ORIGINAL ARTICLE

Conjunctive reward–place coding properties of dorsal distal CA1 hippocampus cells

Zhuocheng Xiao[1] · Kevin Lin[1,2,4] · Jean-Marc Fellous[1,3,4]

Abstract

Autonomous motivated spatial navigation in animals or robots requires the association between spatial location and value. Hippocampal place cells are involved in goal-directed spatial navigation and the consolidation of spatial memories. Recently, Gauthier and Tank (Neuron 99(1):179–193, 2018. https://doi.org/10.1016/j.neuron.2018.06.008) have identified a subpopulation of hippocampal cells selectively activated in relation to rewarded goals. However, the relationship between these cells' spiking activity and goal representation remains elusive. We analyzed data from experiments in which rats underwent five consecutive tasks in which reward locations and spatial context were manipulated. We found CA1 populations with properties continuously ranging from place cells to reward cells. Specifically, we found typical place cells insensitive to reward locations, reward cells that only fired at correct rewarded feeders in each task regardless of context, and "hybrid cells" that responded to spatial locations and change of reward locations. Reward cells responded mostly to the reward delivery rather than to its expectation. In addition, we found a small group of neurons that transitioned between place and reward cells properties within the 5-task session. We conclude that some pyramidal cells (if not all) integrate both spatial and reward inputs to various degrees. These results provide insights into the integrative coding properties of CA1 pyramidal cells, focusing on their abilities to carry both spatial and reward information in a mixed and plastic manner. This conjunctive coding property prompts a re-thinking of current computational models of spatial navigation in which hippocampal spatial and subcortical value representations are independent.

Keywords Reward · Autonomous spatial navigation · Place cells

Communicated by Benjamin Lindner.

This article is part of the special issue entitled 'Complex Spatial Navigation in Animals, Computational Models and Neuro-inspired Robots'.

✉ Jean-Marc Fellous
fellous@email.arizona.edu

1 Program in Applied Mathematics, University of Arizona, Tucson, AZ 85721, USA

2 Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA

3 Department of Psychology, University of Arizona, 1503 E University Blvd, Suite 312, Tucson, AZ 85721, USA

4 Neuroscience Graduate Interdisciplinary Program, University of Arizona, Tucson, AZ 85721, USA

1 Introduction

Be they animals or robots, autonomous agents in ever-changing environments must make dynamic decisions as to where to go next, given a set of goals and an understanding of their surroundings. The ability for animals to explore space in order to reach goals, for example, rewards such as food or water, is essential for survival. Much work has been done in the restricted cases where such an agent has a single task (i.e., retrieve an injured individual). It was proposed early on that animals and robots alike might build a map of their environment in which specific locations are given specific values (Lieblich and Arbib 1982; Guazzelli and Arbib 1997). The neurophysiological nature of the pairing between space and value has, however, remained elusive.

It has long been known that the hippocampus and anatomically related brain areas provide a physiological basis for both memory consolidation and spatial navigation in mammals like rats and mice (Olafsdottir et al. 2018). O'Keefe and

colleagues first established that some CA1 cells exhibited spatially dependent firing rate maps (O'Keefe and Dostrovsky 1971; Smith and Mizumori 2006; Moser et al. 2015). These hippocampal pyramidal cells with spatially sensitive receptive fields (or "place fields") were responsive to specific locations in a given environment. Outside the hippocampus, several other areas have been found to be involved in spatial navigation. Head direction cells in the presubiculum and entorhinal cortex encode the orientation of the head of the animal in the horizontal plane (Mizumori and Williams 1993; Taube 1995). Grid cells in the medial entorhinal cortex and subicular complex have some similarities with place cells but exhibit multiple place fields aligned on a triangular grid which covers at least part of the environment (Hafting et al. 2005; Boccara et al. 2010). Boundary cells, found in the hippocampal formation, i.e., the subiculum, presubiculum, and entorhinal cortex, encode the existence of a boundary at a specific angle and distance (Lever et al. 2009; Barry et al. 2006; Solstad et al. 2008; Boccara et al. 2010). Finally, object vector cells have recently been found in the medial entorhinal cortex and respond to specific directions and distances around spatially confined objects, regardless of their locations (Hoydal et al. 2019). Current theories hold that together, these cells form a value-neutral map of the environment.

Place cells often change their preferred firing location, or "remap," as a result of significant changes in the environment (Wilson and McNaughton 1993; Lenck-Santini et al. 2005; Zhang and Manahan-Vaughan 2015). Significant changes of spatial context are always accompanied by global and apparently random remapping for almost all recorded place cells (Wilson and McNaughton 1993; Schlesiger et al. 2018), and their firing rate can be modulated by small spatial context changes such as adding or subtracting a fraction of spatial cues (Colgin et al. 2008), or changing task demands (Sanders et al. 2019). The mechanisms by which hippocampal information can be preserved and re-instantiated from one environment to the next, across multiple environments, are still theoretically unknown. One possibility is that some cells in hippocampus encode or are strongly influenced by important spatial features, such as the location of rewards, but insensitive to environmental changes. These "importance" cells would act as invariant "anchors" that would index global representations. An alternative is that hippocampal cells are influenced by spatial information and reward information simultaneously. So far, several properties of rewards have been shown to affect population activities (Dupret et al. 2010; Poucet and Hok 2017; Singer and Frank 2009). For instance, as a result of learning, place fields tend to accumulate around the locations of rewards (Dupret et al. 2010). Place cells with place fields away from reward locations could also display excess firing around reward sites (Poucet and Hok 2017).

Interestingly, it has been shown that some cells may exhibit two or more spatial properties simultaneously. These

so-called conjunctive cells indicate cross-talk and integration between the different components of the spatial navigation system. For example, posterior parietal cortical neurons with conjunctive egocentric and allocentric properties have been found (Wilber et al. 2014). In medial entorhinal cortex, position x head-direction conjunctive cells have been proposed to control the network dynamics of a periodic attractor map (Bush et al. 2015; Navratilova et al. 2012), and the place field patterns of grid cells can be affected by both the environment and the velocity of animals (Sargolini et al. 2006). Other studies have shown that hippocampal cells respond to flashes of light (Liu et al. 2018) or the conjunction of place and object location (Deshmukh and Knierim 2013). These single-cell multi-coding properties suggest that there might be neurons in the spatial navigation system capable of carrying spatial and reward information in a conjunctive manner. Motivated by this hypothesis, we studied the coding properties of CA1 hippocampal cells at or near carefully manipulated reward sites and spatial contexts.

The most direct evidence of the relationship between single hippocampal cell activity and reward is given by the recent work of Gauthier and Tank, identifying a dedicated neuron population that explicitly encode reward locations in a virtual environment (Gauthier and Tank 2018). This study found that a small fraction (~4.4%) of pyramidal cells (with place fields in all environments) fired only around reward sites. For cells in this population (called "reward cells"), their place fields shifted with changes of reward location but did not remap during context change of the environment; while place fields of most place cells stayed at the same location during reward location changes and exhibited global remapping after context changes. Furthermore, the study found that one-third of the reward cells predicted the reward. Gauthier and Tank observed a sharp coding properties boundary between place cells and reward cells: All cells active in the two different environments maintained their identity, with no place cells becoming reward cells or vice versa. However, this study used calcium imaging and might have given more weight to high-firing cells. Also, the use of a unidirectional infinite virtual reality one-dimensional corridor may have constrained reward coding properties to a restricted range of the overall possible coding space (Holscher et al. 2005). Overall, it is possible that in a more realistic open-field- type environment in which animals are free to choose movement direction, reward coding properties become richer and lower firing cells can show additional properties within the place-reward spectrum.

Following Gauthier and Tank, we used a set of tasks in which remapping was produced by large changes in spatial context and in which animals were required to learn to obtain rewards from different sets of feeders in different behavioral epochs. By controlling the locations of reward delivery and the remapping, we were able to find reward and place

cells as in Gauthier and Tank. Furthermore, we also found a large population exhibiting conjunctive coding properties that responded in a mixed manner to reward and place.

2 Methods

2.1 Animals

Four adult (6–7 months old) male Brown Norway rats were used in this study. All procedures were approved by the University of Arizona IACUC and followed the NIH Guidelines. Animals were food restricted to 85% of their ad libitum weights and were housed on a reversed 12/12 h cycle. Animals were used in the dark cycle.

2.2 Apparatus and task

Rats were trained in a circular maze with a diameter of 1.5 m, using methods previously described (Jones et al. 2012, 2015). The maze's spatial context included experimenter-chosen odor, floor texture, and 2–3 distal visual cues attached to curtains surrounding the maze. Eight feeders, each with one blinking LED attached, were placed in the maze equidistantly (Fig. 1b). An overhead camera tracked the rat position on the maze (20–25 Hz frame rate), and the feeders and lights were automatically controlled by in-house software (Lab-View, National Instrument).

Five tasks were carried out successively in each experimental session, termed Random1, Set1, Set2, RecallSet1, and Random2 (Fig. 1a). The rat rested on a towel-lined flower pot for 30 min in the center of the maze before and after each task. In Random1 and Random2, all LEDs blinked 15 times each and in a pseudo-random order, cuing the rat to the corresponding feeder, with reward delivered every time after the rat visited the correct blinking feeder. Set1 and Set2 epochs were subdivided into learning and testing phases. In the learning phase of Set1, the rat was cued to run to a selected subset containing only three out of the eight feeders (e.g., 1, 4, and 5). After the rat correctly visited those feeders 25 times each, the task transitioned to the testing phase. In this phase, the light cues were delayed by 15 s (a feeder-to-feeder travel typically lasted 3–5 s). The rats, therefore, were able to visit and collect rewards from the three possible target feeders using memory alone (no cues). The light blinked only if the rat timed out after 15 s. In this phase of the experiment, the behavior of the rat could be of 2 types to obtain a reward. (1) If the rat was at an incorrect feeder, the rat could choose any of the three correct feeders and had a 33% chance of obtaining the reward. (2) If the rat was at a correct feeder, it had two choices, each with a 50% probability of obtaining the reward. The test phase ended (i.e., the set was considered learned) when the rat visited 15 correct feeders in a row

with no more than two timeouts. A similar task was repeated in Set2, with three different feeders (e.g., 3, 6, and 8). In RecallSet1, the rat was cued to recall Set1 by a single blinking light and continued as in the testing phase of Set1 (i.e., memory driven) without any cue.

The context of the maze was manipulated to address the relationship between the coding of reward and the coding of the environment. We carried out three types of sessions: (1) five "same-context" sessions where Set1 and Set2 were learned in the same spatial context (i.e., no context change, Fig. 1a). In these sessions, the any remapping of cells' place fields would therefore only be due to the change in reward locations; (2) five "different-context" sessions, where Set1 and Set2 were learned in different contexts, hence remapping might be due to either change of reward location or change of context; (3) two control sessions where Set1, Set2, and RecallSet1 epochs were replaced by epochs in which the rat remained on the flower pot (Random1 and Random2 were conducted). In all sessions, RecallSet1 was carried out in the same context in which Set1 had been learned.

2.3 Surgery

All rats were implanted with a hyperdrive supporting 14 tetrodes targeted at the dorsal distal CA1 area of the hippocampus (AP $= -3.8$ mm, ML $= 2.1$–2.6 mm, depth 2.0–2.6 mm, bundle diameter 1.0 mm), using previously published surgical and histological methods (Valdes et al. 2015). Two of the electrodes were used for referencing and were placed at or near the hippocampal fissure. Rats were implanted with an EMG electrode in the neck muscle to ascertain periods of sleep (not used in this study). Rats recovered for a week and were re-trained in the experimental paradigm to habituate them to the weight of the implant.

2.4 Data analyses

2.4.1 Sharp wave ripples (SWR) detection

While most SWRs occurred during non-REM sleep, SWRs could also occur briefly during the task epochs (Fig. 1d). Typically, such SWRs occur when the animal stops walking, especially when it consumes rewards (Fig. 1d, e). This type of awake reactivation may unduly bias the computation of place fields. Therefore, we extracted SWRs from local field potential channels, so as to exclude SWR-mediated spiking when computing place fields (see below).

To extract SWRs, we filtered the local field potentials voltage with a frequency between 100 and 240 Hz (fifth-order Butterworth filter, MATLAB). A Hilbert transform was then computed and rectified. The mean μ and variance σ of these traces were computed. Any epoch with an amplitude larger than $\mu + 3.0\sigma$ (with a probability $\leq 0.27\%$ for a normal dis-
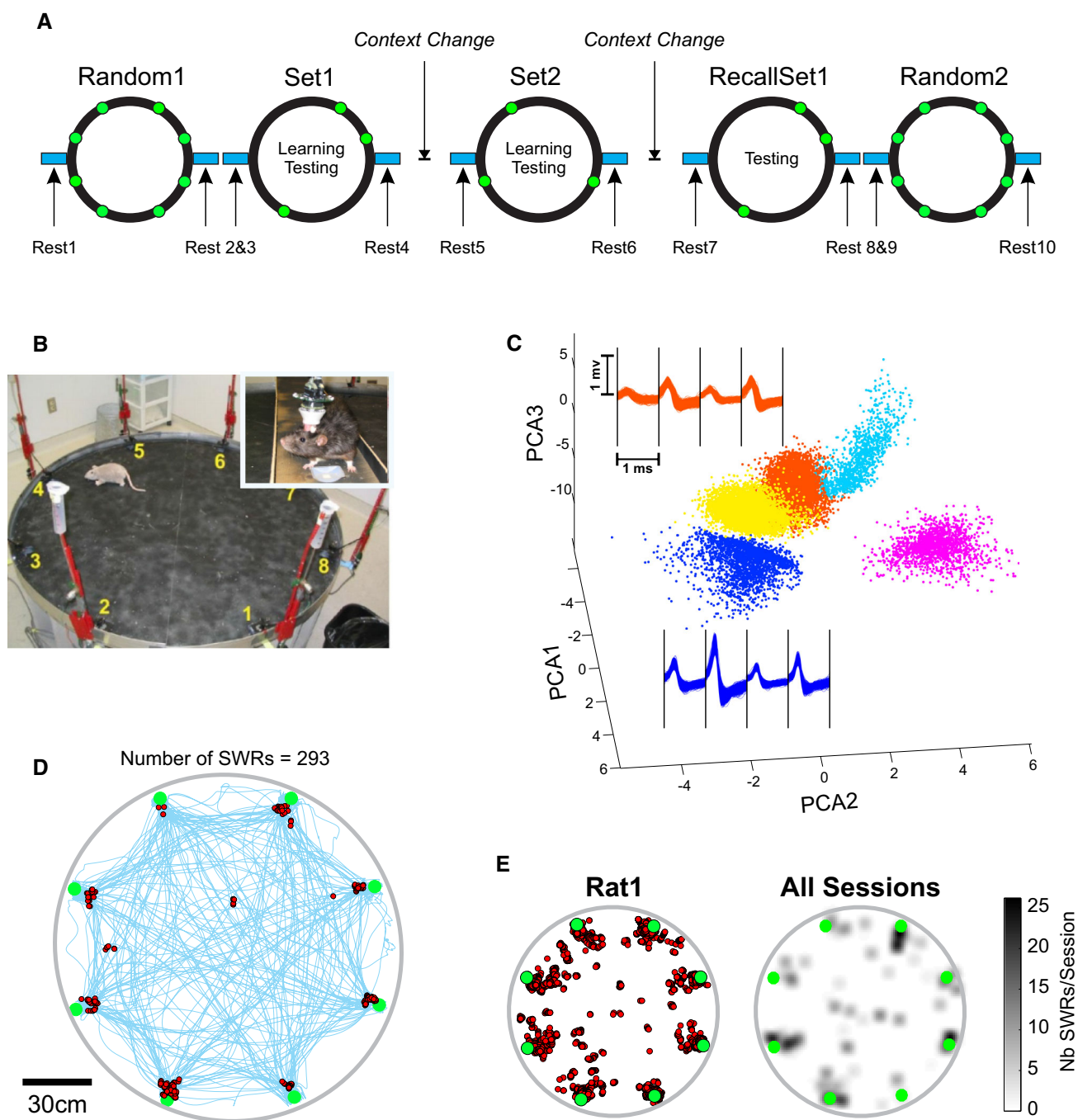
Fig. 1 Behavioral and electrophysiological methods. **a** Behavioral timeline. Set learning was first carried out in context A with 3 feeders (Set1), after which Set2 learning was conducted in either context A (same) or (different) context B with 3 different feeders. Recall of Set1, as well as random sessions, was always in context A. All rest epochs lasted 30 min. **b** Picture of the maze. Each rat was connected to a hyperdrive for electrophysiological recordings with 14 tetrodes (inset). **c** Recorded spikes were sorted and clustered using principal component analysis (PCA) based on the shapes of their waveforms. In this graph, the cyan cluster was a reward cell. **d** Spatial distribution of all SWRs (red dots) recorded during Random1 from one session of Rat1 plotted with tracking data (blue). Awake sharp waves generally took place during immobility, near the feeders. **e** Spatial distributions for all SWRs from Rat1 only (left), and density plot of all SWRs from the entire data sets (right, 4 rats, 12 experiments). Green dots in **d** and **e** are reward sites (color figure online)

tribution), whose duration was between 80 and 200 ms, was designated as an SWR.

2.4.2 Place field computation

Spikes from tetrodes with electrophysiology recording were manually sorted and clustered in the 3 first principal component computed on the basis of their shape (Spike 2, CED and Fig. 1c). Cells were assessed for their theta modulation and for the absence of spikes within their refractory period (< 2%) to evaluate the quality of spike sorting.

For each cell, the place field was computed as the ratio of the firing rate map to the spatial occupancy map. For the occupancy map, the entire space, including the maze, was divided into 53 by 40 bins ($16 \, \text{cm}^2$ per bin). To avoid singularity computations, we deleted the place map bins where animals were there less than 0.08 s. In addition, the track data were filtered for speed > 15 cm/s and the time spent in each bin was computed. For the firing rate map, all spikes occurring with SWRs during the task were deleted, and the number of spikes per spatial bin was computed. Bins with one spike or no spike were set to zero.

Both the occupancy map and rate map were smoothed by a 10-bin Hanning window before calculating the place field map by division. The mean and variance of the place field map were computed. Areas of the map with intensity larger than 1 unit of standard deviation and larger than $288 \, \text{cm}^2$ were considered part of the place field, and an ellipsoid was fitted to these bins.

2.4.3 Place/reward measures/index

Two indexes based on place maps, termed "place score (P)" and "reward score (R)," were used to quantify the extent to which each cell was responding to space or reward location across the five tasks. A perfect place cell should have a consistent place field with a fixed location in different tasks regardless of the reward location, as long as the spatial context stayed the same. In contrast, place field(s) for an ideal reward cell should follow the rewarded feeders as they changed from Random1 to Set1, to Set2, and so on.

When viewing the location of rat, $\vec{x}$, as a random variable of behavior, the place map of cell C in task T, $M^{C,T}(\vec{x})$, is a firing probability function of $\vec{x}$. Therefore, we define the reward score of a cell to be the averaged cosine similarity (or Pearson correlation) between its place maps and the reward-location functions. The Reward-location function in task T, $F^T(\vec{x})$, is the sum of Gaussian surfaces ($\sigma = 5$ cm) centered at the location of correct feeders. That is to say, taking the locations of feeders as $\vec{x}_f$ ($f = 1 - 8$), and the set of correct feeder for each task as $S_1 = S_5 = \{1, 2, 3, 4, 5, 6, 7, 8\}$ (Random1 and Random2), $S_2 = S_4 = \{1, 5, 6\}$ (Set1 and

RecallSet1), $S_3 = \{2, 4, 7\}$ (Set2), then for task T ($i = 1 - 5$) the reward location function is:

$$F^T(\vec{x}) = \sum_{f \in S_i} G_{x_f, \sigma^2}(\vec{x}).$$

Here G_{x_f, σ^2} represents the standard Gaussian distribution function in space. Hence, for task T, the Pearson correlation between cell C's place map $M^{C,T}(\vec{x})$ and $F^T(\vec{x})$ should be the inner product between them normalized by the product of norms:

$$R^{C,T} = \frac{\int_{\vec{x}} F^T(\vec{x}) M^{C,T}(\vec{x}) d\vec{x}}{||F^T|| \, ||M^{C,T}||}.$$

Finally, the reward score of cell C is the average: $\widehat{R^C} = \frac{1}{5} \sum_{T=1}^{5} R^{C,T}$.

Similarly, for a place score to measure the spatial consistency of place maps of a cell, we computed the Pearson correlations between them, i.e., for task epoch T_1 and T_2 ($T_1 \neq T_2$),

$$P^{C,T_1 T_2} = \frac{\int_{\vec{x}} M^{C,T_1}(\vec{x}) M^{C,T_2}(\vec{x}) d\vec{x}}{||M^{C,T_1}|| \, ||M^{C,T_2}||}.$$

Likewise, the place score of cell C is the average: $\widehat{P^C} = \frac{1}{\binom{5}{2}} \sum_{T_1 T_2} P^{C,T_1 T_2}$. Since in different-context experiments, novel place fields for place cells were produced in Task Set2 due to remapping, which could hurt the spatial consistency and place scores, we ruled out the place maps from Set2 when we calculated the Pearson correlation $P^{C,T_1 T_2}$ (but we did not do this in the same-context experiments). That is to say, for different-context experiments, $T_{1,2} \neq 3$.

The variation of reward scores is used to recognize the change in coding properties. For each cell C, if the variation of $R^{C,T}$ for any two consecutive tasks, i.e., $|R^{C,T} - R^{C,T+1}|$ is larger than 0.52, we select it as a "transition cell."

2.4.4 Surrogate place/reward scores

We developed surrogate tests to further categorize place/reward cells. In these tests, we simulated 2000 surrogate place/reward cells (1000 each). For both simulations, we used parameters extracted from the data as follows: First, the number of place fields (PFs) for each cell was randomly generated from the distribution of PF number in the data. Then, for each PF, we extracted its location from the distribution of PFs COM in each task. The distribution of drift distance was extracted from the distances between PFs centers of cells in Random1/2. Likewise, the distributions of size and orientation of PFs were also extracted from the data.

Null Model1: For the reward score, we assumed that all cells were place cells and checked if it could explain the data, i.e., Are the firing maps for a cell in multiple tasks made up by firing around fixed locations plus random drifts of the PFs? The firing maps for each cell in multiple tasks were made up by place fields (PFs) in fixed locations with random drifts (measured in random tasks). The number, size, and orientation of PFs were simulated from the data distributions mentioned above.

Null Model 2: For the place score, we assumed that all cells were reward cells and checked if it could explain the data, i.e., Are the firing maps for a cell in multiple tasks made up by firing around correct feeders? In this case, the number, size, and orientation of PFs were all randomly generated as above, from the data distributions. However, the locations of the PFs were restricted around correct feeders in each task, i.e., randomly generated from the distributions of PFs extracted from data, excluding the possibility of a PF away from any correct feeder.

2.4.5 Measures of coding properties

Following Gauthier and Tank (2018), we studied the extent to which the activities of reward cells were indicative of the expectation of rewards or of its delivery. To address these 2 cases, for the firing patterns of a cell C, we computed their correlations with correct feeder locations/the delivery of reward/the expectation of reward. These correlations were indicated by the proportions of spikes taking place in corresponding time-windows.

In task T, a neuron C fired $n^{C,T}$ times. We assumed $n_{\text{Site}}^{C,T}$ spikes out of the total $n^{C,T}$ spikes were found within 30 pixels (10 cm) from the correct feeders, $n_{\text{Delv}}^{C,T}$ spikes took place within t_d seconds after the delivery of reward, and $n_{\text{Expt}}^{C,T}$ spikes took place within t_e seconds before the rat reached the feeder (see time-windows in Fig. 4c). These time-windows related to expectation and reward delivery were called "expectation-time-windows" and "delivery-time-windows." The time parameters, $t_{d,e}$ (d for delivery and e for expectation), were changeable for the length of the time-windows ranging from 1 to 3 s, based on the average time for a rat spent around a feeder (Fig. 4b). We should note that these three time-windows were not mutually exclusive.

Therefore, in task T, the firing probability (FP) of cell C around the correct feeder locations was defined as $p_{\text{Site}}^{C,T} = n_{\text{Site}}^{C,T}/n^{C,T}$, indicating the correlation of the firing pattern with the locations of correct feeders. FP is also the firing rate of the cell in a window of interest, divided by the overall firing rate of the cell in the given task. Similarly, the FPs in delivery-time-window and expectation-time-window were defined as $p_{\text{Delv}}^{C,T} = n_{\text{Delv}}^{C,T}/n^{C,T}$ and $p_{\text{Expt}}^{C,T} = n_{\text{Expt}}^{C,T}/n^{C,T}$, respectively.

Using the notations above, across different tasks, we computed the averaged FP around the location of correct feeder for cell C as:

$$\overline{p_{\text{Site}}^{C}} = \frac{1}{\#T} \sum_{T} p_{\text{Site}}^{C,T}.$$

Similarly, we computed the average $\overline{p_{\text{Delv}}^{C}}$ for the correlation with reward delivery and $\overline{p_{\text{Expt}}^{C}}$ for the correlation with expectation.

For each cell C, its sensitivity to expectation was computed as the ratio between the number of spikes in the expectation-time-windows and the number of all expectation-time-windows. For example, in an experiment, a rat visited feeders 300 times across all tasks. Therefore, we had 300 expectation-time-windows. If spikes from cell C were found in 240 out of 300 expectation-time-windows, cell C's sensitivity to reward expectation (in Set1 only) was therefore $240/300 = 0.8$. Cell C's sensitivity to reward delivery was computed in the same way, except replacing the expectation-time-windows by delivery-time-windows.

3 Results

Most results were presented in abstract form (Xiao et al. 2019). The results presented below are obtained after the animals reached asymptotic performance. In these conditions, animals learned Sets1/2 in 504 ± 143 s and reached similar performance as in previous studies (Jones et al. 2012, 2015). Place fields were defined on the basis of the trajectory of the rat on the maze, while the animal was moving above a given speed threshold. Spikes were extracted during these trajectory bouts. In general, most awake SWRs occurred when the animal stopped or slowed down, so spiking during SWRs would not contribute to the place field computations. Here, because we were studying the reward dependence of place fields, we sought to ensure that SWRs spikes did not contaminate our place field computations. Figure 1d shows the location of the rat when SWRs were produced during tasks Random1 from one session of rat1. Figure 1e shows these locations for the entire set of sessions used in this study ($N = 12$ sessions, six from Rat1, four from Rat2, one each from Rat3 and Rat4). For different types of tasks, there were 0.338 ± 0.149 SWRs/s in Random1 and Random2, 0.372 ± 0.142 SWRs/s in Set1 and Set2, and 0.321 ± 0.108 SWRs/s in RecallSet1.

From 1217 cells sorted from the raw data (see "Methods" section) of 12 experimental sessions, we selected 584 cells (48.0%) with well-defined place field(s) in different tasks. Here, 286 cells from same-context sessions and 224 cells from different-context sessions had place field(s) in at least
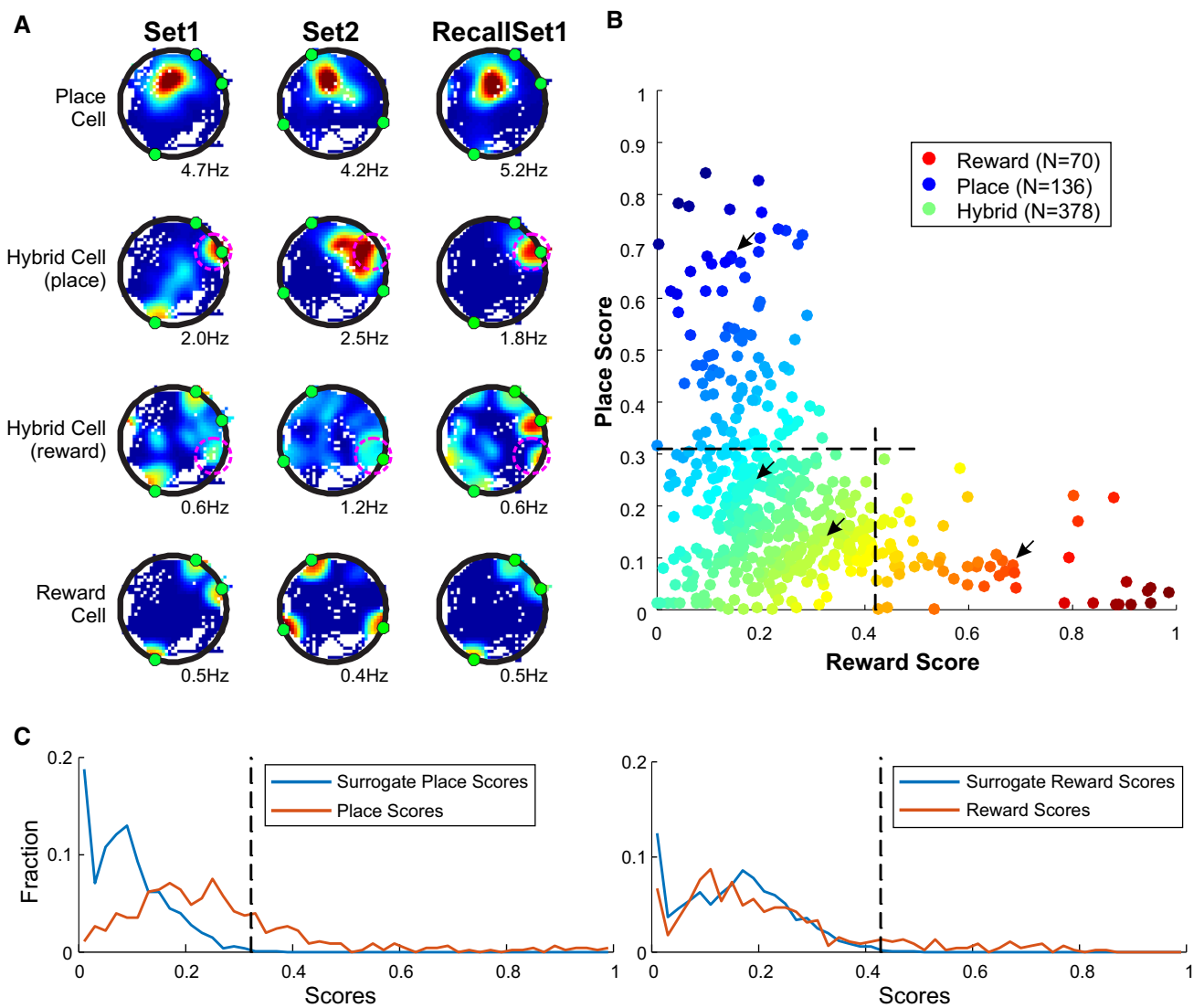
Fig. 2 Cell characterizations. **a** Four representative cells of the 4 categories identified in this study: place cells (top, spatially fixed), reward cells (bottom, follow reward sites), and hybrid cells with place coding dominance (second row) or reward coding dominance (third row). The "place component" of each place field is marked by a dashed pink circle. Rewarded feeders are denoted by a green dot on the periphery of the circular maze. Maximum firing rates are indicated under each graph. **b** Group data showing the skewed nature of the reward/place scores for all cells recorded from all experiments ($N = 584$). Place cells (blue) had small reward scores and high place scores. Reward cells (red) had the opposite properties. Hybrid cells (light blue to light yellow) showed mixed coding property. The separation lines in **b** were chosen above 98% percentiles of surrogate place scores of randomly simulated reward cells (black dash line, horizontal) and surrogate reward scores randomly simulated place cells (black dash line, vertical). The four cells in panel **a** are indicated by arrows. **c** Comparison between the distributions of data and surrogate scores (color figure online)

4 out of 5 tasks, and the remaining 74 cells from control sessions had place fields in both Random1 and Random2. For different-context sessions, we also included cells with place field in Set2 only. The other 633 cells had no place fields in at least two tasks and were excluded from our analyses. We noticed that place fields had four different spatial characteristics (Fig. 2a). Like most classical place cells, some cells had 1 or 2 place fields with fixed locations regardless of the change of reward-delivery location (Fig. 2a, top, green dots, $N = 136/584$). Another group of cells called "reward cells"

had fields tightly bound to the reward location (Fig. 2a, bottom row, $N = 70/584$). The reward coding property of these cells was obtained from 5 different-context sessions, where we obtained 25/224 reward cells. Twelve of 25 cells (48%) fired around all the correct feeders irrespective of the change of context because they fired at all the Set2 feeders, which were also activated in Random1. Notably, we also found a large number of cells with conjunctive properties, referred to as "hybrid cells," for which part of the place fields of the cell exhibited consistency, yet the cell also had a high firing prob-

ability around correct feeders irrespective of their location (Fig. 2a, rows 2 and 3, $N = 378/584$). Some hybrid cells had more highly consistent place fields with lower firing probability around correct feeders (253/378), whereas the other hybrid cells preferentially coded reward locations (125/378). Furthermore, we observed a small fraction (18/584) of the cell population exhibited sharp changes in their coding properties between different tasks (called "transition cells"), i.e., their firing patterns could code some or all of the correct feeders in a task even though they had a fixed place field in the previous tasks, or vice versa. We will discuss the transition phenomena later in this paper.

These three types of cells (place cell, reward cell, and hybrid cell) were classified based on two different indexes. To measure the extent to which a cell was correlated with absolute spatial location or with reward location, we computed its place score $\widehat{P^j}$ and reward score $\widehat{R^j}$ (see "Methods" section). Figure 2b shows that cells fell on a skewed continuum, with place cells following the place score axis (high place score $\widehat{P^j}$, low reward score $\widehat{R^j}$. Figure 2b, blue) and reward cells following the reward score axis (low place score $\widehat{P^j}$, high reward score $\widehat{R^j}$. Figure 2b, red). Hybrid cells had lower mixed scores (Fig. 2b, black, with reward scores of 0.25 ± 0.09, and place scores of 0.18 ± 0.03). Observed from Fig. 2a, an ideal place cell should have identical place fields in different tasks, whereas an ideal reward cell should have place fields at some or all locations of correct feeders. The separation line between place cell and hybrid cell, $y = 0.31$, was selected as the 98.4 percentile of place scores of randomly simulated ideal reward cells. Hence, cells with larger place scores had a high probability of being a place cell. The separation line between the hybrid cell and reward cell, $x = 0.42$ was selected similarly. The classification was not dichotomous since the entire cell population was distributed continuously. When comparing the data to surrogate sets (see "Methods" section), we found that the probability distributions for both place (Fig. 2c, left) and reward (Fig. 2c, right) scores had much longer tails than the surrogate scores in the high-score areas (above the thresholds, dash lines). These results suggested that our data could not be explained by the null models.

Although these two scores described the place and reward components of each cell, from their definition, they were not entirely independent from each other. However, we found no cell that had both high place and high reward scores. A high reward score meant that the cell only had a high firing probability at correct feeders, but since Set1 and Set2 were designed to contain different feeders, this led to low correlations between place maps of Set1 and Set2, as well as Set2 and RecallSet1. On the other hand, a cell with both low place/reward scores was possible as long as its place maps were strongly affected by the collection of feeders, and the place fields were not close to them.

We next compared the physiological properties of the three classes of cells. Among all cells analyzed, many cells showed a single clear place field (136/584 cells, Figure 3a). The majority of place cells had at most 1 or 2 place fields, compatible with previous findings (Park et al. 2011). Hybrid cells, however, had several place subfields since they had stronger responses to the reward locations (black bars in Fig. 3a). For the reward cells, the number of place subfields showed that the selectivity to feeders was different from cell to cell. While some reward cells had as many as 5-8 subfields because they responded to all feeders in the random tasks (21/70), others responded only to a smaller subset of correct feeders (49/70). Overall, the number of place subfields was not significantly correlated with spatial information (see definition from Royer et al. 2010) since the coefficients of determination $r^2 < 0.01$, regardless of the type of the cells (Fig. 3b). We noted that, especially for reward cells (due to the higher percentage of spikes around feeders, Fig. 3c), some spikes around the correct feeders were not successfully recognized as place fields because the place map values failed to exceed the $1 \times$ standard deviation threshold. The distribution of place subfield area differed between categories (Fig. 3d), with the reward cells exhibiting smaller place fields than other cells since they only fired around feeders. (Reward cells: $586 \pm 497 \, cm^2$. Place cells: $1632 \pm 1094 \, cm^2$, Hybrid cells: $1309 \pm 828 \, cm^2$.) This indicated that even though reward cells remapped when reward contingencies changed, their spatial selectivity still allowed for a more precise encoding than place cells. Many reward cells also had averaged firing rates in tasks as low as 0.1 Hz (0.12 ± 0.11 Hz, Fig. 3e), i.e., they only fired about 120 times in a 20-min task. Hybrid and place cells had larger firing rates which resulted in wider distributions (0.70 ± 0.67 Hz).

Despite their differences between place fields and firing rates, reward cells were found mixed with other cells in the recoding from the tetrodes (Fig. 3f), instead of recorded together from a small number of tetrodes. The 1217 analyzed cells came from 65 tetrodes, from which we collected 70 reward cells. In order to assess whether reward cells might be spatially clustered, we computed the proportion of each cell type per tetrode. Sixty of 65 tetrodes contained no more than 2 reward cells, and only 5/65 tetrodes had 3 or 4 reward cells. When compared with other kinds of excitatory pyramidal cells, the distribution of reward cells for each tetrode yielded a weak positive correlation with the number of all cells ($\beta_1 = 0.03, r^2 = 0.04$). These results suggest that although reward cells were putative pyramidal cells, the number of reward cells was not correlated with the number of pyramidal cells recorded. Reward cells therefore may be a subset of place cells that receive strong or more specific reward-related inputs.

We further studied the coding properties of reward cells. For all Set1 and Set2 tasks, we checked if the coding of reward
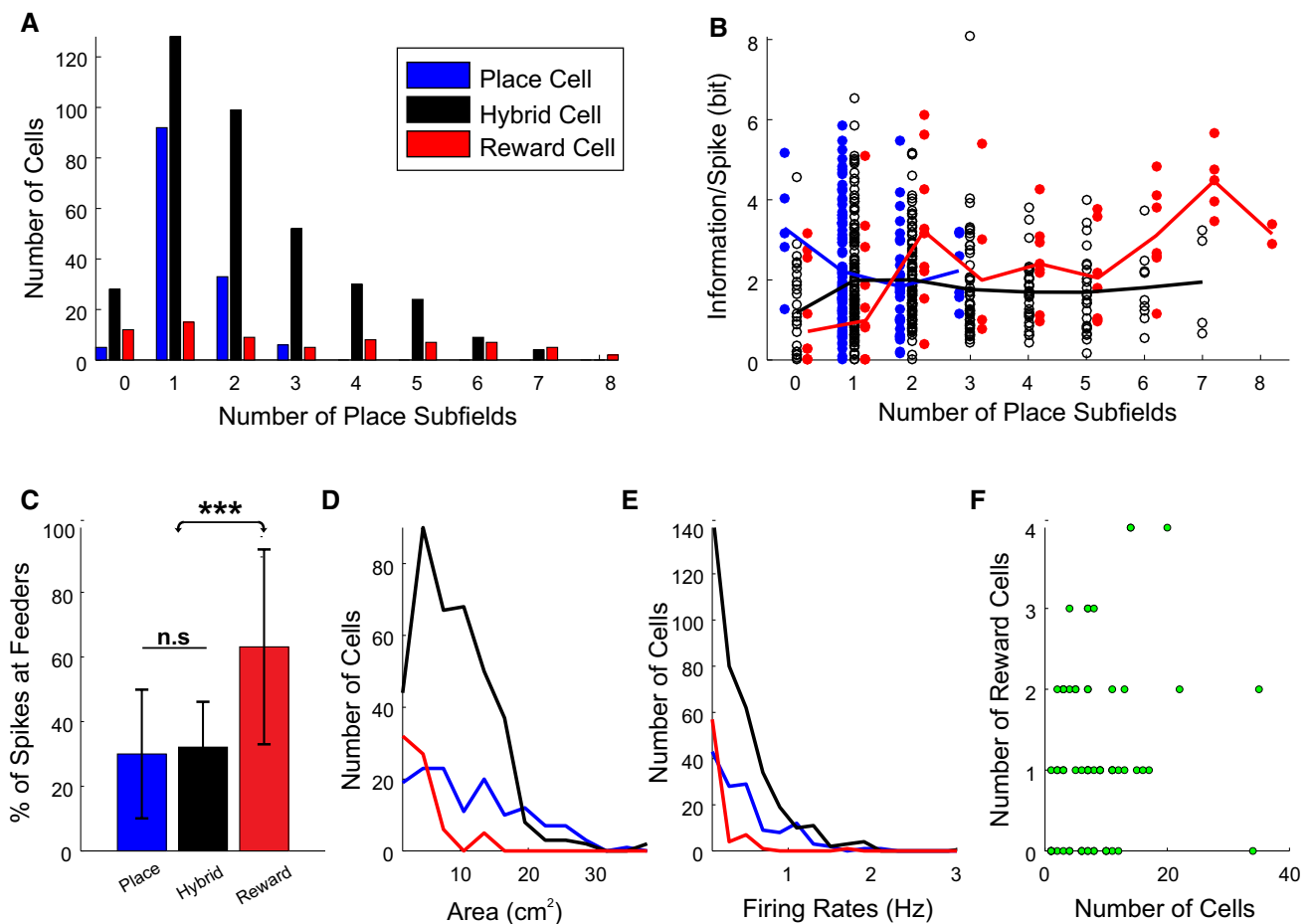
Fig. 3 Place-field properties for each cell type. **a** Number of place subfields across different tasks. Most place cells had no more than 2 subfields (blue), whereas some reward cells could have 8 subfields (red, one for each feeder in random tasks). **b** Information per spike for each cell. Average values indicated by continuous line. **c** Fraction of spikes at correct feeders for place, hybrid, and reward cells. **d** Reward cells had smaller place fields than other cell types. **e** The averaged firing rates of reward cells were lower than those of the other classes. **f** Tetrode-specific numbers of reward cells as a function of the number of all cells recorded on that same tetrode (color figure online)

cells had any preference for any of the three correct feeders. For each Set1 or Set2 task featuring reward cells, the number of spikes produced at each of the three feeders was counted, as well as the number of visits to each of the three feeders. After each task, to assess the firing probability of each cell and the visit probability to each feeder, the counts were normalized by the total number of spikes produced by the cell and the total number of visits during the task, respectively. Figure 4a shows firing probabilities plotted against visits probabilities across all same- and different-context sessions ($N = 58$ cells, $N = 10$ sessions). Even though the size of our sample was relatively small, we saw that for most reward cells the firing probabilities at a feeder were close to the visit probability to that same feeder (Fig. 4a, dots in the green circle). On the other hand, there were a small number of reward cells that strongly preferred specific feeders, and were unlikely to fire at other feeders (Fig. 4a, both in blue circles).

We next proceeded to study the extent to which reward firing was due to the actual delivery of rewards or to the expectation of reward. Because of the probabilistic nature of the reward delivery (see "Methods" section), rats could visit a correct feeder with or without actual reward delivery. For all rats, the averaged time spent around a correct feeder without reward and a correct feeder with reward was different (Fig. 4b): Due to reward consumption time, rats stayed ~ 3 s (Fig. 4b, solid line) around a feeder that delivered a reward, while they left a feeder faster (~ 1 s, Fig. 4b, dashed line) if the feeder did not deliver the reward. These two distributions of staying time were statistically different ($p < 0.001$, standard t test).

Rat trajectories were analyzed around the reward sites, whether rewards were delivered or not during the memory-driven portions of the Set1 and Set2 learning tasks, and during recall. Figure 4c shows an example of a time-window spent
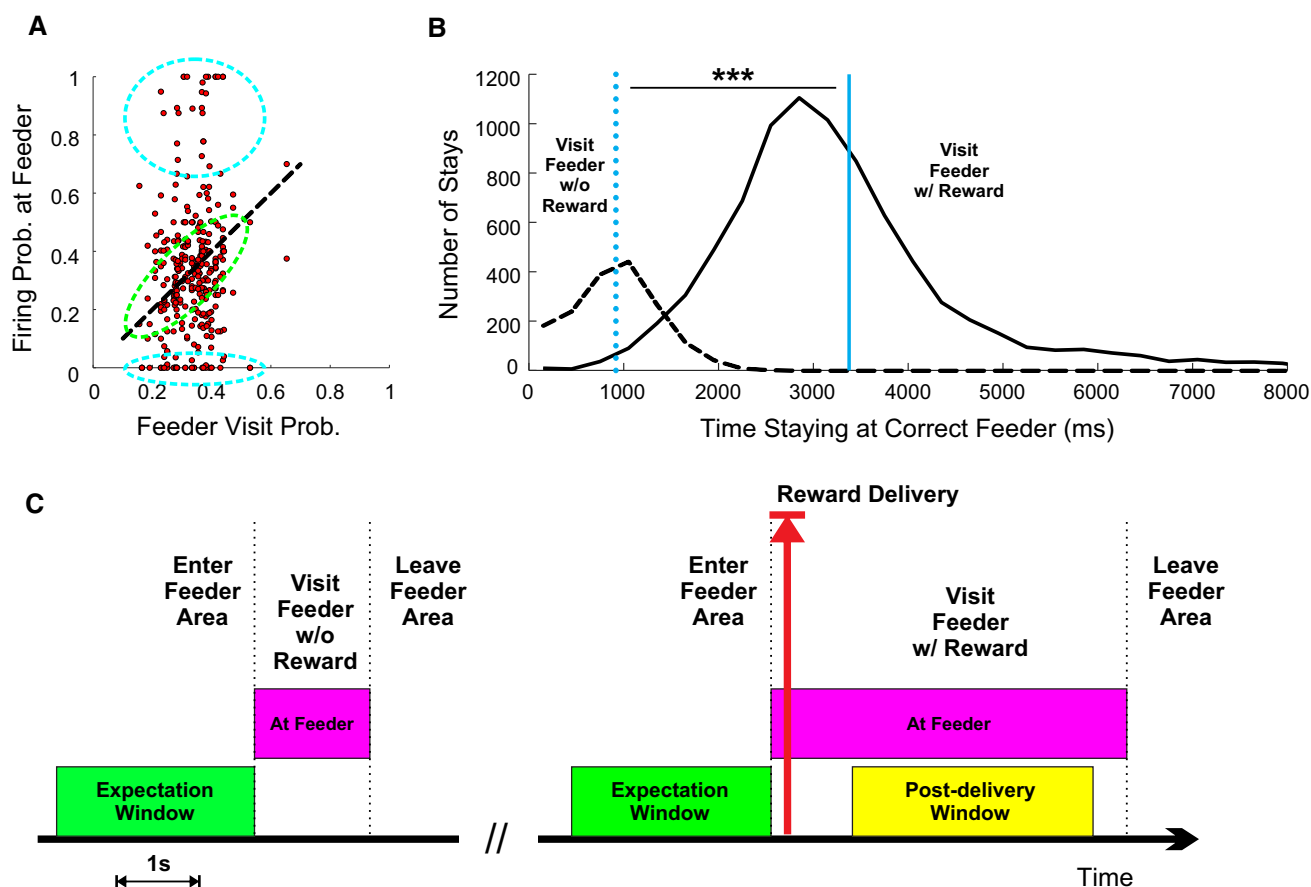
Fig. 4 Properties of the cell population at correct feeders. **a** Firing probability of a cell at a correct feeder versus the visiting probability to that same feeder. While firing probabilities of most cells were proportional to visiting probabilities (around the 45 degree line, marked by the green circle), some cells exhibited strong preference for specific feeders and hence were not likely to encode the visit to other feeders (blue circles). **b** Distributions of time rats spent at a correct feeders. Average time is indicated by vertical lines (dashed and solid). **c** Illustration of the time-windows used for analyses: at-feeder intervals (purple), expectation-time intervals (green), and consumption time interval (yellow). The time of reward delivery is indicated by a red arrow (color figure online)

at correct feeders that were not rewarded, and a time-window spent around a rewarded feeder. The red arrow indicates the time when the sugar water reward was released. For each visit to a correct feeder, a time-window was chosen from the time of the feeder triggering, lasting t_d second (d for delivery and e for expectation), to assess the response of the cells to reward delivery (Fig. 4c). On the other hand, for the response to reward expectation, we set another spike counting time-window lasting t_e second before the rat entered the feeder area (10 cm from the feeder). Entrance to this area triggered the reward solenoid which produced an audible click immediately followed by reward delivery. Therefore, in the reward expectation window animals did not know whether they would receive a reward or not. Based on the distributions of staying time around feeders, the length of time-windows was set as $t_{d,e} = 1 \sim 3$ s (Fig. 4c) to count the spikes responding to delivery/expectation of the reward and then obtained the firing probabilities (FP) $\overline{p_{\text{Site}}^C}$, $\overline{p_{\text{Delv}}^C}$, $\overline{p_{\text{Expt}}^C}$, i.e.,

cell C's FPs around feeder locations (in at-feeder windows), in delivery-time-windows, and in expectation-time-windows (Fig. 4c. See "Methods" section). Different choices of time-window length were made as $t_{d,e} = 1$ or 3 s to test if the results were sensitive to the length of time-windows. Specifically, larger time-windows could result in a bigger overlap between delivery/expectation-time-windows, since rats usually took 5 s to cross the maze, going from one feeder to its diametrically opposite one. In addition, FP in delivery windows $\overline{p_{\text{Delv}}^C}$ did not necessarily represent the correlation to "acquisition" of the reward, but rather a reasonable compromise, since the rat consumed reward after the delivery.

An decoding approach, using instantaneous firing rate in moving 3-s windows, yielded inconclusive results as to whether a reward cell. When compared to a place cell, could be indicative of reward expectation or reward delivery (reward delivery: 53% accuracy, reward expectation: 29.5% accuracy, details not shown). This observation suggested that
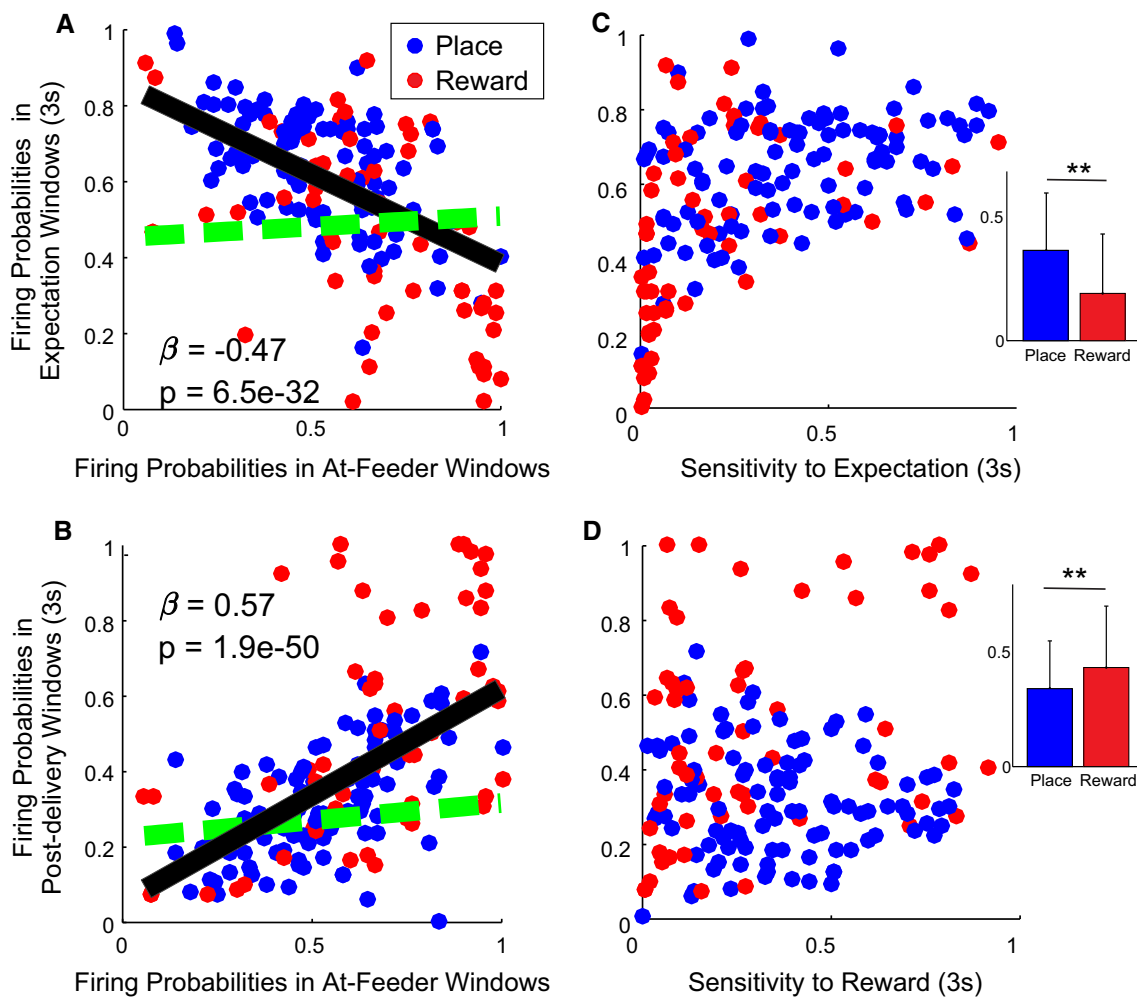
Fig. 5 Firing probabilities in different time-windows reveal the firing preference of reward cells. **a** Firing probabilities for place and reward cells in 3-s expectation windows. Black line indicates the linear regression across all cells (hybrid cells not plotted). Linear coefficient and p-value are indicated in inset. Green dashed line: linear regression for random 3-s windows (data points not plotted for clarity). **b** Firing prob-

abilities in post-delivery 3-s windows. Labels and regressions as in **a**, **c**, **d**. Specificity (represented by the firing probabilities in corresponding windows) versus sensitivity to expectation windows (**c**) and delivery windows (**d**) across all place and reward cells. Insets: mean sensitivity (color figure online)

instantaneous firing rate might not be sufficient to assess reward delivery or reward expectation. We therefore analyzed these epochs more coarsely, using firing probability within each task. The correlations between FPs ($\overline{p^C_{Site}}$, $\overline{p^C_{Delv}}$, $\overline{p^C_{Expt}}$) revealed a strong preference to the first hypothesis, i.e., the recorded cells, especially the reward cells, preferably responded to the acquisition of rewards (Fig. 5, hybrid cells were omitted for clarity). Generally, a higher FP around correct feeders resulted in a higher probability for cell C to be a reward cell. Most of the reward cells had a higher FP than other cells within the 1-s time-window after delivery of reward (not shown). Furthermore, FPs around correct feeders and FPs in 3-s delivery-time-window formed a strongly positive correlation (Fig. 5b black line, $\beta = 0.57$, $\overline{p^C_{Delv}}$ ver-

sus $\overline{p^C_{Site}}$), showing that the spikes produced after delivery of rewards made up most of the spikes in at-feeder windows, although the latter did not strictly contain the former. Accordingly, hybrid cells also had higher FP around correct feeders than place cells, but lower than the reward cells (not shown). In contrast, FPs around the correct feeders and FPs in expectation-time-windows showed a negative correlation (Fig. 5a black line, $\overline{p^C_{Expt}}$ vs $\overline{p^C_{Site}}$) where most reward cells had a lower response to expectation than other cells. To assess the significance of these regressions, we also computed FPs in random time-windows (regression lines shown as green dash lines in Fig. 5a, b). For example, in Fig. 5a, for each cell, we calculated the FPs in as many randomly-chosen 3-s time-windows, as the number of expectation windows used for the

analyses and found that FPs in such windows were not significantly correlated with the FPs around feeders (the green dash line, Fig. 5a). Similar results were obtained for the post-delivery windows (Fig. 5b). We also noted that most reward cells had very low sensitivity to the reward expectation because they only fired in a small fraction of expectation-time-windows prior to entering the feeder areas (Fig. 5c, d, see definitions in "Methods" section). A small fraction of reward cells (10 out of 70) exhibited both high sensitivity and specificity to expectation (high FPs in expectation-time-windows) when taking $t_e = 1$ (not shown). Some hybrid cells also exhibited a mixed coding property of delivery and expectation (not shown), though the actual mechanism of reward coding for pyramidal cells could be more complicated than a simple mixture. Finally, we found that 70% of the reward cells (43/61 from same/different-context sessions) had a larger number of spikes in post-delivery windows than in expectation windows (paired t test, $p < 0.001$) and that 100% of the reward cells had a larger number of spikes in post-delivery windows than in randomly chosen windows (not shown).

We next asked whether the place and reward coding properties of each cell were fixed or could change dynamically. Figure 6a shows an example of a cell that initially had a place cell coding property in Random1 and Set1, but then switched to firing near some of the correct feeders in Set2, RecallSet1, and Random2. Note that the cell did not fire at the same correct feeder in these three tasks. This cell showed a transition from place to reward coding (all tasks in the same context). Note that the firing rates during sleep epochs (only included time during in-REM sleep) did not change significantly before and after the transition (Fig. 6b), which ruled out the possibility we had unstable recordings or errors in spike cutting.

The cells exhibiting transition in coding properties, or "transition cells," were detected by the variation of reward score (see "Methods" section). Encoding a fixed location resulted in a low reward score, whereas encoding reward resulted in a high reward score. Hence, a transition should bring a significant variation to the reward score of such a cell. By looking at the largest variation of reward scores for consecutive two tasks, we found that 18/584 cells exhibited transitions of coding property between tasks, whose reward scores underwent sharp change (variation of reward score > 0.52, dashed green box in Fig. 6c). These transition cells were classified as hybrid cells at an earlier stage due to their medium averaged reward and place scores, yet the sharp change in reward scores rather suggested a change of coding properties.

We studied the coding properties of these 18 transition cells in different tasks and found that 11 cells transferred from place cell to reward cell, whereas the other 7 cells transferred from reward to place coding. Figure 6d shows the dynamic

of the reward scores for the 21 cells mentioned above. Solid curves show 11 cells for place-to-reward transitions (increase in reward scores), and dashed curves show 7 cells for reward-to-place transitions. The red curve shows the cell in Fig. 6a, and the black dots indicate when the transition took place. Since we also found 222/584 stable place cells and 67/584 stable reward cells (not shown), cell exhibiting transitions only made up a small fraction of the population recorded.

To exclude the possibility that the transition phenomena were the results of a place cell spuriously classified as a reward cell in some tasks (or vice versa), we built a null model in which we assumed that there was no transition taking place during the sessions, and therefore, that the change of the place maps of a place cell was due to random drift, whereas for a reward cell it was due to the change of reward locations. We built the model from the parameters obtained from the data as above, except that the drift of place fields was extracted from task to task (not only from Random1/2 as it was above). In addition, we excluded the possibility that a surrogate cell had no place field in some task, since the transition cells we observed had indeed place fields in every task. We simulated 1000 surrogate place/reward cells separately. We produced surrogate hybrid cells by linearly combining the place maps of the 1000 samples of surrogate place and reward cells with weights from 0.1:0.9, 0.2:0.8, …, to 0.9:0.1, respectively (0:1 and 1:0 are place/reward cells themselves), which yielded 9000 samples of hybrid cells.

We found that no place cells out of 1000 and 2 reward cells out of 1000 exhibited a change of reward scores larger than 0.52 in two consecutive tasks. The reward score changes in these 2 cells were due to variation change in the number of place fields around rewarded feeders. For hybrid cells, only one out of 9000 exhibited the same threshold crossing. In all, with a null model in which we assumed no transition between reward and place cells, the probability to find a sharp change of reward score larger than 0.52 was < 0.2% across place/hybrid/reward cells, which was significantly lower than the finding in our experiments (3% = 18/584), suggesting that this result was not due to chance.

4 Discussion

One of the basic tenets of motivated spatial navigation in robots, animals, and humans is the notion that choosing to navigate to a specific location is related to the immediate or delayed expectation of a reward at that location. Many studies have focused on understanding the neural mechanisms of spatial navigation (Hinman et al. 2018) or the neural basis of reward processing (Luo et al. 2011) separately. Because in laboratory rats, reward and spatial navigation are intimately linked, it is reasonable to think that their neural systems tightly interact to the extent that reward information
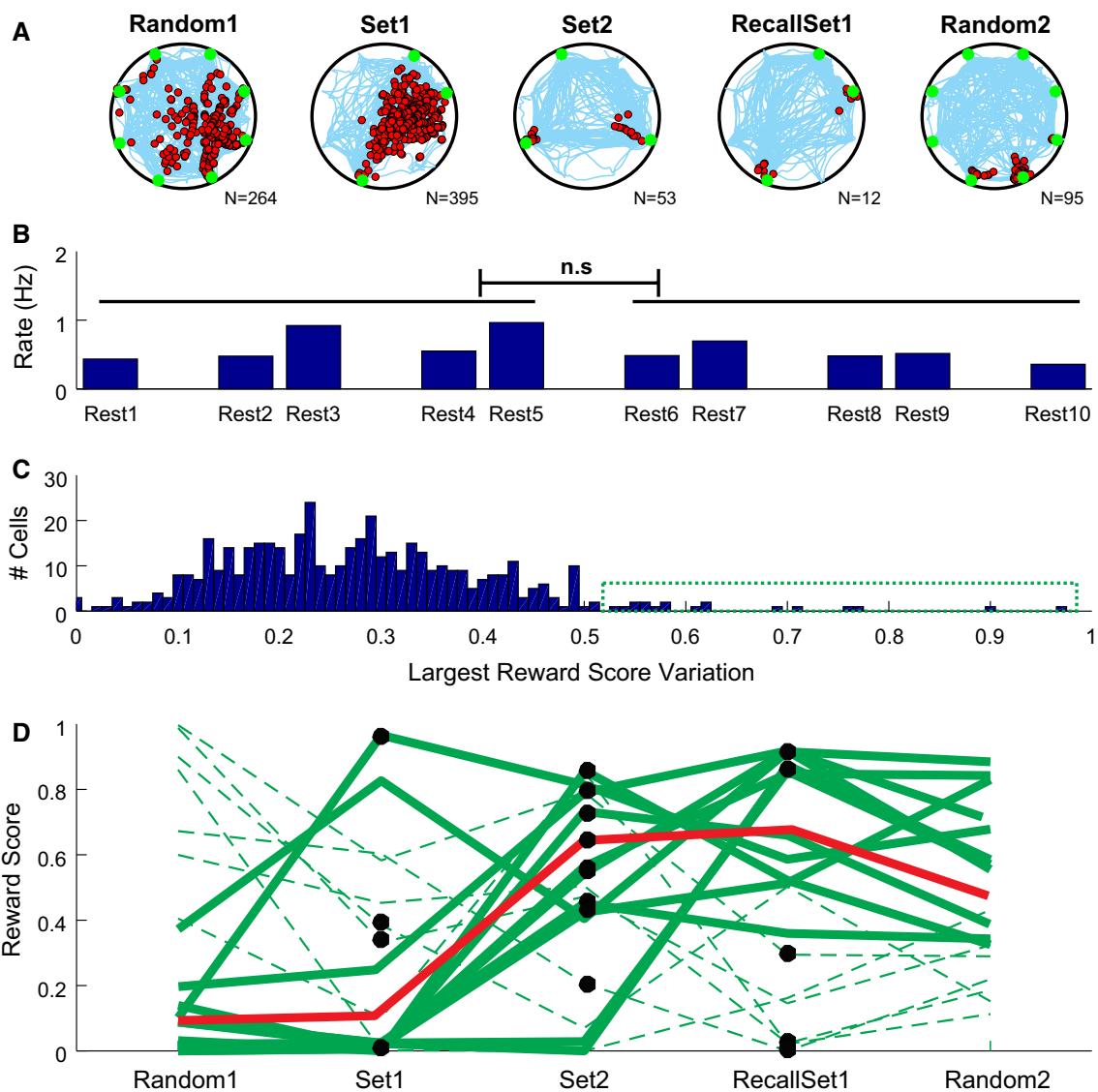
Fig. 6 Transitions between place cell and reward cell. **a** Example of transition from place cell (Random1, Set1) to reward cell (Set2, RecallSet1, Random2), with unequally represented reward sites. The number of spikes contributing to rate map computations is indicated below each graph. **b** Firing rates of the cell in **a** in different sleep epochs. No sharp changes of sleep firing rates indicate that the decrease in spike numbers in the last 3 tasks was not the result of recording stability loss or spike-sorting errors. **c** Distribution of largest variations of reward scores between consecutive tasks across all cells. A small fraction of cells exhibited sharp changes (variation of reward score > 0.52, green dashed rectangle), which were labeled as transition cells ($N = 18$). **d** Reward scores for the 18 transitions observed in **c**. Eleven cells transitioned from place to reward coding (solid) and 7 transitioned from reward to place coding (dashed). Black dots indicate the task after which the transition took place. The red line shows the cell in **a** (color figure online)

may be integrated, at least in part, in the spatial navigation code. Motivated by this hypothesis, we studied the coding properties of CA1 hippocampal cells at or near carefully manipulated reward sites. Compatible with others (Dupret et al. 2010; Poucet and Hok 2017; Singer and Frank 2009), we demonstrated the existence of reward-location coding cells in dorsal CA1, complementing previous studies using virtual reality on a 1D track (Gauthier and Tank 2018).

4.1 Anatomical and physiological considerations

Many studies have shown that hippocampal CA1 cells may encode more than absolute spatial information (Rueckemann and Buffalo 2017; Eichenbaum 2018; Zhang and Manahan-Vaughan 2015; Liu et al. 2018). Some cells fire relative to the locations of neutral objects placed on the maze (Deshmukh and Knierim 2013), while others are strongly associated with the location of rewards (Gauthier and Tank 2018; Poucet and

Hok 2017; Rolls and Xiang 2005). CA1 firing may reflect the actual value of a location in complex probabilistic decision-making tasks (Lee et al. 2012) but not in simpler tasks (Duvelle et al. 2019). While CA1 is not generally known to receive direct inputs from the reward systems, CA3, a synapse away, does (Berridge and Kringelbach 2015; Luo et al. 2011). It has been suggested that the influence of the reward system on the hippocampus and associated circuitry could at least in part be attributed to inputs from the ventral striatum or ventral tegmental area during sleep-induced reactivation (Lansink et al. 2009; Pennartz et al. 2004; Valdes et al. 2015) or during behavior (Mamad et al. 2017). Notably, reactivation can also occur in the awake state during awake SWRs and often occurs near reward sites, as our results and that of others showed (Diba and Buzsaki 2007; Jadhav et al. 2012; Malvache et al. 2016; Olafsdottir et al. 2018). As such, awake reactivation is likely to contaminate the strictly spatial component of hippocampal firing when place fields are computed and artificially introduce a reward dependency in the spatial code. Action potentials occurring during SWRs should therefore be excluded when place fields are analyzed near reward locations.

Compatible with others, while our results suggest that there may be a subpopulation of CA1 cells able to encode specifically for reward zones, our experiments did not allow for a determination of whether these cells were found anatomically in different subregions of dorsal distal CA1. Others have suggested that there may be a CA1-sublayer specificity in the manner in which place fields aggregate around reward sites, with deep layers (close to *stratum oriens*) being more sensitive than superficial (close to *stratum radiatum*) layers (Danielson et al. 2016). More work needs to be done to ascertain these findings.

4.2 Complex spatial navigation and the need for reward coding

There has been conflicting evidence as to whether place fields aggregate around or overrepresent reward areas. Notwithstanding the issue of SWR contamination, it is also possible that the differences seen can be attributable to the nature of the task. It is possible that when tasks are simple, with no cognitive demands or significant decision-making components, such as running on a track for rewards (Diba and Buzsaki 2007; Pennartz et al. 2004) or other simple tasks (Duvelle et al. 2019; Speakman and O'Keefe 1990) place fields do not bear any specific relationships to reward location and uniformly cover the environment. If the task is slightly more complex, involving working memory and simple decision making such as in a continuous alternating T maze, some place fields may shift coherently toward the reward location (Lee et al. 2006), and if the same task is made more challenging, some place field remap to reward

zones (Mamad et al. 2017). These results suggest that if the task is sufficiently complex and requires a combination of working memory or short-term memory, probabilistic decision making among multiple locations, or multiple possible routes to a goal, such as in our study (memorization of 2 sets of 3 feeders with probabilistic reward delivery, separated by several hours), the reward–place system dynamically adjusts through remapping, presumably to improve performance (Dupret et al. 2010; Gauthier and Tank 2018; Tryon et al. 2017; Mamad et al. 2017; Rolls and Xiang 2005). The finding of goal-related place field accumulation may also be related to the stressful nature of the task, though that relationship has not been explicitly tested (Hollup et al. 2001).

4.3 A dynamic place-reward conjunctive code

Our results show that a small population of cells dynamically remap from spatial coding to reward coding or vice versa. Different forms of remapping have been observed in response to changes in sensory cues or changes in the nature of the task (Latuske et al. 2017; Ainge et al. 2012; Markus et al. 1995). Interestingly, studies using a block design across multiple sessions showed that a small portion of place cells may remap toward a goal area (Kobayashi et al. 2003). This study did not vary goal location (Set1/2 as in our study), so it is not possible to ascertain whether the cells became reward coders or whether they underwent classic remapping because of learning-mediated plasticity. Also, rewards in this study were through medial forebrain bundle self-stimulation, which may be significantly different from natural rewards. This dynamic shift between place and reward coding indicates that the field aggregation around reward/goal areas, as reviewed above, is a dynamic rather than static feature of the hippocampal circuit in complex tasks.

4.4 Neuromodulation

Most computational theories and robotic implementations of motivated spatial navigation rely on reinforcement learning (Chersi and Burgess 2015; Llofriu et al. 2015; Cazin et al. 2019; Scholkopf and Mallot 1995; Strosslin et al. 2005). The conundrum is that reinforcement learning approaches rely on the usually anatomically and temporally diffuse actions of neuromodulators such as dopamine. This diffuseness and low temporal resolution seem at odds with the selectivity and precision of reward-directed spatial navigation. Recent work, however, has shown that VTA neurons could selectively and precisely reactivate during sleep, indicating that specific VTA dopaminergic cells could in principle carry reward information to specific hippocampal or cortical memories (Valdes et al. 2015). The finding of reward cells within the hippocampus furthers this idea and suggests that, with training, the information about reward might become inte-

grated into the hippocampal neural code and continues to work in concert with the VTA. This phenomenon may be akin to the dynamics between the hippocampus and cortex during memory consolidation in general (Nadel et al. 2000; Hardt and Nadel 2018; Sekeres et al. 2018). These reward cells may be the positive counterpart to the fear "engram cells" found in the hippocampus (Bittner et al. 2015).

4.5 Comparison with recent work

As others have demonstrated, the presence of reward-sensitive cells was not due to artifactual correlations with behavioral or perceptual cues bound to the rewards (Gauthier and Tank 2018). We furthered this finding and confirmed that the spikes of reward cells were indeed induced by rewards, rather than LED blinking or SWR-related bursts. Although LED blinking played an important role in random tasks and in the learning phases of Set1/2 to provide cues, it was delayed for 15 s in RecallSet1 and the test phases of Set1/2. In these cue-less conditions, reward cells still responded to the delivery of reward within 3 s of reaching a correct feeder. Also, because we had excluded the population bursts during SWRs before computing place maps, the reward dependence of the cells was guaranteed to be independent of SWR-related activity.

Using the rate maps of each cell, we used a place score and reward score for quantifying the similarity and correlations of the place fields across several tasks, leading to the division of the cell population into place, reward, and hybrid categories. Hybrid cells could have a dominance for place coding or reward coding. Although largely compatible, some of our findings of the reward coding properties of CA1 neurons stood in contrast to those from Gauthier and Tank (2018).

First, it was interesting to note that the hybrid-coding phenomenon was not reported by Gauthier and Tank (2018). They instead used a dichotomy of place coding and reward coding properties to explain their results. While we are unsure of the source of this discrepancy, this may have been reasonable given that their place fields were one-dimensional on a linear track, and perhaps the shift and the mixing of place fields were more likely to have been overlooked. Our observations furthermore confirmed previous results that hippocampal place fields could shift toward the reward location (Dupret et al. 2010) and that place cells might exhibit excess firing around the reward locations (Poucet and Hok 2017). Whether the finding of such hybrid cells comes from our more distal CA1 recording site remains to be tested. Second, in our experiments, although a small fraction of reward cells responded to the expectation of reward, most cells responded after reward delivery. We obtained a higher proportion of reward cells and hybrid reward cells (195 out of 1217 recorded cells, 16.0% vs 4.4% resp.) than found in this previous study. This difference may be due to the fact that a reward cell could fire only one or two spikes after the delivery of reward, making it difficult for its calcium signal to be detected (Deneux et al. 2016). Third, we observed a few transitions between place cells and reward cells that put into question the notion of "sharp boundary" asserted by Gauthier and Tank (2018). We could not observe any preference of the transition's direction from only 18 cells, despite that the transition itself was a reliable phenomenon based on the loss or gain of firing fields. Place cells and reward cells carried different spatial information in the pyramidal cell population. Hence, the transitions could represent the potential for pyramidal cells to change their coding roles in spatial navigation.

The changes of spatial context or task goals included in our experiments could not explain the transitions since the coding properties did not transition back by the recovery of spatial context or task contingencies (Random2 vs Random1, RecallSet1 vs Set1). This leads to the possibility that the transition could be a result of the synaptic plasticity of hippocampal and hippocampal-related circuits (Neves et al. 2008). The finding of a seemingly continuous population of cell in the reward/place coding dimensions raises the possibility that hippocampal pyramidal cells in dorsal CA1 carry both spatial and reward information in an integrative and plastic manner to various extents. Place or hybrid place cells are simply dominated by spatial information, while reward or hybrid reward cells are dominated by reward information. A transition is, therefore, a simple plastic change in input weights yielding a change in coding properties.

In sum, our study suggests that a large fraction of hippocampal pyramidal cells may receive a dynamic mixture of spatial and reward information that manifests itself in the spatial firing properties of the neurons across multiple tasks. Such a mixture can be used to encode and direct motivated spatial behavior appropriately. Further computational and robotic work is needed to elucidate the functional advantages of this type of continuous and dynamic conjunctive coding.

References

Ainge JA, Tamosiunaite M, Worgotter F, Dudchenko PA (2012) Hippocampal place cells encode intended destination, and not a discriminative stimulus, in a conditional T-maze task. Hippocampus 22(3):534–543. https://doi.org/10.1002/hipo.20919

Barry C, Lever C, Hayman R, Hartley T, Burton S, O'Keefe J, Jeffery K, Burgess N (2006) The boundary vector cell model of place cell firing and spatial memory. Rev Neurosci 17(1–2):71–97

Berridge KC, Kringelbach ML (2015) Pleasure systems in the brain. Neuron 86(3):646–664. https://doi.org/10.1016/j.neuron.2015.02.018

Bittner KC, Grienberger C, Vaidya SP, Milstein AD, Macklin JJ, Suh J, Tonegawa S, Magee JC (2015) Conjunctive input processing drives feature selectivity in hippocampal CA1 neurons. Nat Neurosci 18(8):1133–1142. https://doi.org/10.1038/nn.4062

Boccara CN, Sargolini F, Thoresen VH, Solstad T, Witter MP, Moser EI, Moser MB (2010) Grid cells in pre- and parasubiculum. Nat Neurosci 13(8):987–994. https://doi.org/10.1038/nn.2602

Bush D, Barry C, Manson D, Burgess N (2015) Using grid cells for navigation. Neuron 87(3):507–520. https://doi.org/10.1016/j.neuron.2015.07.006

Cazin N, Llofriu Alonso M, Scleidorovich Chiodi P, Pelc T, Harland B, Weitzenfeld A, Fellous JM, Dominey PF (2019) Reservoir computing model of prefrontal cortex creates novel combinations of previous navigation sequences from hippocampal place-cell replay with spatial reward propagation. PLoS Comput Biol 15(7):e1006624. https://doi.org/10.1371/journal.pcbi.1006624

Chersi F, Burgess N (2015) The cognitive architecture of spatial navigation: hippocampal and striatal contributions. Neuron 88(1):64–77. https://doi.org/10.1016/j.neuron.2015.09.021

Colgin LL, Moser EI, Moser MB (2008) Understanding memory through hippocampal remapping. Trends Neurosci 31(9):469–477. https://doi.org/10.1016/j.tins.2008.06.008

Danielson NB, Zaremba JD, Kaifosh P, Bowler J, Ladow M, Losonczy A (2016) Sublayer-specific coding dynamics during spatial navigation and learning in hippocampal area CA1. Neuron 91(3):652–665. https://doi.org/10.1016/j.neuron.2016.06.020

Deneux T, Kaszas A, Szalay G, Katona G, Lakner T, Grinvald A, Rozsa B, Vanzetta I (2016) Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. Nat Commun 7:12190. https://doi.org/10.1038/ncomms12190

Deshmukh SS, Knierim JJ (2013) Influence of local objects on hippocampal representations: landmark vectors and memory. Hippocampus 23(4):253–267. https://doi.org/10.1002/hipo.22101

Diba K, Buzsaki G (2007) Forward and reverse hippocampal place-cell sequences during ripples. Nat Neurosci 10(10):1241–1242. https://doi.org/10.1038/nn1961

Dupret D, O'Neill J, Pleydell-Bouverie B, Csicsvari J (2010) The reorganization and reactivation of hippocampal maps predict spatial memory performance. Nat Neurosci 13(8):995–1002. https://doi.org/10.1038/nn.2599

Duvelle E, Grieves RM, Hok V, Poucet B, Arleo A, Jeffery KJ, Save E (2019) Insensitivity of place cells to the value of spatial goals in a two-choice flexible navigation task. J Neurosci 39(13):2522–2541. https://doi.org/10.1523/JNEUROSCI.1578-18.2018

Eichenbaum H (2018) What versus where: non-spatial aspects of memory representation by the hippocampus. Curr Top Behav Neurosci 37:101–117. https://doi.org/10.1007/7854_2016_450

Gauthier JL, Tank DW (2018) A dedicated population for reward coding in the hippocampus. Neuron 99(1):179–193. https://doi.org/10.1016/j.neuron.2018.06.008

Guazzelli A, Arbib MA (1997) NeWG: in search of the rat's world graph. J Braz Comput Soc. https://doi.org/10.1590/s0104-65001997000200001

Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. Nature 436(7052):801–806. https://doi.org/10.1038/nature03721

Hardt O, Nadel L (2018) Systems consolidation revisited, but not revised: the promise and limits of optogenetics in the study of memory. Neurosci Lett 680:54–59. https://doi.org/10.1016/j.neulet.2017.11.062

Hinman JR, Dannenberg H, Alexander AS, Hasselmo ME (2018) Neural mechanisms of navigation involving interactions of cortical and subcortical structures. J Neurophysiol 119(6):2007–2029. https://doi.org/10.1152/jn.00498.2017

Hollup SA, Molden S, Donnett JG, Moser MB, Moser EI (2001) Accumulation of hippocampal place fields at the goal location in an annular watermaze task. J Neurosci 21(5):1635–1644

Holscher C, Schnee A, Dahmen H, Setia L, Mallot HA (2005) Rats are able to navigate in virtual environments. J Exp Biol 208(Pt 3):561–569. https://doi.org/10.1242/jeb.01371

Hoydal OA, Skytoen ER, Andersson SO, Moser MB, Moser EI (2019) Object-vector coding in the medial entorhinal cortex. Nature 568(7752):400

Jadhav SP, Kemere C, German PW, Frank LM (2012) Awake hippocampal sharp-wave ripples support spatial memory. Science 336(6087):1454–1458. https://doi.org/10.1126/science.1217230

Jones B, Bukoski E, Nadel L, Fellous JM (2012) Remaking memories: reconsolidation updates positively motivated spatial memory in rats. Learn Mem 19(3):91–98. https://doi.org/10.1101/lm.023408.111

Jones BJ, Pest SM, Vargas IM, Glisky EL, Fellous JM (2015) Contextual reminders fail to trigger memory reconsolidation in aged rats and aged humans. Neurobiol Learn Mem 120:7–15. https://doi.org/10.1016/j.nlm.2015.02.003

Kobayashi T, Tran AH, Nishijo H, Ono T, Matsumoto G (2003) Contribution of hippocampal place cell activity to learning and formation of goal-directed navigation in rats. Neuroscience 117(4):1025–1035. https://doi.org/10.1016/s0306-4522(02)00700-5

Lansink CS, Goltstein PM, Lankelma JV, McNaughton BL, Pennartz CM (2009) Hippocampus leads ventral striatum in replay of place-reward information. PLoS Biol 7(8):e1000173. https://doi.org/10.1371/journal.pbio.1000173

Latuske P, Kornienko O, Kohler L, Allen K (2017) Hippocampal remapping and its entorhinal origin. Front Behav Neurosci 11:253. https://doi.org/10.3389/fnbeh.2017.00253

Lee I, Griffin AL, Zilli EA, Eichenbaum H, Hasselmo ME (2006) Gradual translocation of spatial correlates of neuronal firing in the hippocampus toward prospective reward locations. Neuron 51(5):639–650. https://doi.org/10.1016/j.neuron.2006.06.033

Lee H, Ghim JW, Kim H, Lee D, Jung M (2012) Hippocampal neural correlates for values of experienced events. J Neurosci 32(43):15053–15065. https://doi.org/10.1523/JNEUROSCI.2806-12.2012

Lenck-Santini PP, Rivard B, Muller RU, Poucet B (2005) Study of CA1 place cell activity and exploratory behavior following spatial and nonspatial changes in the environment. Hippocampus 15(3):356–369. https://doi.org/10.1002/hipo.20060

Lever C, Burton S, Jeewajee A, O'Keefe J, Burgess N (2009) Boundary vector cells in the subiculum of the hippocampal formation. J Neurosci 29(31):9771–9777. https://doi.org/10.1523/JNEUROSCI.1319-09.2009

Lieblich I, Arbib M (1982) Multiple representations of space underlying behavior. Behav Brain Sci 5(4):627–640. https://doi.org/10.1017/S0140525X00013959

Liu YZ, Wang Y, Tang W, Zhu JY, Wang Z (2018) NMDA receptor-gated visual responses in hippocampal CA1 neurons. J Physiol 596(10):1965–1979. https://doi.org/10.1113/JP275094

Llofriu M, Tejera G, Contreras M, Pelc T, Fellous JM, Weitzenfeld A (2015) Goal-oriented robot navigation learning using a multi-scale space representation. Neural Netw 72:62–74. https://doi.org/10.1016/j.neunet.2015.09.006

Luo AH, Tahsili-Fahadan P, Wise RA, Lupica CR, Aston-Jones G (2011) Linking context with reward: a functional circuit from hippocampal CA3 to ventral tegmental area. Science 333(6040):353–357. https://doi.org/10.1126/science.1204622

Malvache A, Reichinnek S, Villette V, Haimerl C, Cossart R (2016) Awake hippocampal reactivations project onto orthogonal neuronal assemblies. Science 353(6305):1280–1283. https://doi.org/10.1126/science.aaf3319

Mamad O, Stumpp L, McNamara HM, Ramakrishnan C, Deisseroth K, Reilly RB, Tsanov M (2017) Place field assembly distribution encodes preferred locations. PLoS Biol 15(9):e2002365. https://doi.org/10.1371/journal.pbio.2002365

Markus EJ, Qin YL, Leonard B, Skaggs WE, McNaughton BL, Barnes CA (1995) Interactions between location and task affect the spatial and directional firing of hippocampal neurons. J Neurosci 15(11):7079–7094

Mizumori SJY, Williams JD (1993) Directionally selective mnemonic properties of neurons in the lateral dorsal nucleus of the thalamus of rats. J Neurosci 13(9):4015–4028

Moser MB, Rowland DC, Moser EI (2015) Place cells, grid cells, and memory. Cold Spring Harb Perspect Biol 7(2):a021808. https://doi.org/10.1101/cshperspect.a021808

Nadel L, Samsonovich A, Ryan L, Moscovitch M (2000) Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. Hippocampus 10(4):352–368. https://doi.org/10.1002/1098-1063(2000)10:4%3c352:AID-HIPO2%3e3.0.CO;2-D

Navratilova Z, Hoang LT, Schwindel CD, Tatsuno M, McNaughton BL (2012) Experience-dependent firing rate remapping generates directional selectivity in hippocampal place cells. Front Neural Circ 6:6. https://doi.org/10.3389/fncir.2012.00006

Neves G, Cooke SF, Bliss TV (2008) Synaptic plasticity, memory and the hippocampus: a neural network approach to causality. Nat Rev Neurosci 9(1):65–75. https://doi.org/10.1038/nrn2303

O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res 34(1):171–175. https://doi.org/10.1016/0006-8993(71)90358-1

Olafsdottir HF, Bush D, Barry C (2018) The role of hippocampal replay in memory and planning. Curr Biol 28(1):R37–R50. https://doi.org/10.1016/j.cub.2017.10.073

Park E, Dvorak D, Fenton AA (2011) Ensemble place codes in hippocampus: CA1, CA3, and dentate gyrus place cells have multiple place fields in large environments. PLoS ONE 6(7):e22349. https://doi.org/10.1371/journal.pone.0022349

Pennartz CM, Lee E, Verheul J, Lipa P, Barnes CA, McNaughton BL (2004) The ventral striatum in off-line processing: ensemble reactivation during sleep and modulation by hippocampal ripples. J Neurosci 24(29):6446–6456. https://doi.org/10.1523/JNEUROSCI.0575-04.2004

Poucet B, Hok V (2017) Remembering goal locations. Curr Opin Behav Sci 17:51–56. https://doi.org/10.1016/j.cobeha.2017.06.003

Rolls ET, Xiang JZ (2005) Reward-spatial view representations and learning in the primate hippocampus. J Neurosci 25(26):6167–6174. https://doi.org/10.1523/JNEUROSCI.1481-05.2005

Royer S, Sirota A, Patel J, Buzsaki G (2010) Distinct representations and theta dynamics in dorsal and ventral hippocampus. J Neurosci 30(5):1777–1787. https://doi.org/10.1523/JNEUROSCI.4681-09.2010

Rueckemann JW, Buffalo EA (2017) Spatial responses, immediate experience, and memory in the monkey hippocampus. Curr Opin Behav Sci 17:155–160. https://doi.org/10.1016/j.cobeha.2017.08.008

Sanders H, Ji D, Sasaki T, Leutgeb JK, Wilson MA, Lisman JE (2019) Temporal coding and rate remapping: representation of nonspatial information in the hippocampus. Hippocampus 29(2):111–127. https://doi.org/10.1002/hipo.23020

Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser MB, Moser EI (2006) Conjunctive representation of position, direction, and velocity in entorhinal cortex. Science 312(5774):758–762. https://doi.org/10.1126/science.1125572

Schlesiger MI, Boublil BL, Hales JB, Leutgeb JK, Leutgeb S (2018) Hippocampal global remapping can occur without input from the medial entorhinal cortex. Cell Rep 22(12):3152–3159. https://doi.org/10.1016/j.celrep.2018.02.082

Scholkopf B, Mallot HA (1995) View-based cognitive mapping and path planning. Adapt Behav 3(3):311–348. https://doi.org/10.1177/105971239500300303

Sekeres MJ, Winocur G, Moscovitch M (2018) The hippocampus and related neocortical structures in memory transformation. Neurosci Lett 680:39–53. https://doi.org/10.1016/j.neulet.2018.05.006

Singer AC, Frank LM (2009) Rewarded outcomes enhance reactivation of experience in the hippocampus. Neuron 64(6):910–921. https://doi.org/10.1016/j.neuron.2009.11.016

Smith DM, Mizumori SJ (2006) Hippocampal place cells, context, and episodic memory. Hippocampus 16(9):716–729. https://doi.org/10.1002/hipo.20208

Solstad T, Boccara CN, Kropff E, Moser MB, Moser EI (2008) Representation of geometric borders in the entorhinal cortex. Science 322(5909):1865–1868. https://doi.org/10.1126/science.1166466

Speakman A, O'Keefe J (1990) Hippocampal complex spike cells do not change their place fields if the goal is moved within a cue controlled environment. Eur J Neurosci 2(6):544–555. https://doi.org/10.1111/j.1460-9568.1990.tb00445.x

Strosslin T, Sheynikhovich D, Chavarriaga R, Gerstner W (2005) Robust self-localisation and navigation based on hippocampal place cells. Neural Netw 18(9):1125–1140. https://doi.org/10.1016/j.neunet.2005.08.012

Taube JS (1995) Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. J Neurosci 15(1 Pt 1):70–86

Tryon VL, Penner MR, Heide SW, King HO, Larkin J, Mizumori SJY (2017) Hippocampal neural activity reflects the economy of choices during goal-directed navigation. Hippocampus 27(7):743–758. https://doi.org/10.1002/hipo.22720

Valdes JL, McNaughton BL, Fellous JM (2015) Offline reactivation of experience-dependent neuronal firing patterns in the rat ventral tegmental area. J Neurophysiol 114(2):1183–1195. https://doi.org/10.1152/jn.00758.2014

Wilber AA, Clark BJ, Forster TC, Tatsuno M, McNaughton BL (2014) Interaction of egocentric and world-centered reference frames in the rat posterior parietal cortex. J Neurosci 34(16):5431–5446. https://doi.org/10.1523/JNEUROSCI.0511-14.2014

Wilson MA, McNaughton BL (1993) Dynamics of the hippocampal ensemble code for space. Science 261(5124):1055–1058

Xiao Z, Nagl S, Lin K, Fellous J-M (2019) Continuous reward-place coding properties of dorsal distal CA1 hippocampus cells. Society for Neuroscience, Chicago

Zhang S, Manahan-Vaughan D (2015) Spatial olfactory learning contributes to place field formation in the hippocampus. Cereb Cortex 25(2):423–432. https://doi.org/10.1093/cercor/bht239
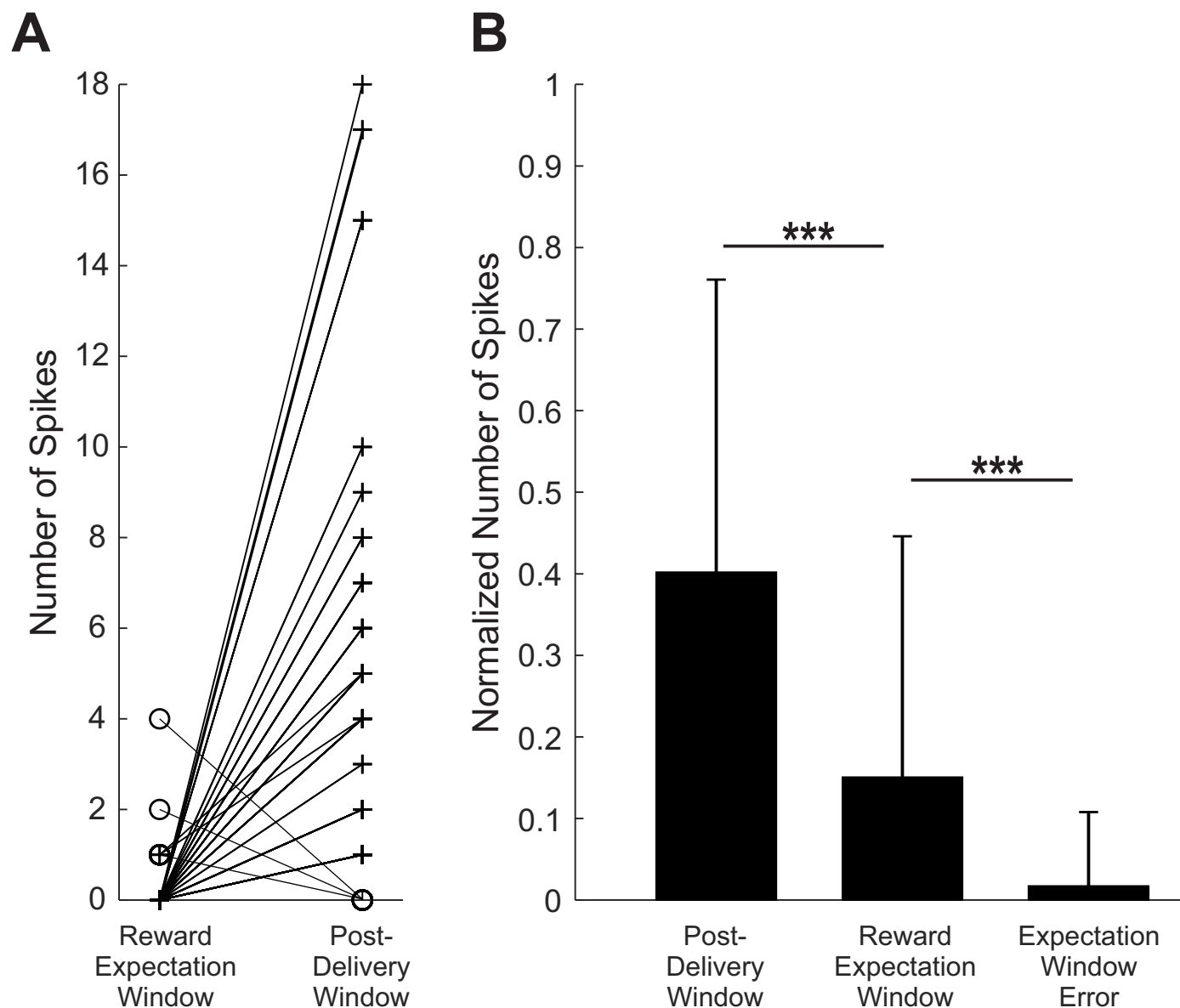
Supplemental Figure 1.1
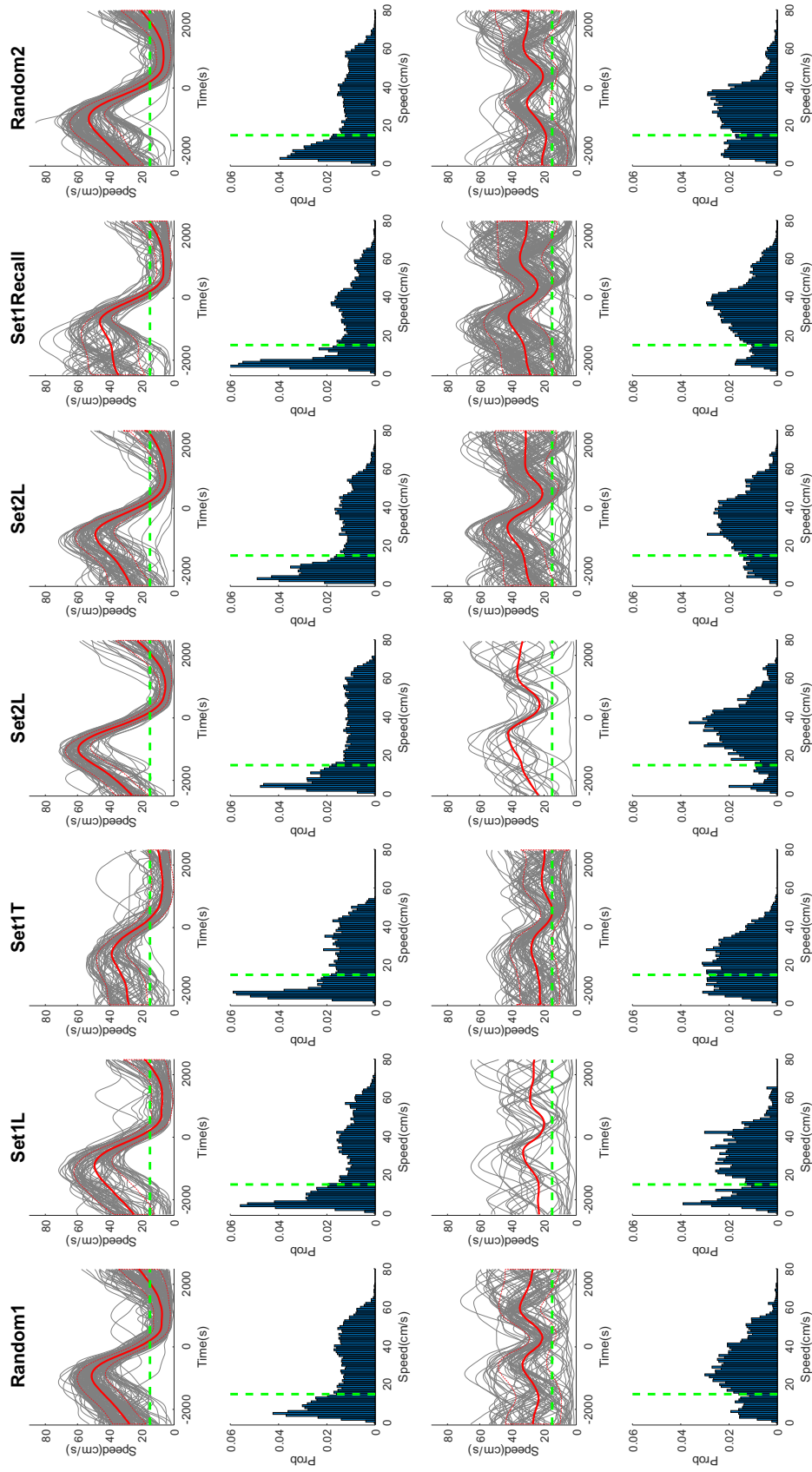
Supplemental Figure 1.1: Animal speed in each task in one experiment. **A:** Rewarded feeders. In each panel, time 0 indicates the time point when the reward was delivered. The top row show the instantaneous speed around the feeders (1s smoothing window), the red thick lines represent the average. The bottom row shows the histogram of the instantaneous speed throughout the corresponding tasks. **B:** same as A for cases when reward were not delivered. The 15 cm/s threshold is indicated in all panels by a green dashed line.

Supplemental Figure 1.2
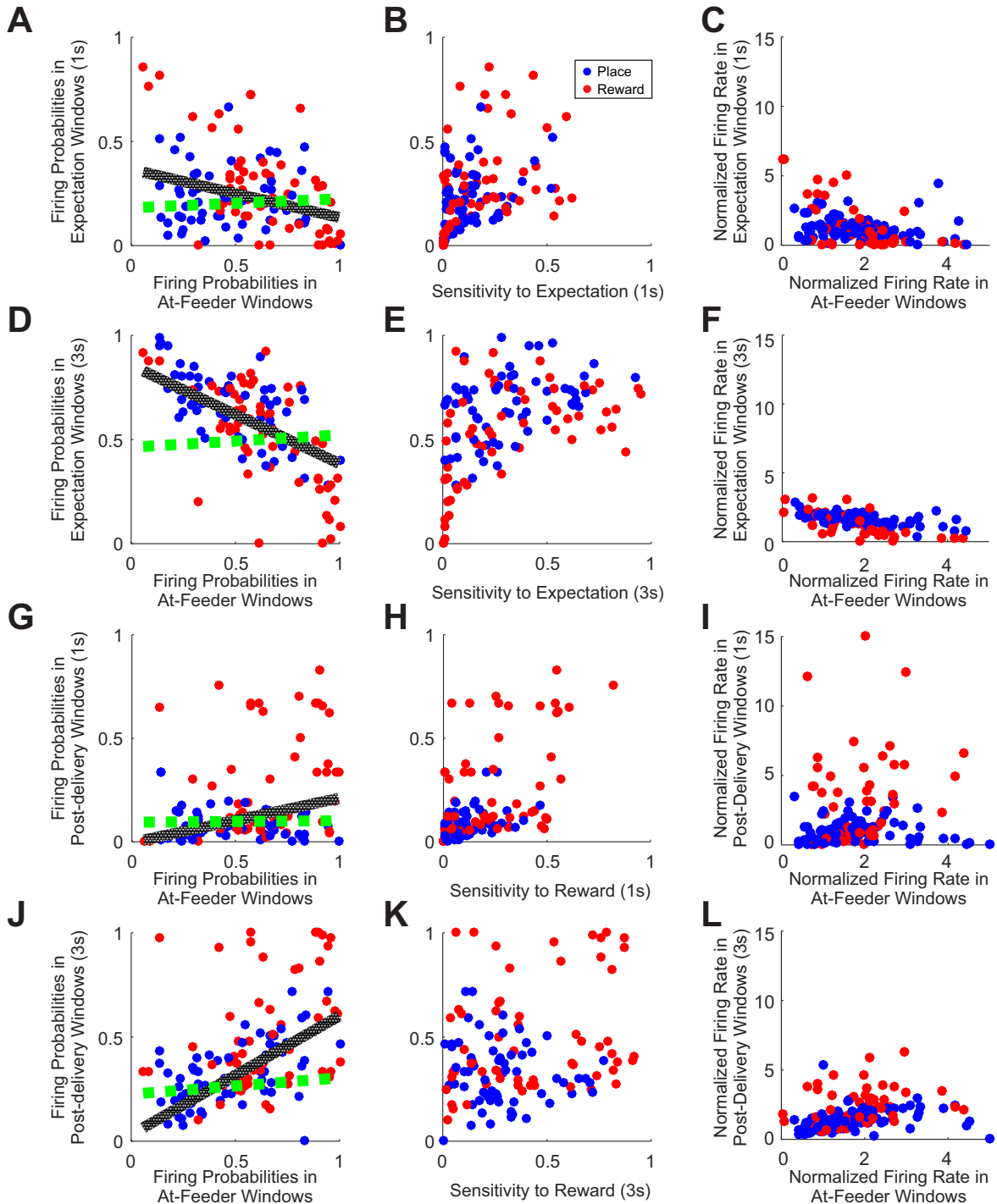
Supplemental Figure 1.2: Extension of Figure 5. The subpanels A-D of Figure 5 correspond to D, E, J, K in this figure. To assess the choice of the 3-second time window size in Figure 5, we add the firing probability analysis (A, G) as well as sensitivity/specificity analysis (B, H) for 1-second windows. We compare the normalized firing rates in the corresponding time windows by the firing rates in randomly chosen time windows with the same window size (C, F, I, L), as in Figure 5. The subpanels in the first 2 columns reveal that delivery preference was insensitive to the choice of window size. Also, the normalized firing rates of reward cells are larger than those of place cells, especially for the 1-second case (I).

Supplemental Figure 1.3

Supplemental Figure 1.3: Comparison of the number of spikes in different time windows for reward cells. **A:** Number of spikes of one example reward cell (Figure 2A, fourth row) in reward-expectation windows and the post-delivery windows after every visit to a rewarded feeder. The visits with more spikes in post-delivery windows are marked by plus signs (n = 81), and the visits with more spikes in expectation windows are marked by circles (n = 22). **B.** Number of spikes in windows normalized by their maximum task rate, for all reward cells. A paired t-tests reveals that the numbers of spikes in post-delivery windows are larger than the numbers of spikes in reward-expectation windows (p<0.001). In addition, an unpaired t-test reveals that the latter are also larger than the number of spikes in the expectation windows when visiting an incorrect feeder (p<0.001).
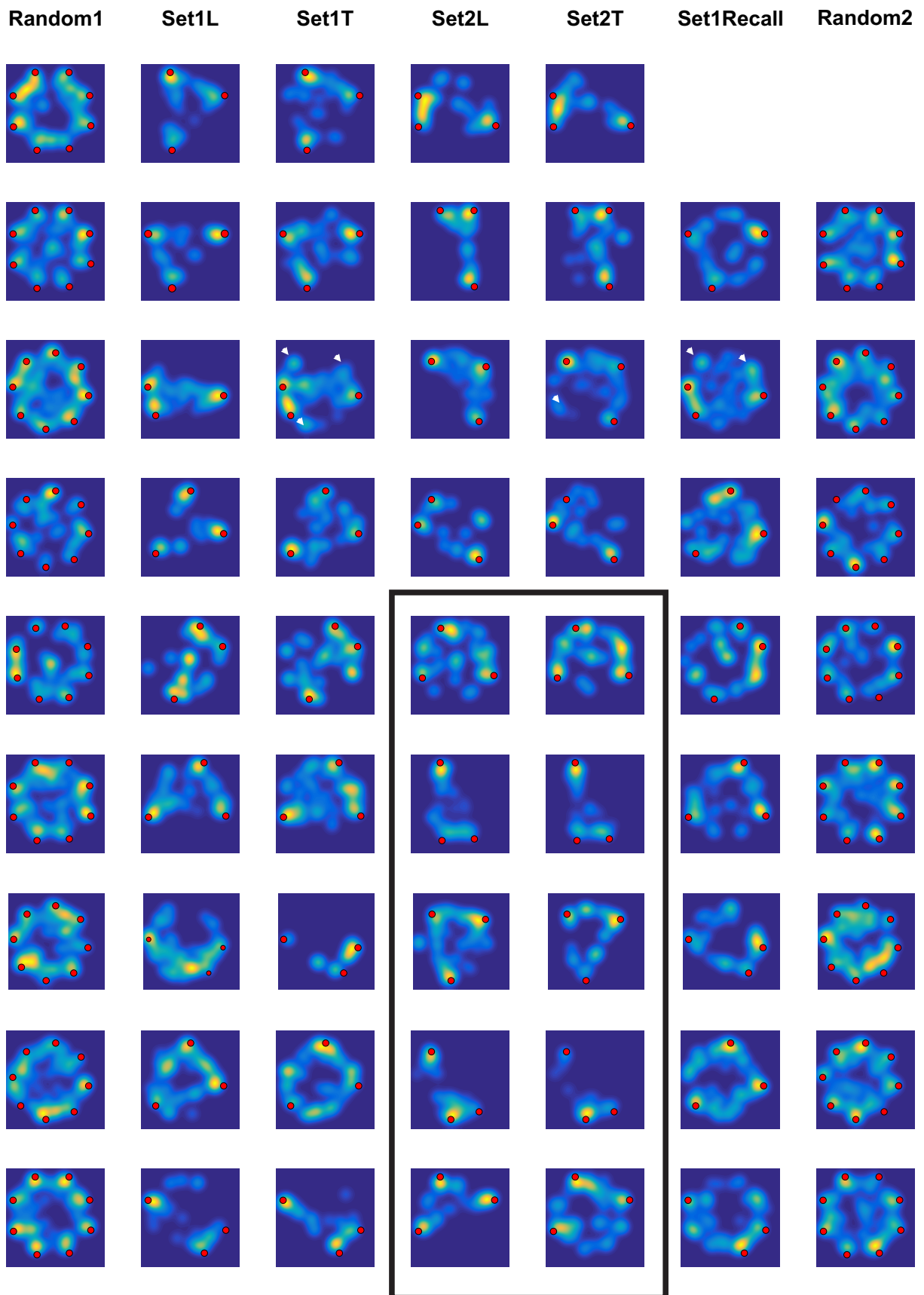
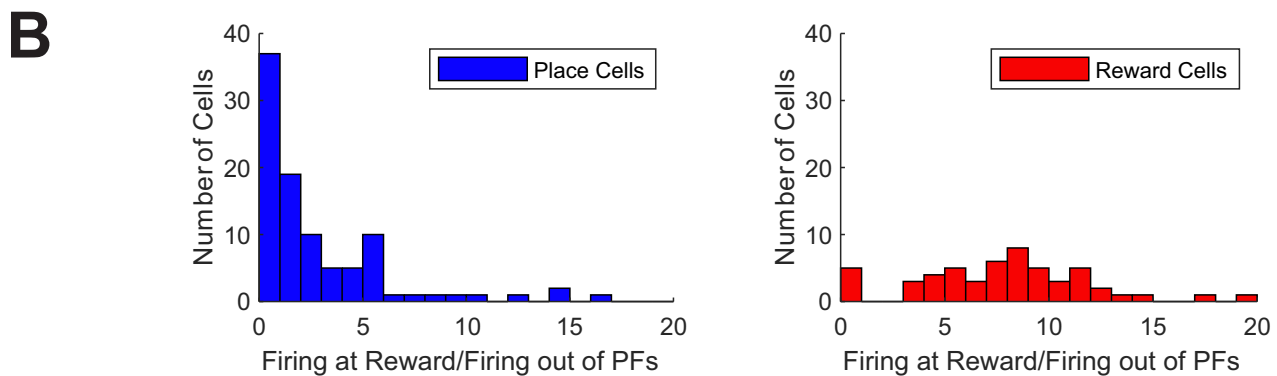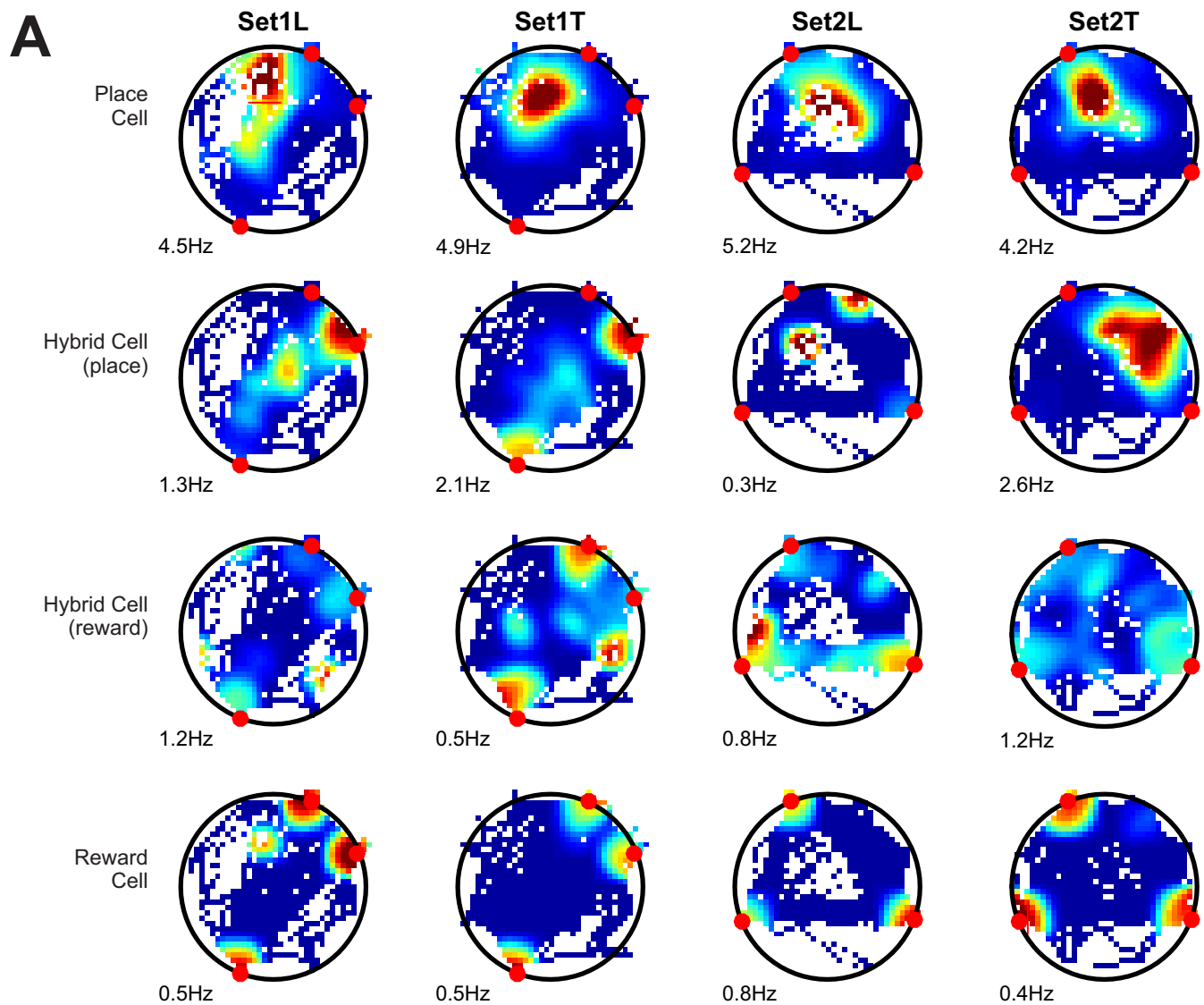# Supplemental Figure 2.1

A

B
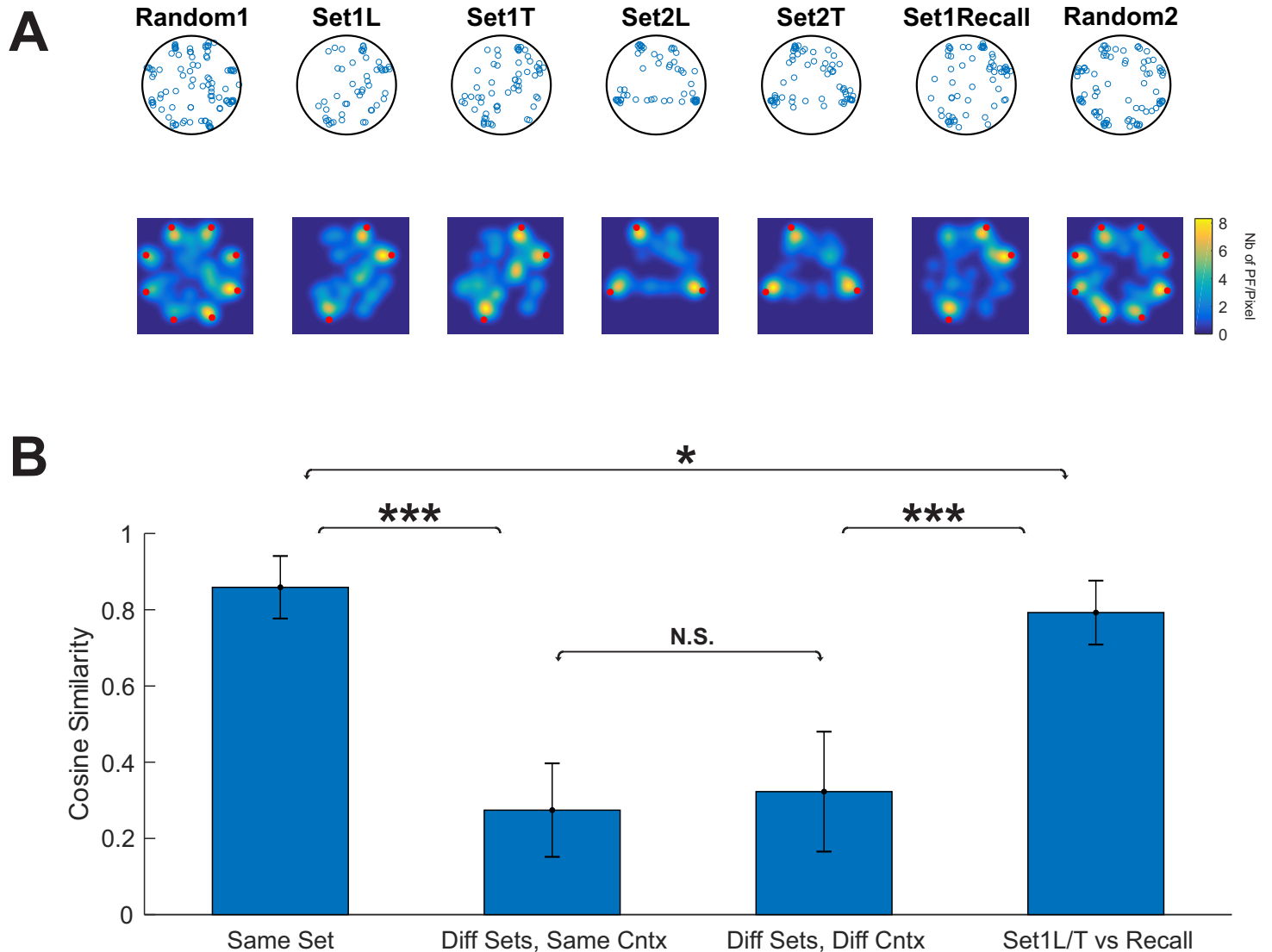
Supplemental Figure 2.1. Density of place fields. **A**. Scatter plots of the location of the center of mass of place fields of all cells recorded in one experiment (upper row). The task epochs are displayed separately for learning (e.g. Set1L) and testing (e.g. Set1T) phases. The lower row shows the density of place field locations shown in A, where the red dots indicate the location of correct feeders. **B.** Cosine similarities between density maps withing the same sets (left), between different sets in the same context conditions (2nd bar), between different sets in different context conditions (3rd bar) and between Set and Recall epochs of the same session (right-most bar). **C.** (next page) The density of place field locations for the other nine same/different context sessions is shown, one session per row. There is no data for the last two tasks in one of the sessions due to a recording failure. Phases in Set2 for different-context sessions are marked by the black box.

Supplemental Figure 2.1 (cont)

C

| Random1	Set1L	Set1T	Set2L	Set2T	Set1Recall	Random2

Supplemental Figure 2.2

Supplemental Figure 2.2 A: Re-plotting of the fields for the four cells presented in Figure 2A with separation of the learning and testing phases. Though not identical, the place maps for different phases in the same tasks are very similar. **B:** We computed the baseline level of the place fields as the average value of the place map outside the cell's activity near the reward. We also computed the average value of the map near the correctly rewarded feeder. The ratios of these values are plotted separately for place cells and reward cells (x-axes).
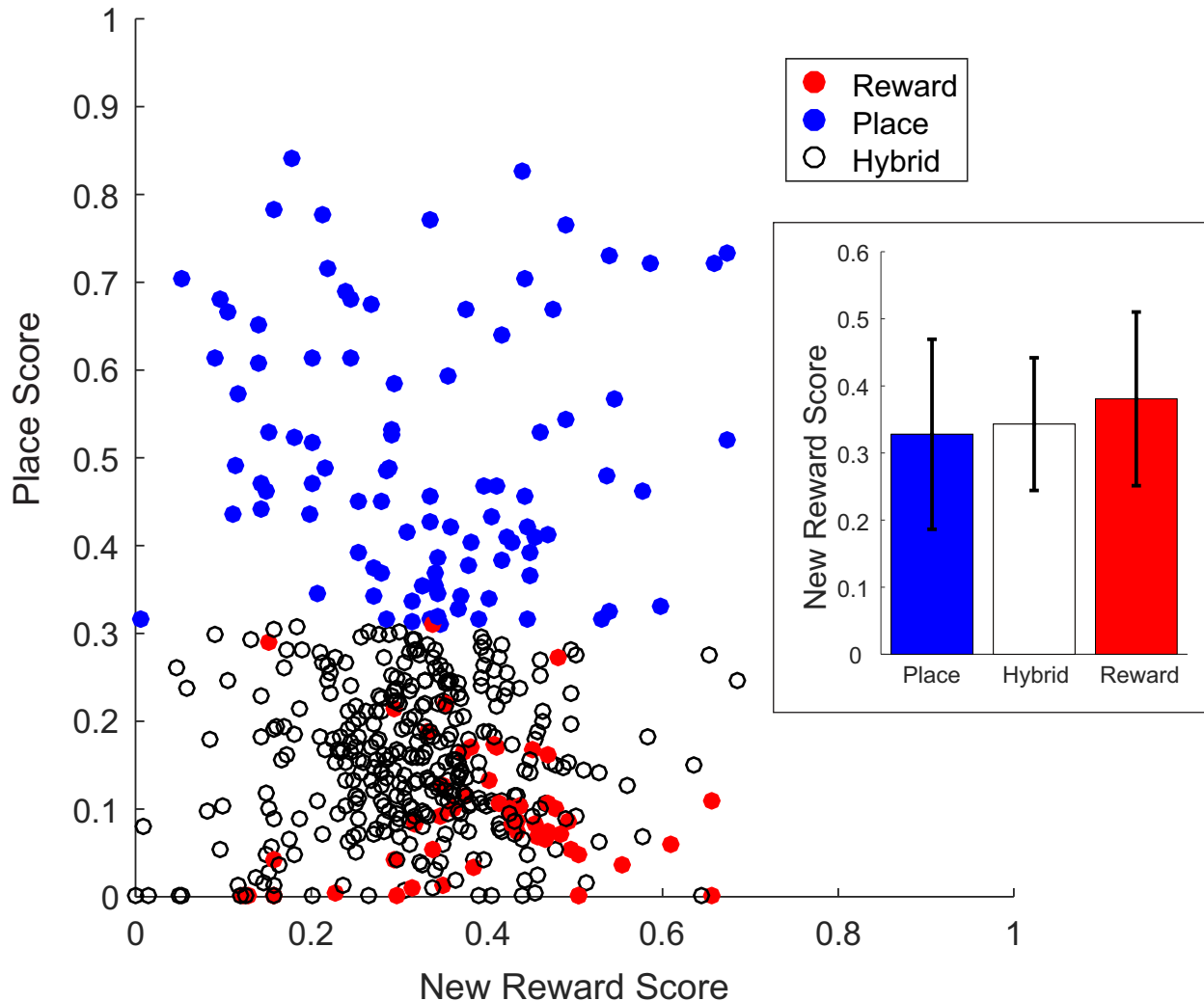
Supplemental Figure 2.3
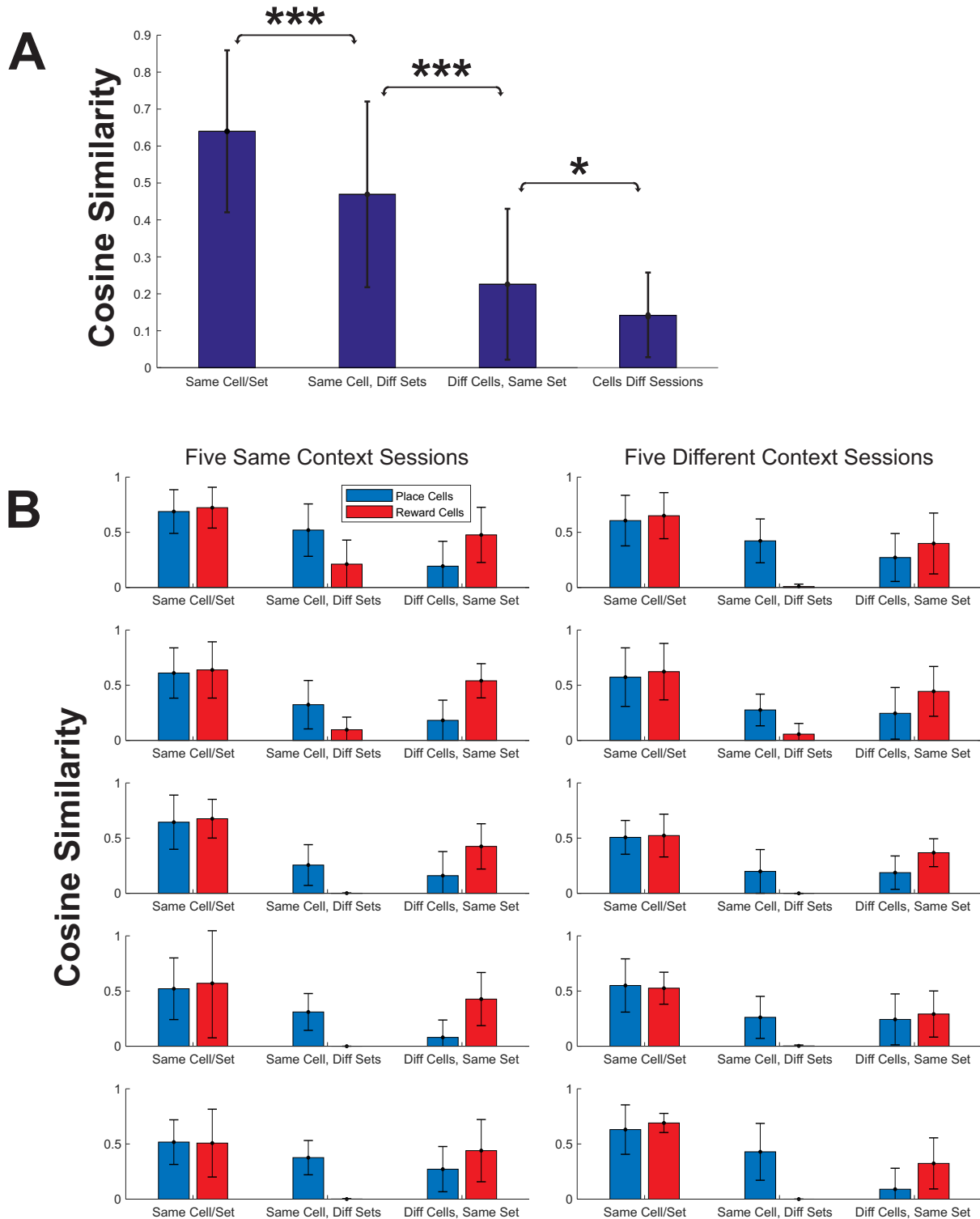
Supplemental Figure 2.3. Cosine similarity of place maps. **A:** Averaged similarity for the same cells in the same and different sets (left 2 bars), and across cells (right 2 bars). **B:** Cosine similarity between place maps for each session for place and reward cells separately. For all sessions, the cosine similarity between the place maps of the same cells was high for both place and reward cells when computed between learning/testing phases of the same sets (left-most bar pairs) . The cosine similarity of same cells in different sets was lower for reward cells than for place cells due to the change of reward locations between sets. The similarity between different reward cells in the same sets had higher similarity than place cells, since reward cells fired around the same feeders.
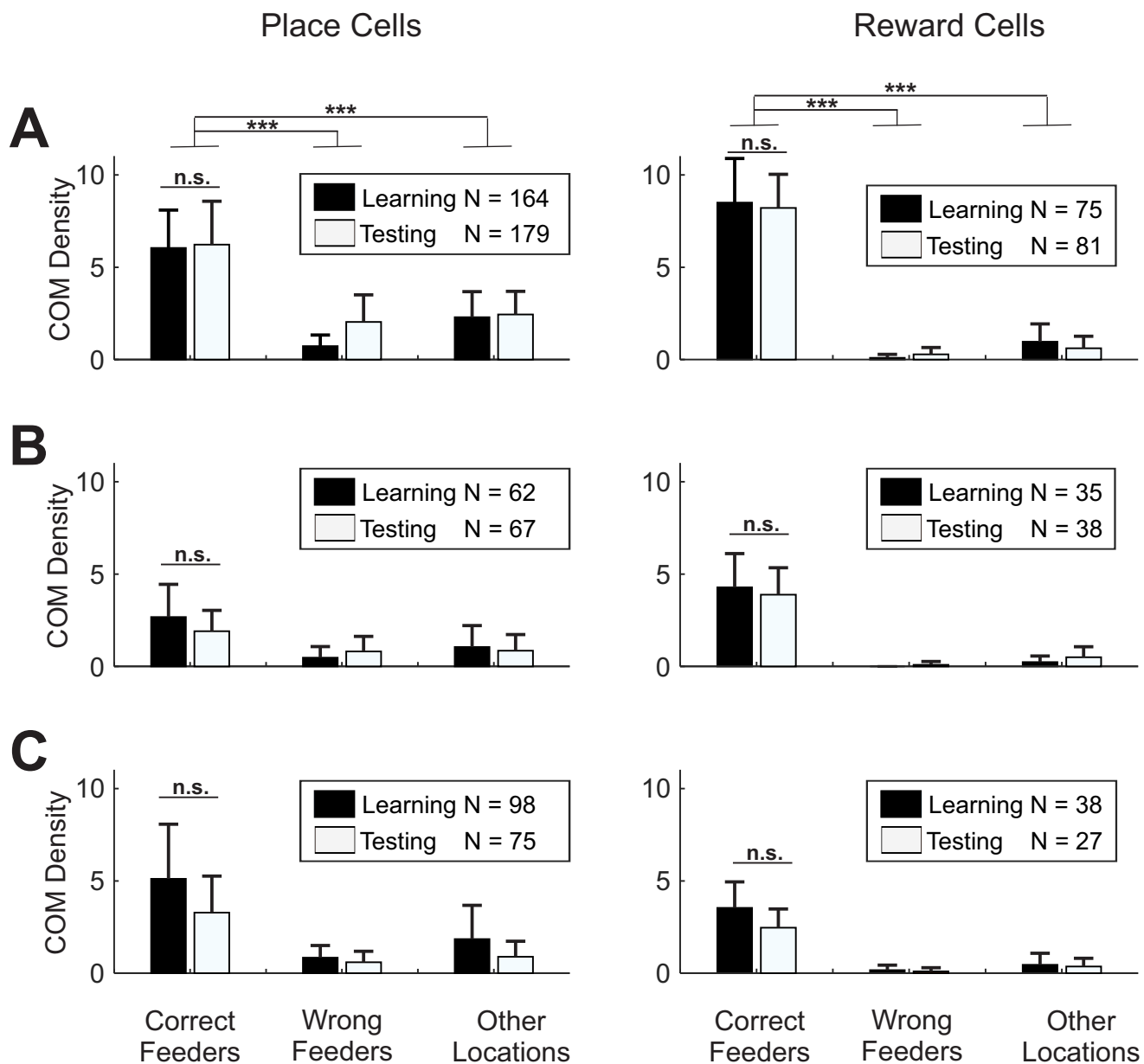
Supplemental Figure 2.4

Supplemental Figure 2.4: Replotting Figure 2 with a new reward score that measures the correlation with single feeders but not all correct feeders. Different colors indicate the categories of cells we previously determined, and we found that under the new definition of reward score, the reward cells become inseparable from hybrid or place cells. Unlike with the skewed distribution found in Fig 2, no group of cells with high reward score and low place score can be found. The average new reward scores for all three cell categories are included in inset.

Supplemental Figure 2.5

Supplemental Figure 2.5: Density of fields Center of Mass around correct feeders, wrong feeders, and other locations on the maze. **A**: Density of subfields center of mass for Place cells (left) and Reward cells (right) in the Set1 learning/testing experiment (10 experiments). Both cell types exhibit higher density of place fields around the correct feeders than around error feeders or other locations on the maze. No difference was found between the learning (cue driven) and testing (memory driven) phases of the task. **B**: same as A for Set2 in the same-context experiments only (5). **C**: same as A for Set2 in the different-context experiments only (5). ***=p<0.001. The pattern of significance in B and C were identical to that in A (not shown for clarity). N indicates the number of COM entered in the analyses.