscopix.com). Frames were linearly interpolated to match the sampling rate of the microscope. All analysis of foraging trajectories was performed offline using stored position data in Matlab 2018 (www.mathworks.com).

Data analysis: calcium imaging

In NTF task, mice (N=5) were head-fixed in a custom made head restraint box using the RIVETS head-fixation apparatus and the microscope was connected to the baseplate when the animal was headfixed, after adjusting to the best imaging focal plane, the imaging session started. In the STF task (N = 2 mice), first, the microscope was connected to the baseplate when the animal was head-fixed, after adjusting to the best focal plane, mice were removed from head fixation and were put into the free moving behavior box.

Fluorescence images were acquired at 10 Hz (GP4.3 mice) or 17 Hz (Ai93(TITL-GCaMP6f)-D;ROSA26-ZtTA×Kcnd IRESCre 3G5 mice; www.jax.org) and the LED power was set 10–35% (0.1–0.35mW) with analogue gain 3–3.5. To decrease the photo bleaching effect, in NTF task, imaging recording was performed every 10 trials; in STF task, imaging record was performed 40 trials each block.

All recorded calcium videos from one animal in one day were concatenated in Fiji. The concatenated video was spatially down sampled 2x, movement corrected using Mosaic (www.inscopix.com). Then, the corrected video was cropped to remove correction artifacts and exclude areas with no GCaMP6f³⁸ activity. The cropped video was further spatially down sampled 2x (usually resulting in 350 x 300 pixel videos). CNMF-E package^{107,108} was used to automatically segment neurons from the preprocessed videos. The neuron ROIs from CNMF-E were manually examined and corrected. Calcium signals within these corrected ROIs were extracted with CNMF-E. Spike trains were inferred with the deconvolution function in CNMF-E package (constrained FOOPSI).

Data analysis: place fields

To analyze place fields, we identified ‘movement periods’ when the mouse ran in open field arenas at the speed ≥ 1 cm/s. These criteria rejected small movements such as grooming, rearing, or head turning. We spatially binned the open field arenas into 4 cm x 4 cm bins. To suppress noise, we also identified ‘foraging bins’, into which bins the mouse ran ≥ 5 times in one session. We divided the number of calcium transients in each foraging bin by the mouse's total occupancy time there, applied a Gaussian smoothing filter ($\sigma = 4$ cm), and normalized each place field by its maximum value.

Data analysis: neural correlates of behavior

All data analysis was accomplished with custom written code in Matlab (www.mathworks.com) and will be made available at the lab GitHub (<https://github.com/DudLab>) upon publication.

Briefly, in both tasks individual movements in trained mice were quite well isolated (see extended traces in Fig. 1, Fig. 3). In the NTF task, analysis was preceded by identifying the start and stop time of each individual movement. Movements were required to be at least 1 sec in duration with at least 1 sec between well separated movements. Raw position data was centered around either the reward collection port (STF task) or the true 0 position of the joystick (NTF task). Speed was computed by taking a simple pointwise difference and smoothing with a Savitzky-Golay filter. In the NTF task, a threshold was used to estimate the onset and offset of movement events. A number of statistics of movement were then computed from these events. Whether a movement event was rewarded or not was determined by

looking for reward triggers occurring during an event. >95% of rewards could be attributed to a single well-isolated movement event in all sessions used for analysis.

Cross-validated PMTH alignment in Fig. 3 was determined by taking a random half of trials, sorting by time of peak response magnitude, and then using that ROI index array to sort the held out half of trials. The results were plotted in Fig. 3 for both tasks. For continuous plots of data shown in Fig. 3, we accomplished a hierarchical sorting of activity by first dividing ROIs around the median of average activity over the session and then within each group re-sorting by latency to peak response from movement onset. This array of ROI indices was used in all subsequent plotting.

Data analysis: decoder and classifier construction

To decode the continuous behavior from inferred spikes in the imaging data we took an approach we recently described ⁷⁰ that is inspired by the use of committee machines in machine learning. Briefly, we sought to identify a linear decoder to estimate the joystick movement or body position. The decoder defines linear mapping (W_{decode}) between the neural population activity and the two dimensional position:

$$K = F^T \cdot W_{decode}$$

where F is the data matrix comprising the population vector of spike counts with the dimension of the number of units concatenated across all time bins and trials in the training data set. The matrix K comprises two vectors each corresponding to the decoded position ($\{x,y\}$ or $\{angle,radius\}$). We solve for W_{decode} as $W_{decode} = (F^T F)^{-1} F^T K$ using the Moore-Penrose inverse on a subset of randomly permuted and concatenated trials. This approach yielded noisy decoder performance on cross-validation. To reduce noise and provide better generalization, we computed a family of linear decoders from N folds of P permuted trials. Typical values were $N=50$ and $P=75$. We then took the mean of the family of decoders ($N \times$ Number of units) to yield a 'consensus' decoder. Decoding performance is illustrated with this consensus decoder applied to a unique permuted sequence of trials.

To classify the block identities of trials from the neural ensemble, we first computed a 1D population vector comprised of the total activation (trapezoidal integration) during the movement trajectory (± 1 sec around target interception) for each trial. We then trained a BoostedTrees ensemble decoder with 10-fold cross validation (*trainClassifier*, Matlab Statistics and Machine Learning toolbox) and reported the summary performance (average of true negative and true positive rates).

Data analysis: Perturbation effects in NTF task

To assess the effects of optogenetic inactivation of hippocampal neuronal ensemble on NTF task, forelimb movements were aligned with the reward event (Fig. 6b) within 12 s after that event. Because only ~30% trials were inactivation catch trials, we randomly resampled (with replacement) k trials of the aligned movement from the catch and control trials respectively, where k is the number of catch trials. Then, we used the aligned movement in the resampled catch and control trials to compute their post event time histograms (PETHs) of movement. To statistically evaluate the difference in PETH between catch and control trials, we repeated the resampling and PETHs calculation procedure 1000 times. Mean PETHs and 95% (2.5%-97.5%) confidence intervals (CIs) of PETHs under inactivation and control condition were calculated with the 1000 resampled PETHs. To remove transient noise, only the time spans greater than 200 ms and no overlapping between the 95% CIs were marked with red horizontal lines in Fig. 6b to show when inactivation significantly affected VOA behavior.

Data analysis: Perturbation effects on STF task

To assess the effects of optogenetic inactivation of hippocampal neuronal ensemble on STF task, we extracted each individual foraging trajectory run as described above. We then identified runs that contained a laser perturbation either in the collection area immediately following run termination or during interception of the target area. For each run we calculated a number of statistics including its probability of correctly intercepting the target, latency to initiation relative to end of prior trial, path length, initial heading angle when leaving collection or heading angle when returning to the collection area. We used 2-way ANOVAs with main effects for mouse and stimulation/no-stimulation to compare groups. For most comparisons there were significant differences between mice that are not described directly in the text.

Computational modeling: Trajectory MeSH Learning (tML) Agent

The tML agent model is based around the idea that an agent can learn to scale the parameters of a structured representation of foraging trajectories. For a central place forager we might demand a trajectory that forms a closed out and back loop which begins at a 'home' location, transits through an extrema and returns home. The goal of the learning agent is to update the heading and amplitude of this trajectory so that it reliably intercepts a target location according to the specific rules of the environment. For example, interception may need to occur at the trajectory extrema or perhaps anywhere along the trajectory or perhaps for some fixed duration. In the specific cases for this study we consider interception either at the extrema (STF task) or for a fixed duration (NTF task) that correspond to the practical requirements of our real-time behavior analysis used in the experimental task designs. We note that similar results to those reported have been obtained with a range of different simulated environments.

Returning to the notion of a structured, closed-loop trajectory, we consider the problem as a control signal that determines behavior at each time step. First consider a locomoting animal. At each time step we assume it is governed by a heading angle and an instantaneous speed. Under such a model a closed-loop trajectory will be produced by a smooth rotation in heading angle (a linear function from $-\pi$ to π). For a fixed speed this would produce a rotation about a circle. However, to produce the observed, roughly elliptical paths speed is inhomogeneous and reaches maxima along specific heading angles - outward runs ($\pi/2$) and return runs ($-\pi/2$). Given the expected bell-shaped distribution of speeds that minimize jerk along a trajectory this can be modeled as a sequence of gaussian speed profiles. These dynamics for heading and speed can be generated by an artificial neural network, but for simplicity we have used simple generative functions perturbed by noise. A schematic of the model architecture can be found in Fig. 2 and Fig. 7.

$$\Theta(t) = L(-\pi, \pi) + \omega[i] + \varepsilon \quad (1)$$

$$S(t) = G(\tau, \sigma) \times a[i] + \varepsilon \quad (2)$$

Where, L is a linear mapping across the range $\{-\pi, \pi\}$ spanning time t , $\omega[i]$ is the heading offset sampled from a distribution of mean $\Omega[i]$ and constant variance for each trial i . G is either a single (NTF) or double (STF) peaked gaussian function with offsets $\tau = \{\tau_1\}$ or $\{\tau_1, \tau_2\}$ and width, scaled by a gain $a[i]$ sampled from a distribution with a mean $A[i]$ and constant variance for each trial i . ε is a normally distributed and smoothed noise term matched to observed variability in observed behavioral trajectories.

Given a model of structured trajectories defined by continuous speed ($S(t)$) and heading ($\Theta(t)$), the learning problem for an agent is to learn to scale the amplitude ($A(i)$) and orientation ($\Omega(i)$) offsets of trajectories trial by trial in order to reliably intercept target locations. Behavioral data indicated bidirectional

and rapid learning for changes in the scaling of movement trajectories, thus we used a modified version of a learning rule (MeSH) previously described to account for rapid, bidirectional movement scaling^{32,35}.

$$A[i+1] = A[i] + (a[i] - A[i]) [i] - \beta(a[i] - A[0])) \quad (3)$$

$$\Omega[i+1] = \Omega[i] + (\omega[i] - \Omega[i]) [i] - \beta(\omega[i] - \Omega[0])) \quad (4)$$

where , i is the index of the i th trial, $[i]$ is a smoothed estimate of the local reward rate, $a[i]$ is the magnitude of the speed on the current trial as sampled from a normal distribution centered on $A[i]$ with rate parameters μ and σ . Learning rate parameters and the standard deviation of the distribution (A, σ) were explored using grid search optimization. The equivalent learning rule is also expressed for $\Omega(i)$ in equation 4.

In order to account for effects of inactivation we considered two implementation modifications to the tML model corresponding to the distinct behavioral context of the STF and NTF tasks (see discussion in Main Text). Simulation data and schematics depicting these model formulations are shown in Fig. 7.

First, we consider the STF task in which CA1 SPEs were observed immediately upon completion of a foraging trajectory and return to the reward location. The critical computation for learning in the tML model is the MeSH update (equation 3) that depends upon the signed difference between the current trial speed, $a[i]$, (or heading) and the current policy speed, $A[i]$ (or heading). A precise circuit mechanism for this computation is unclear and beyond the scope of the current study, but one possibility consistent with our experimental data is that CA1 SPEs encode information about the current trajectory. In such a formulation, we consider a model in which the SPE is necessary to update the policy and in the absence of an SPE the policy reverts to its default $A[0]$.

Second, we consider the NTF task in which CA1 SPEs were observed just prior to initiation of a joystick movement. Again we postulate that the occurrence of an SPE is critical for a learning update, however, we note some key differences in the control of skilled forelimb movements as contrasted with navigational trajectories (see Discussion in Main Text). Previous modeling work in the context of tasks like the NTF task have been consistent with the possibility that the putative MeSH update is produced in the form of an eligibility trace at the time of movement initiation^{32,35}. Here we consider the additional possibility that movement initiation is facilitated by the occurrence of a CA1 SPE. We note that this would be a particularly useful formulation to ensure that a viable eligibility trace is present when movements are initiated given the width of the distribution of movement initiation times and relatively low frequency of SPEs of about 1Hz. To model such an initiation process we generated a hazard function that matched the observed latency distribution and determined the probability of initiating a trial. The hazard function is given by:

$$H = g(\mu, \sigma) / 1 - G$$

Where, g is a gaussian function with mean 3 seconds and standard deviation of 0.48 seconds. G is the cumulant density of g . Individual trial latencies were determined by sampling a uniform random variable for the time point at which it exceeded probability H if an SPE had occurred. We used the observed empirical distribution of intervals between SPEs for all datasets to draw event times for an SPE. In the case of optogenetic inactivation we assume that the probability of an SPE by ~75%, but also resulted in an SPE with high probability at offset of inactivation due to rebound excitation⁶⁹.

For simulations of an epsilon ϵ -greedy Q learning agent (Extended Data Fig. 1) we used a standard approach³⁴. For simulations shown $\epsilon=0.1$ and $\alpha=\{0.3,0.5,0.7\}$ (many other parameterizations were explored with qualitatively similar results).

Code for running all simulations in paper will be made available (<https://github.com/DudLab>) and/or (<http://dudmanlab.org>) upon publication.

Figure Legends

Fig. 1| Foraging tasks with similar task structure but distinct spatial contexts

a. Schematic representation of the freely moving spatial target foraging (STF) task in which mice must find an unmarked and variable target area on the floor of their arena. Mouse position was tracked in real time below the floor and brief occupancy within the target box yielded a water reward delivered through a port positioned on the near wall within a collection area. Upper right inset shows the trial structure of the STF task. Lower traces show continuous running speed traces of multiple trials from an example session. **b.** Schematic representation of the head-fixed Non-navigational target foraging (NTF) task in which mice move the joystick to variable locations in search of a target area (distance window). Brief occupancy of the target area (~100 ms) yields a delayed water reward through a spout positioned in front of the head position. Upper right inset shows the trial structure of the NTF task. Lower traces show continuous joystick speed traces of multiple trials from an example session. **c,g.** Change in path length of foraging runs as a function of block in STF task (c; close (blue); further (green)) and change in path length of joystick movements in the NTF task (g; close (blue), far (green), intermediate (purple), close (red)). **d,i.** Errors (movements that failed to intersect target) per correctly executed STF trial (d) or NTF trial (i). Dots indicate the first and last trials of a block; colors as in c,g. **e,h.** Trajectories taken on each trial of the first (blue) and second (green) block for an example STF session (e) and joystick amplitude trajectories from NTF task (h) as a function of target location blocks. Thick lines indicate mean trajectories, colors as in c,g. **f,j.** Latency distribution of attempted trial initiating movements after previous reward for all sessions of STF (f) and NTF (j) tasks. Gray area indicates a time out period (unsigned) in which no reward could be obtained. **k.** Comparison of within block (all block average) change in errors for NTF (cyan) and STF (red) tasks overlaid for comparison. All shaded areas indicate the standard error of mean.

Fig. 2| A common learning model accounts for behaviors in navigational and non-navigational contexts

a. Schematic representation of the tML model (see Methods). The model assumes an agent whose foraging trajectory is governed by two dynamical outputs: heading (pink) and speed (cyan). The time evolution of the heading and speed control signals is assumed to be the product of dynamics in a recurrent neural network or can be simulated by specifying an analytic function $\{\Theta(t) \& S(t)\}$ (see methods) as indicated and used for simulations herein. A learning rule which is a variant of the MeSH rule is used to update the heading offset and speed scaling dependent upon performance feedback ('reward', red). **b.** Examples of the time-varying heading (pink) and speed signals (cyan) plus a learned shift (Δ). **c.** Trajectories produced by the control signals (and their shifted versions) shown in b. **d.** The tML model was tested in simulated version of the STF and NTF task. First a schematic of the STF task is shown with trajectories from an example simulation across the close (blue) and far (green) blocks are illustrated. Heavy line indicates the mean and thin traces are individual trials. **e.** Summary statistics of the trajectory amplitude (upper) and error rate per trial (lower) are plotted similarly to data plotted in Fig. 1c,d. Note: for these simulations every trajectory attempt is plotted whereas for data only successful trials are shown. **f.** Grid search was used to identify optimal model parameters (learning rate and exploration scaling) that minimized difference between simulation (pink) and observed data (black). **g.** Schematic of NTF task simulation and example trajectories as in d. Colors in lower simulated movements are the same as Fig. 1g. **h.** Performance statistics for simulations of the NTF task showing trajectory length (upper) and error rate (lower) that can be compared to data in Fig. 1g,i. **i.** Best fit model (pink) performance compared to behavioral data (black) for the block 1 to block 2 (blue to green) transition. Shaded area indicates the standard error of the mean in all plots.

Fig. 3| Sequential activity of dCA1 neural ensembles in navigational and non-navigational contexts

a,h. Example imaging field of view (standard deviation projection of $\Delta F/F$ image) for STF (a) and NTF (h) tasks with overlaid ROI detection in red. Activity traces of several example ROIs are shown below. **b,i.** Number of ROIs recorded per session in STF (b) and NTF (i) tasks. **c,j.** Mean detected calcium event rate

(left) and maximum inferred firing rates (right; for spike inference details see methods) per ROI in STF (c) and NTF (j) tasks. **d,k.** Mean (\pm s.e.m.) movement speed (black) and mean reliability of calcium activation profile (red) in STF (d) and NTF (k) tasks. **e,l.** Halfwidth of suprathreshold calcium events within trials in STF (e) and NTF (l) tasks. **f,m.** Sequential alignment of normalized calcium profiles for a held out half of trials for all ROIs in all sessions imaged in STF (f) and NTF (m) tasks. Mean speed trajectories as in d,k plotted below for comparison. **g,n.** Imaging data for a run of 14 trials during performance of STF task (g) and NTF task (n). ROIs were sorted by first sorting ROIs above and below median activation then within those groups sorting by latency of peak activation from movement onset. Average of example data for all trials within the session shown at right. Lower traces show corresponding movement speed traces and successful reward triggers indicated by asterisk.

Fig. 4| *Decoding of behavior trajectory and classification of task epoch with dCA1 ensemble activity*

a. Example traces showing amplitude (upper) and angle (lower) of concatenated foraging runs (black traces) and the performance of a linear ensemble decoder trained on simultaneous dCA1 activity for a single session of STF task (red traces). **b.** Summary data showing consensus decoder (see Methods) performance (coefficient of determination; R^2) in the STF (red, left) and NTF (cyan, right) tasks compared against a shuffled (permuted decoder weights) control (gray). **c.** Decoder performance in the STF task depended significantly upon the number of ROIs imaged in a given session. **d.** Cross validate performance of ensemble classifiers trained on active dCA1 population vectors to discriminate the task episodes (different target locations) from which a given trial was taken for all sessions in the STF (red, left) and NTF (cyan, right) tasks.

Fig. 5| *Distinct timing of synchronous population events in navigational and non-navigational contexts*

a. Example data from a single session of STF task. Upper trace indicates the speed of foraging run. Raster plot shows inferred spikes from ROIs sorted by average latency relative to trial initiation. Lower trace indicates the framewise summed spike across the entire imaged population. Red dashed lines indicate detected SPEs in the population activity. Blue dashed lines indicate when a reward was triggered. **b.** Mean position profile (cyan) and probability of observing an SPE (red; $P(\text{SPE})$) for all sessions of STF task. Shaded area indicates the s.e.m. **c.** Mean cross-correlation between displacement from reward port and summed population activity. At negative time lags population activity leads and at positive time lags trajectory speed leads. Shaded area indicated s.e.m. **d.** Count of SPEs plotted as a function of position within the arena in the STF task. **e.** Fraction of successful trials across the entire session plotted against max $P(\text{SPE})$ per trial for all sessions of STF task. Colors indicate individual animals. **f-j.** Analogous plots as in a-e but shown for NTF task performance. Note the reversed phase at which peak $P(\text{SPE})$ occurs between the two tasks. P values in e,j indicate significance testing for the pearson's correlation coefficient.

Fig. 6| *Optogenetic inactivation of dCA1 produces dissociable effects in navigational and non-navigational contexts*

a. Schematic of freely moving navigational STF task and scheme for optogenetic inactivation experiments. Optogenetic inactivation was achieved by delivering blue light through a flexible fiber above dCA1 in VGAT-ChR2-EYFP mice (N=4). Inactivation was delivered only on catch trials ($\leq 30\%$ of trials). **a1.** Blue light was delivered in closed-loop upon entry to the collection area (reward consumption) and ceased upon exit from the collection area - the time epoch in which SPEs were observed. We saw no overt reduction in movement. To quantify changes in performance we compared the latency to initiate the next trial ('latency', n.s.), probability of correctly completing the subsequent trial (' $P(\text{corr})$ ', n.s.), the variance in heading at beginning to next trial (' $\text{var}(\Theta)$ '; significant reduction; ANOVA), and the max distance travelled to on subsequent foraging run ('out dist.'; significantly increased; ANOVA). Note: significant differences between animals were also observed in all measures. **a2.** Blue light was delivered in closed-loop upon leaving the collection area and ceased upon exit from the target area - task phases

with robust place cell activity, but when SPEs were rarely observed. No overt reduction in movement was observed. To quantify changes in performance we compared the time required to complete the trial ('duration', n.s.), probability of correctly completing the subsequent trial ('P(corr)', n.s.), the variance in heading at return to collection area ('var(Θ)', n.s.), and the distance travelled to complete foraging run ('in dist.', n.s.). **b.** Schematic of the head-fixed NTF task used for optogenetic inactivation experiments (N=3). Preparation was similar as a. However, optogenetic inactivation was achieved by delivering blue light from a fiber through the imaging cannula preparation (lacking a camera and a grin lens). On individual catch trials ($\leq 30\%$ of trials) inactivation was delivered in closed-loop 1 (b1) or 0 (b2) sec after reward delivery (times at which SPEs were observed) or a visible blue light flash outside the brain was delivered as a distractor (b3). For each experimental condition histogram of movement initiations per trial are shown for control (black) and inactivation (cyan) conditions. Black dashed lines indicate beginning and end of blue light. Purple line indicates end of intertrial interval (uncued trial start, see Fig. 1 for task details). Shaded lines in b1-b3 are the mean movement rates and 0.025-0.975 confidence intervals by 1000 shuffling. Red lines indicate times in which a significant difference between control and inactivation were observed (permutation test). **c.** Peak velocity of movement plotted for movement initiated during long inactivation condition (b2; 'Inact.') and control, trial-initiating movements ('Ctrl'). No significant difference in maximum velocity (Wilcoxon signed-rank test) was detected.

Fig. 7 | Blockade of policy updates in tML agent can explain effects of optogenetic perturbations

tML model architecture (Fig. 2) is extended to incorporate context-specific roles of CA1 SPEs for the STF (a-e) and NTF (f-j) task variants. **a-c.** tML model (a) in which CA1 SPE activity (c) occurs at the end of a trajectory and used for the learning (MeSH rule) update comparison between the trajectory for the current trial $\{a[i], \omega[i]\}$ and the policy $\{A[i], \Omega[i]\}$ (b) (see Methods). **d.** Return heading angle (in radians) for the current trial is plotted as a function of the loading of the SPE ensemble for that trial onto the first principal component of SPE ensemble matrix (PC_{spe}) for an example session. Absolute Pearson correlation ($|\rho|$) between Return heading angle (Θ) and PC_{spe} loading for all sessions. **e.** Simulations in which inactivation of CA1 prevented SPEs was modeled as the absence of an update and reversion to default heading and speed vectors for 50 simulations with independent initializations (gray points). Median values for simulations are shown in cyan points. Plots are the same quantities as shown in experimental data in Fig. 6a1; left, variance in initial heading in trials after inactivation ('Inact. var(Θ)') vs trials without inactivation before or after ('Control var(Θ)'); right, amplitude of run following inactivation ('Inact. out dist.') vs control ('Control out dist.'). **f-h.** tML model (f) in which initiation of a foraging trajectory depends upon an intertrial hazard function and the occurrence of a CA1 SPE (g) and is used for the learning (MeSH rule) update comparison between the current trial command $\{a[i], \omega[i]\}$ and the policy $\{A[i], \Omega[i]\}$ (h) (see Methods). **i.** Histogram of absolute Pearson correlation ($|\rho|$) between Block ID vs PC_{spe} loading (left) and estimated $A[i]$ vs PC_{spe} (right) are plotted for all sessions. **j.** Simulations of a tML model in which initiation is governed by hazard \cap SPE (described in f-j and Methods) for control (black) and simulated inactivation (cyan) of 3 (left) and 5 (right) seconds as in Fig. 6b1 and 6b2. Significant correlation scores ($p < 0.05$; in d,i) are represented with shaded bars.

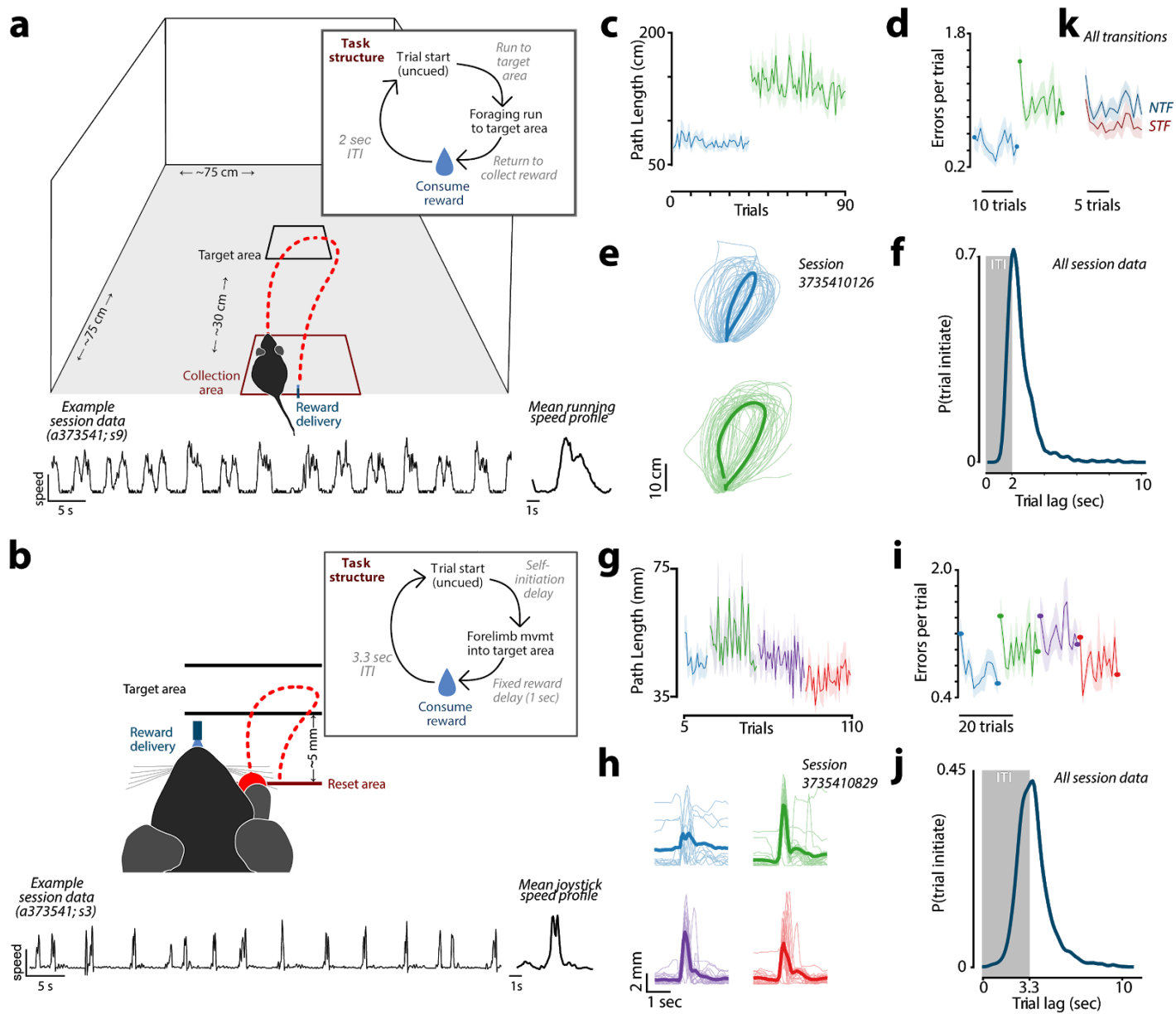


Figure 1

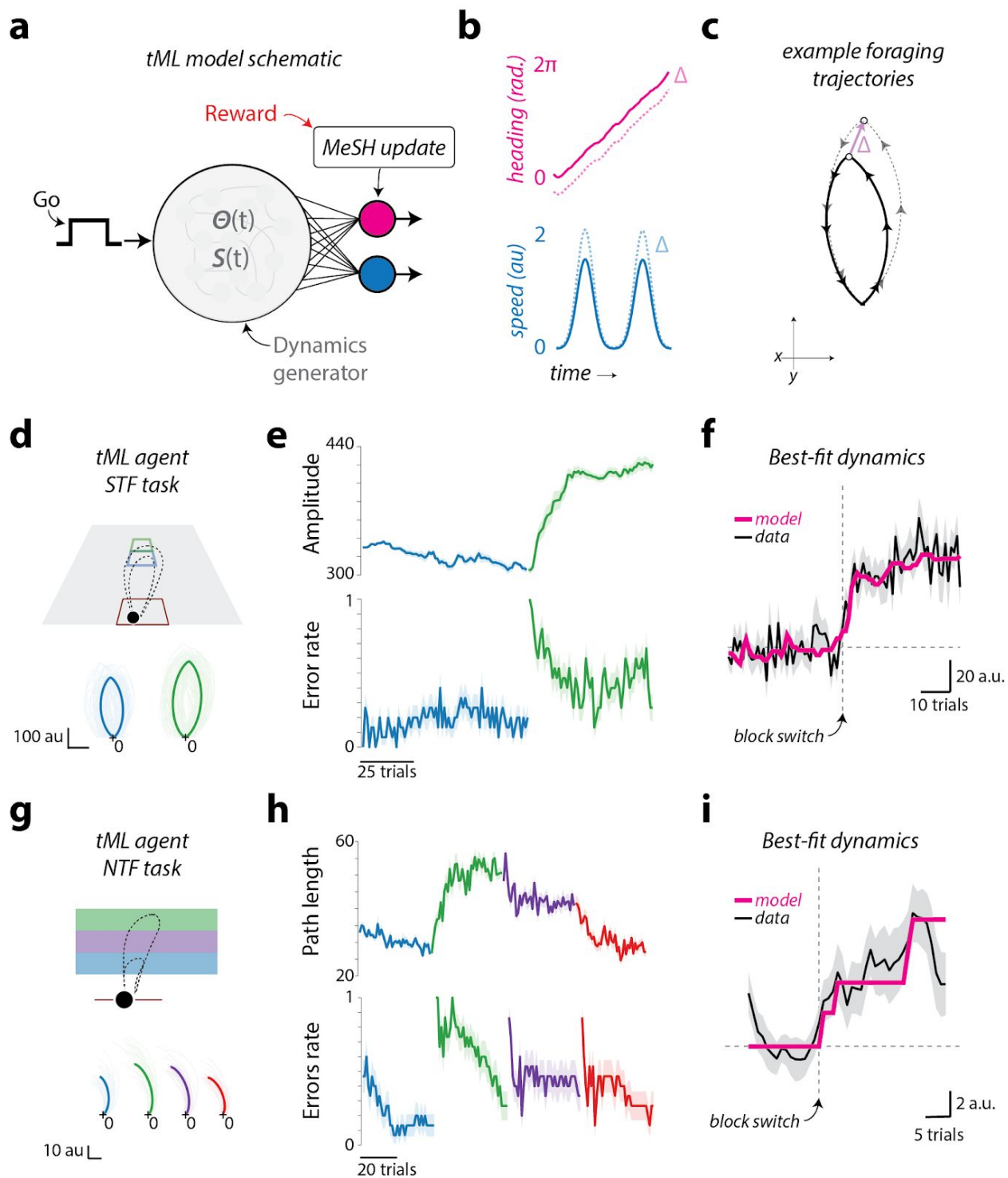


Figure 2

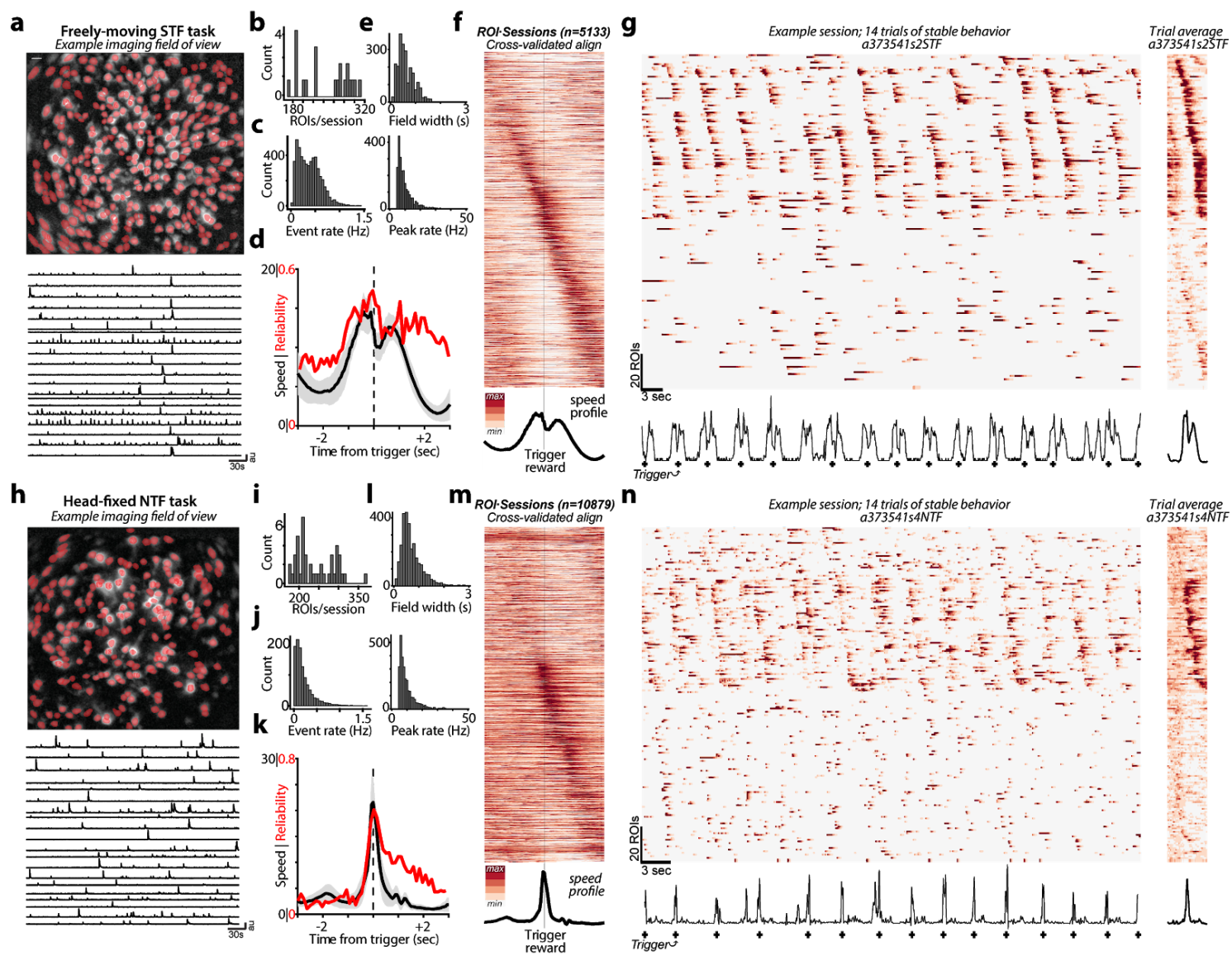


Figure 3

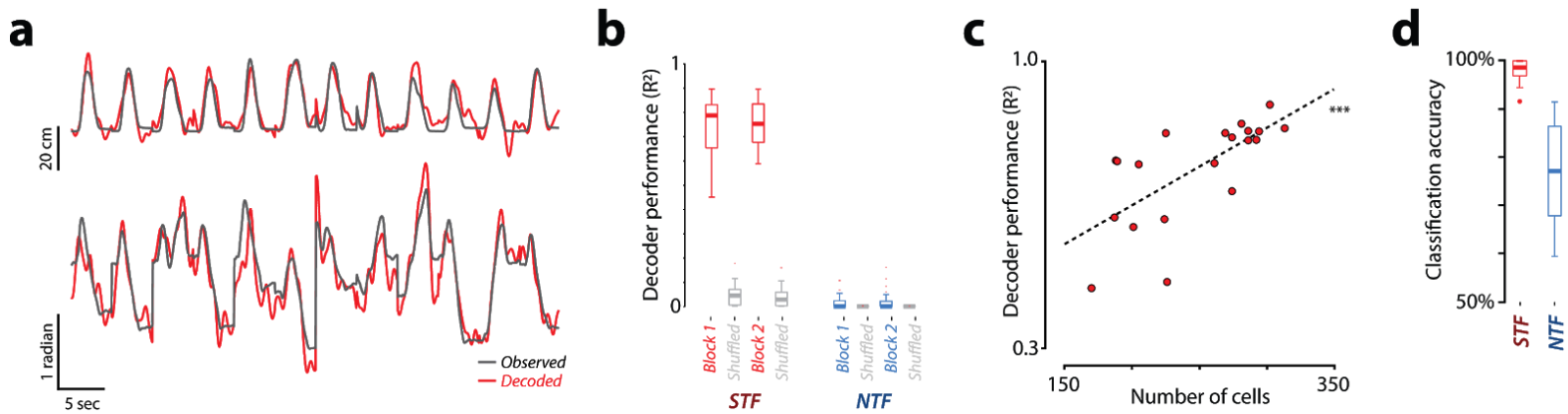


Figure 4

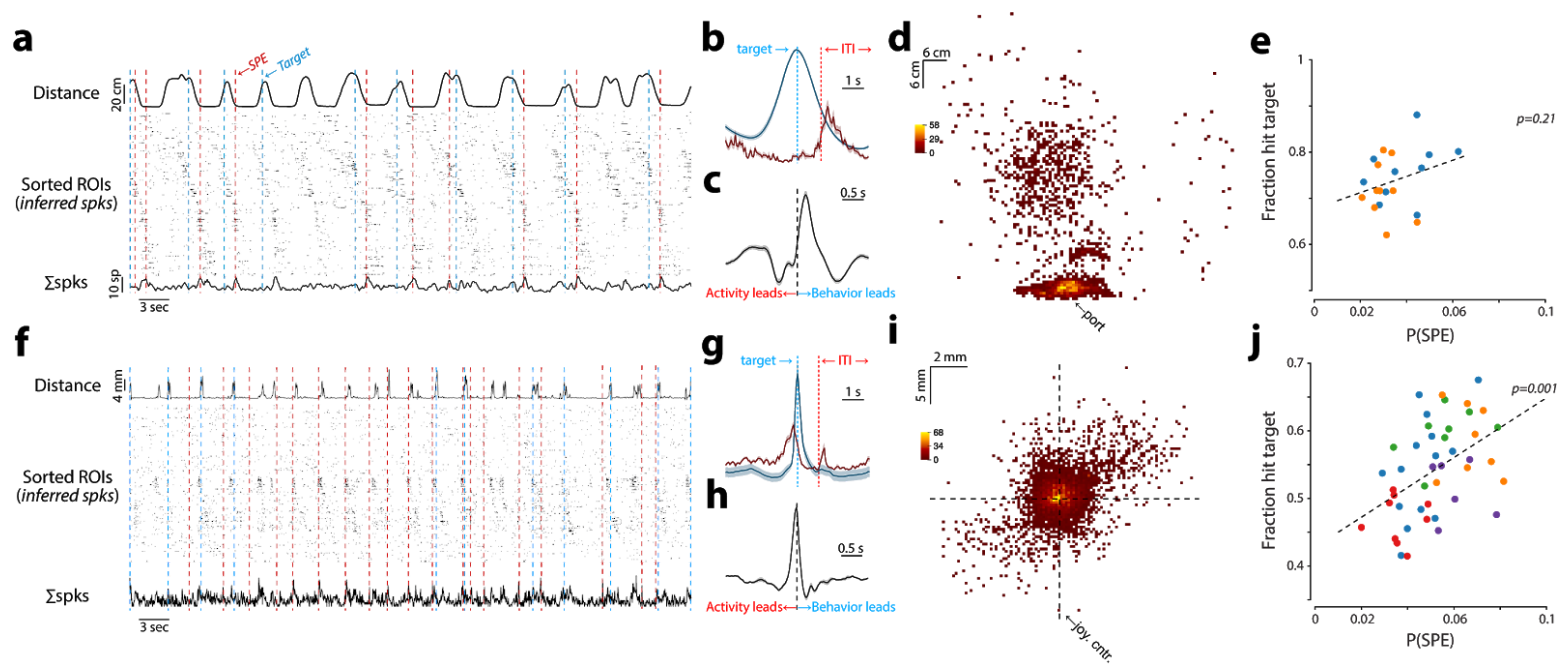


Figure 5

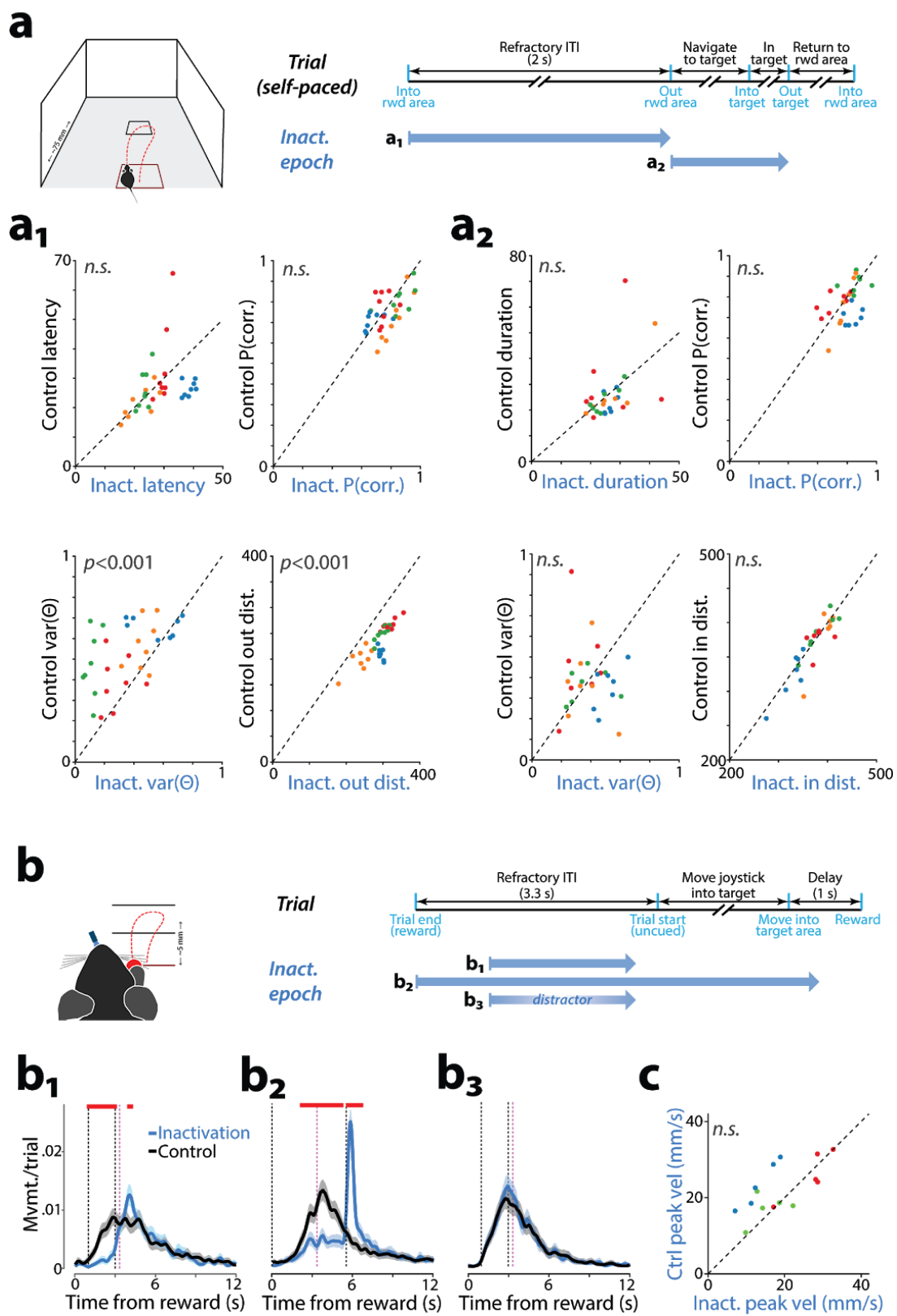
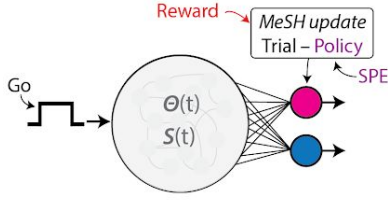


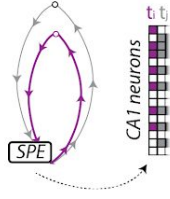
Figure 6

Spatial Target Foraging

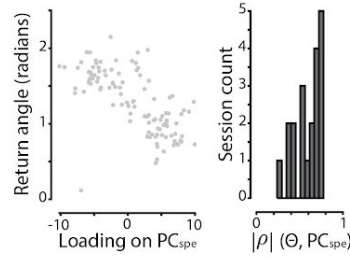
a. tML model schematic



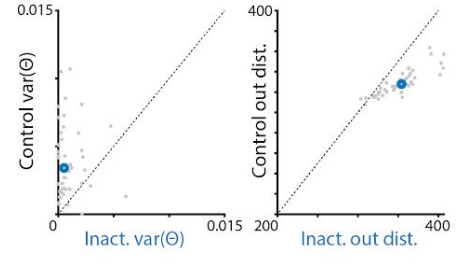
b. trajectories c. activity



d. SPE ensemble correlates

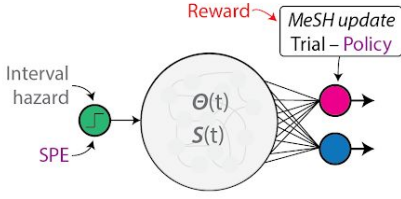


e. tML simulation of optogenetic inactivation

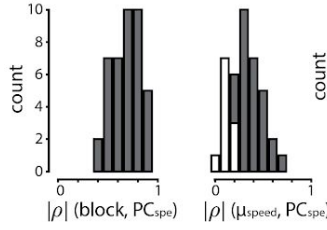
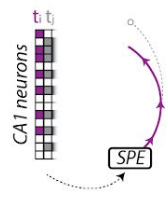


Non-navigational Target Foraging

f. tML model schematic



g. activity h. trajectories i. SPE ensemble correlates



j. Simulation of optogenetic inactivation

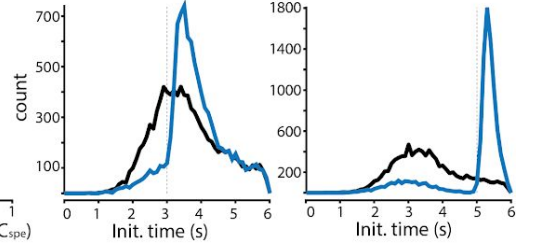
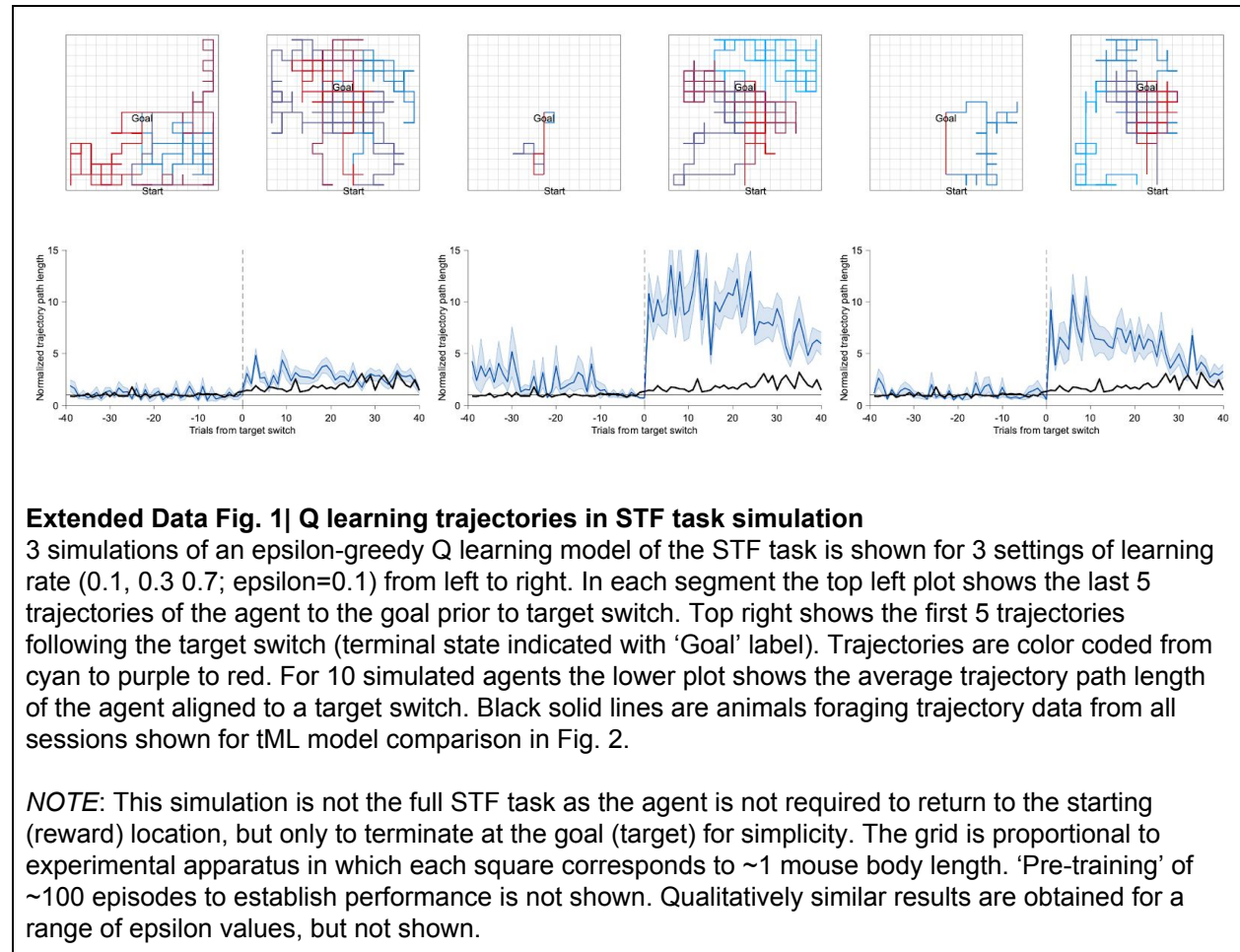


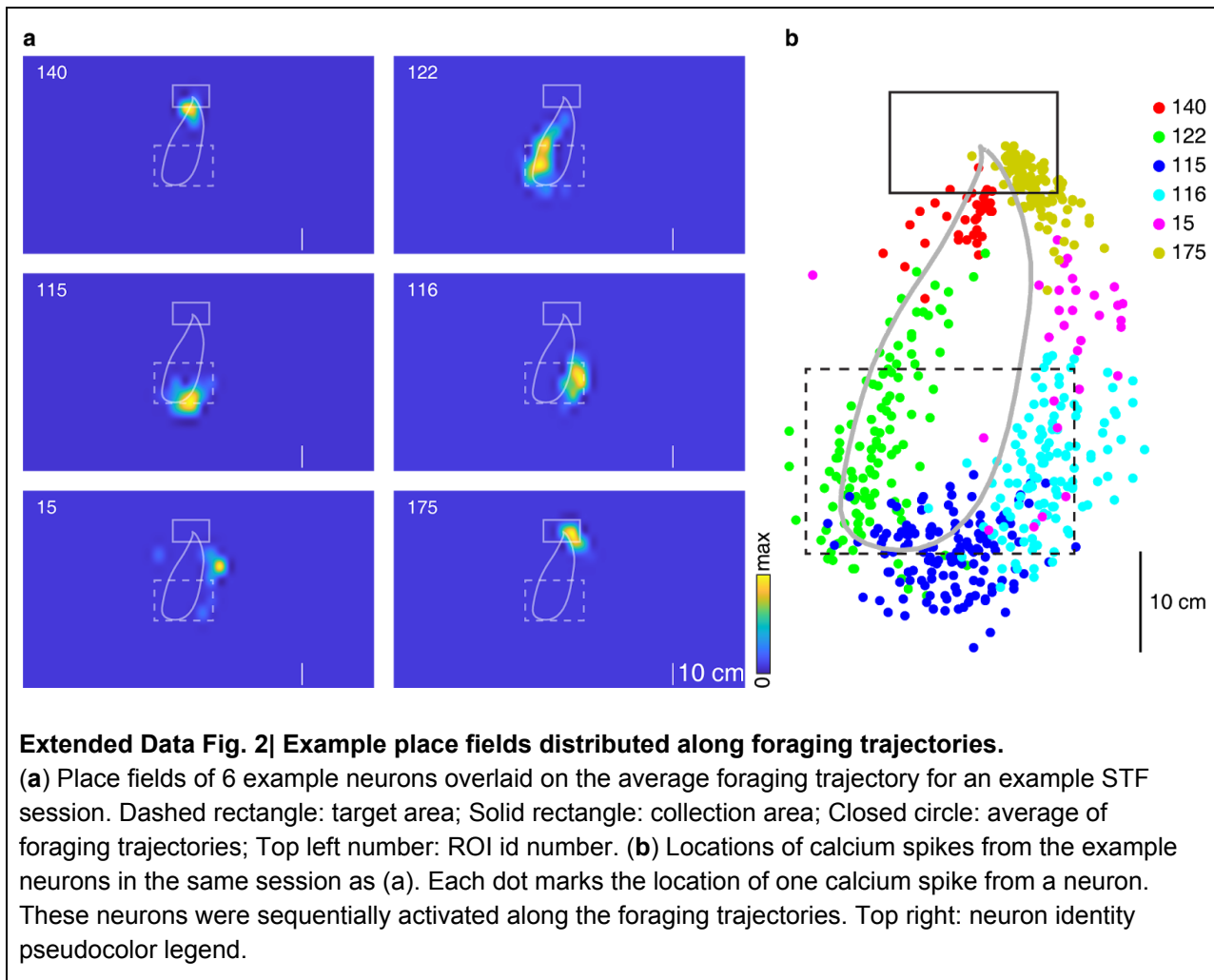
Figure 7

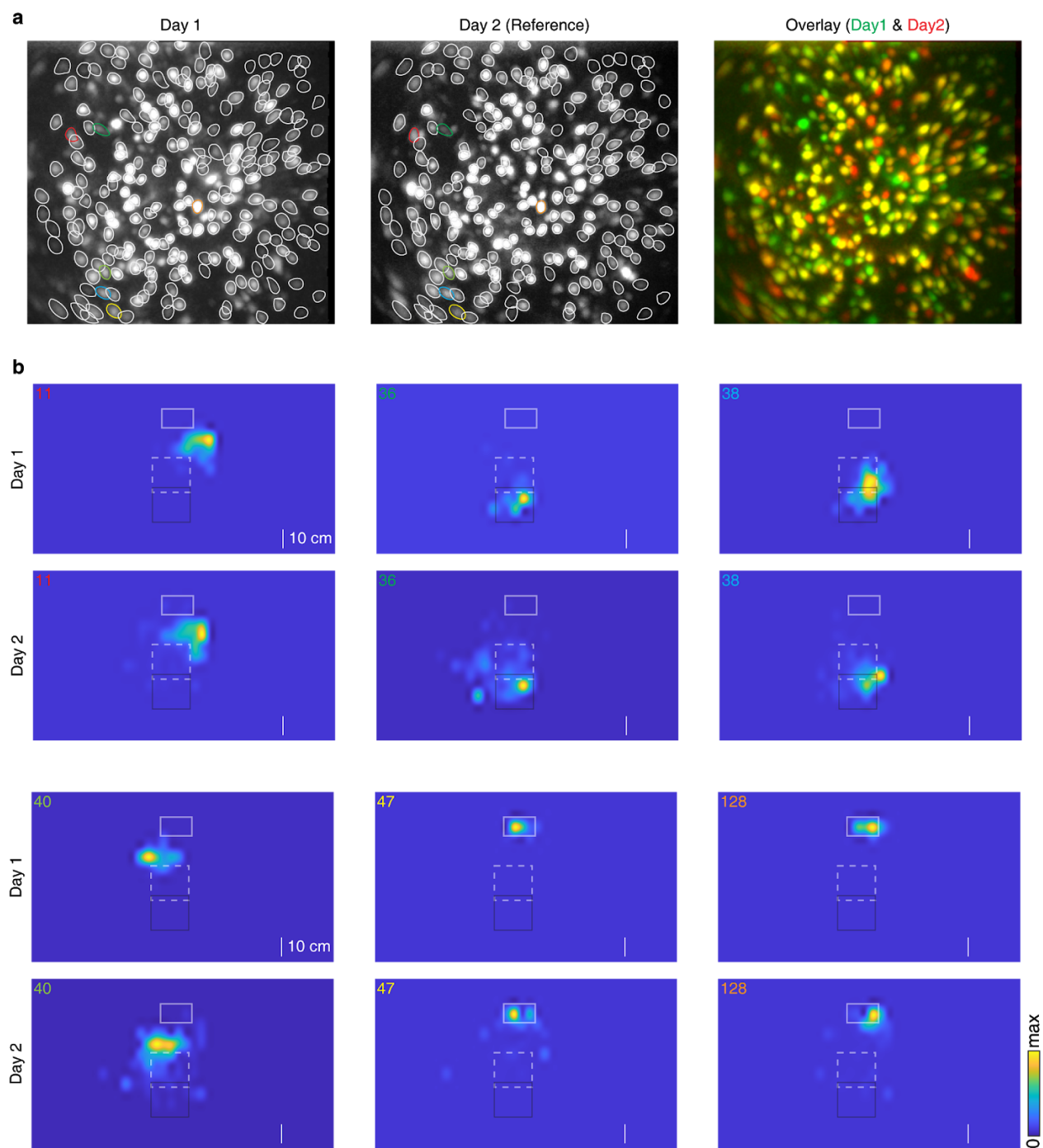
Supplementary Video 1

STF task recorded with 2 cameras to illustrate real-time tracking performance.

Extended Data Figures



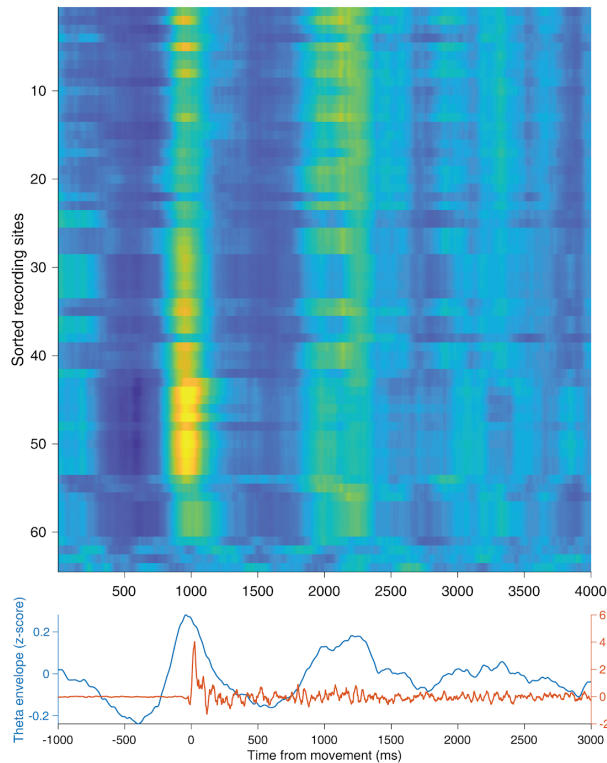




Extended Data Fig. 3| Stable imaging of individual dCA1 neurons across days.

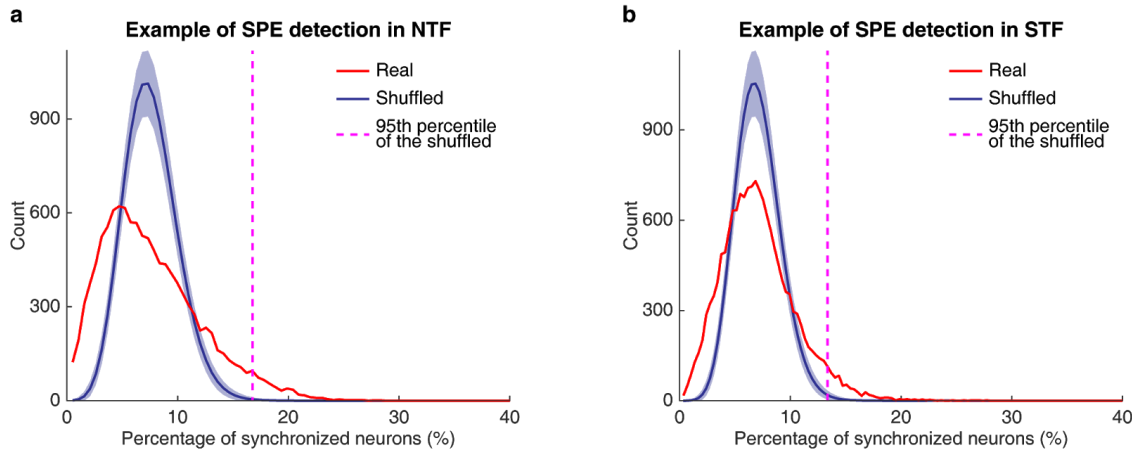
(a) Example response images (maximum intensity projection - average intensity projection) of the same field of view across days. Left: aligned response image on Day 1 overlaid with the ROIs detected on both days; Middle: reference response image on Day 2 overlaid with ROIs detected on both days; Right: pseudocolor overlay of the response images on both days. (b) Stable place fields of 6 example neurons

across days. Dashed white rectangle: target area1; Solid black rectangle: target area2; Solid white rectangle: collection area; Top left number: neuron identity with the same color as that of its ROI in (a).



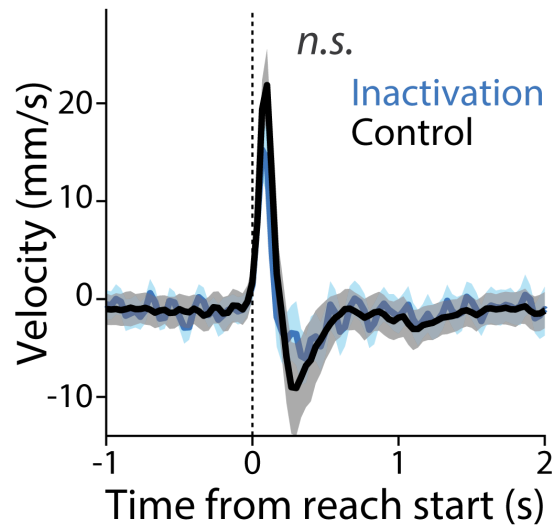
Extended Data Fig. 4|. Forebrain activity was modulated at theta frequency in NTF task

Data from a 64 channel probe recording from dorsal striatum in the joystick task from previously published data ³⁰. The mean amplitude envelope of filtered LFP signal (bandwidth: 4-12 Hz) is plotted for all 64 channels averaged across all joystick movements. Lower plots show the mean amplitude envelope change across all trials, all channels (cyan) compared to the mean joystick movement (red; sampled at 1kHz). Note: second peak in theta envelope is around the time of mean reward delivery (1 second after threshold crossing).



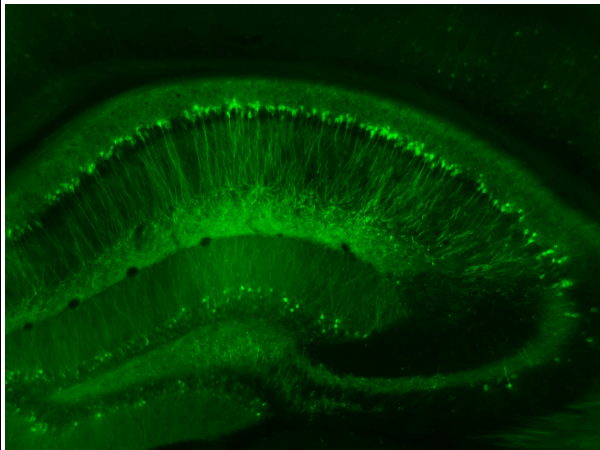
Extended Data Fig. 5| Detection of SPEs using a shuffling approach.

(a) Histograms of synchronized neurons in an example NTF session and in its corresponding shuffled datasets. The shuffled datasets were generated by shuffling the spike-timing but keeping the interspike intervals for each neuron and repeating the shuffling procedure 1000 times. Shaded blue line: mean \pm s.e.m of the 1000 shuffled datasets; Magenta dashed line: statistical threshold for SPEs detection. Any imaging frame with more synchronized neurons than the threshold was detected as a SPE (Fig. 5f). (b) same as (a) but for an example of STF session.

a

Extended Data Fig. 6| Effects of long inhibition at SPEs period on movement velocity.

(a) Comparisons of movement velocities with and without inhibition in the long SPE inhibition experiments. Shaded line: mean and 0.025-0.975 confidence interval by 1000 shuffling. n.s.: no significance (permutation test).



Extended Data Fig. 7| GCaMP-expressing hippocampal neurons in Ai93(TITL-GCaMP6f)-D;ROSA26-ZtTA×Kcnd2-IRES-Cre 3G5 mouse.
Example image of GCaMP6 expression in dorsal CA1 neurons.