Readout of a spatial memory by interruption of ramping activity in the retrohippocampal cortex

Sarah A. Tennant[1,5], Harry Clark[1,5], Ian Hawes[1,5], Wing Kin Tam[1,2], Junji Hua[1], Wannan Yang[1,4], Klara Z. Gerlei[1], Emma R. Wood[1,2], Matthew F. Nolan[1,2,3]

[1]Centre for Discovery Brain Sciences, [2]Simons Initiative for the Developing Brain, [3]Centre for Statistics, University of Edinburgh, Edinburgh, United Kingdom, [4]Current address: Center for Neural Science, New York University, New York, United States. [5]Equal contributing first authors.

Abstract

Neurons in the retrohippocampal cortices play crucial roles in spatial memory. Many retrohippocampal neurons have firing fields associated with specific locations, but whether this is the sole strategy for representing spatial memories is unclear. Here, we demonstrate that during a spatial memory task retrohippocampal neurons also encode location through ramping activity within discrete regions of a track, with rewarded locations represented by offsets or switches in the slope of the ramping activity. These representations do not require cues that mark reward locations indicating that they result from recall of the track structure. Similar representations emerged in recurrent neural networks trained to navigate to a rewarded location. We hypothesize that retrohippocampal ramping activity and its interruption mediates readout of learned models for goal-directed navigation.

Introduction

Spatial memory relies on complex brain networks, with neurons in the hippocampus and associated retrohippocampal cortices playing essential roles. Neurons in these brain areas have receptive fields selective for location (Figure 1A). For example, place cells in the hippocampus and grid cells in retrohippocampal cortices have been widely studied (Hafting et al., 2005a; Moser et al., 2017; O'Keefe and Dostrovsky, 1971), and are central to theories of navigation and memory (e.g. (Eichenbaum, 2000, 2017; McClelland et al., 1995; O'Keefe and Nadel, 1978; O'Reilly and Rudy, 2001; Whittington et al., 2019)). However, theoretical studies suggest that continuous coding schemes in which location is directly proportional to neuronal firing rate (Figure 1B) may also be suited to control of navigational behaviours (Burgess and O'Keefe, 1996; Bush et al., 2015; Fiete et al., 2008; Foster et al., 2000; Gershman, 2014; Gustafson and Daw, 2011; Stachenfeld et al., 2017; Stemmler et al., 2015). While in principle such codes could be generated downstream from place or grid cells (Burgess and O'Keefe, 1996; Bush et al., 2015; Fiete et al., 2008; Foster et al., 2000; Gershman,

2014; Gustafson and Daw, 2011; Stachenfeld et al., 2017; Stemmler et al., 2015), it is unclear whether neurons in these circuits represent information using continuous codes or if coding schemes of this kind can be engaged by recall of memories.

Results

To test whether continuous coding schemes are engaged during recall of learned environments we used a location memory task in which external sensory cues can be controlled on a trial by trial basis (Tennant et al., 2018)(see Methods). The task requires recall of the location of a reward zone on a virtual linear track (Figure 1C)(Tennant et al., 2018). On standard trials the reward zone is marked by a prominent visual cue but otherwise the track lacks unique spatial cues. With training, mice efficiently obtain rewards by running directly to the reward zone (Figure 1D, Supplemental Figure 1). This spatial behaviour is maintained when the visual cue marking the reward zone is removed, indicating that mice can solve the task by comparing estimates of their distance from the start of the track obtained using a path integration strategy with estimates stored in memory (Tennant et al., 2018); while the path integration process may be driven either by sensory input or internal command signals, this alone is insufficient for accurate stopping unless the mouse also remembers where to stop. As mice carried out the task we recorded neural activity using tetrodes that were targeted to the retrohippocampal cortices (Supplemental Figure 2), which include the entorhinal cortex, presubiculum and parasubiculum. Following spike sorting we obtained 1499 isolated single units recorded during sessions from trained mice that contained > 30 correct trials (193 sessions, 10 mice).

Rewarded locations are encoded by interruption of ramp-like firing rate trajectories

We find that on trials in which the reward zone is visible (beaconed trials) a majority of recorded retrohippocampal neurons generated ramping activity within discrete regions bounded by the start of the track and the reward zone, and by the reward zone and the end of the track (n = 940 / 1499 neurons)(Figure 1E). We found many neurons for which movement through the reward zone was associated with a striking interruption in the ramping trajectory, which manifested either as a change in the slope of the ramp or as a resetting of the spike rate (Figure 1E-G and Supplemental Figure 3).

To quantitatively evaluate ramp firing and its interruption at the reward zone, we first classified neurons on the basis of their activity in the track segments before and after the reward zone (Figure 1E-G); while there was a slight bias towards neurons with negative ramp slopes, all possible combinations of trajectory (positive ramps, negative ramps or stationary activity) were observed and ramp slope on one part of the track did not predict ramping properties elsewhere (Supplemental Figure 3C). Classifications were similar when ramping activity was identified by fitting a linear model ($p < 0.01$ after correction for multiple comparisons, see Methods) or using a ‘ramp score’ based on

correlations between position and firing rate (Figure 1G and Supplemental Figure 4). Ramp-like activity was present in similar proportions of neurons from all animals (mean 64.4 ± 6.28 %, range 55 - 71.88), had firing rate peaks for ramps with positive slopes (or minima for ramps with negative slopes) that were strongly biased towards locations adjacent to the reward zone (Supplemental Figure 4E), was independent of theta modulation of spiking activity (Supplemental Figures 5), was similar in medial entorhinal cortex and pre/para-subiculum (Supplemental Figures 6), and was absent in 99.4% of shuffled data sets (n = 1488273 / 1497501 shuffles)(Supplemental Figure 7). For most ramping neurons their activity could best be accounted for by encoding of position, either alone or with position as the dominant coefficient alongside speed and / or acceleration (n = 618 / 940 neurons)(Supplemental Figures 8, 9). Ramping activity was also better accounted for by encoding of position rather than time (Supplemental Figure 10). Ramping trajectories on trials in which mice ran through the reward zone without slowing to stop were also consistent with encoding of location (Supplemental Figure 11).

Thus, firing rate trajectories in which ramping activity encodes location within discrete track segments and reconfigures at a reward location are a common feature of neural activity in retrohippocampal cortices. These activity patterns contrast with discrete firing fields that are typical of place and grid cell fields recorded previously on linear virtual tracks (Chen et al., 2013; Domnisoru et al., 2013; Harvey et al., 2009)(Supplemental Figure 14).

Ramp-like trajectories and their reconfiguration are maintained in the absence of reward zone cues

To test whether interruption of ramp firing at the reward zone results from recall of a memory of the track structure, we investigated activity on probe trials in which rewards and the visible cues indicating the reward zone were omitted (Figure 2A). We focus on neurons with firing rate trajectories that depend on track position and for which there were sufficient numbers of probe trials (see Methods) to classify their activity (n = 252 / 618 positional neurons, 141 sessions, 10 mice).

For many neurons, encoding of the reward zone, by changes in the slope of ramping activity or by resetting of the spike rate, was maintained on probe trials (Figure 2B). The slope of ramping activity on each track section (Figure 2C-D), and resetting of spike rates at the reward zone (Figure 2E), were both strongly correlated between probe and standard trials. Using our criteria for classification of ramping trajectories, the firing patterns of 113 / 252 neurons (mean 39.9 ± 5.5 %, n = 10 mice) had identical classifications on standard and probe trials (Figure 2G). For those neurons with classification that changed, this was either because the ramp slope approaching the reward zone was reduced, indicating that in these neurons the approaching visual cue drove ramping activity (Supplemental Figure 12), or because the ramp slope after the reward zone was reduced, indicating that the cue or

reward was directly driving resetting of activity (Supplemental Figure 13). For neurons in which firing rate trajectories were maintained on probe trials, this appears unlikely to be explained by encoding of movement variables as position rather than speed or acceleration continued to be the major predictor of firing rate (Supplemental Figure 12C). Thus, recall of the track structure is sufficient for interruption of ramping activity at the reward zone, while ramping activity approaching the reward zone can result from path integration or can be driven by the visual cue. These observations are consistent with activity of neurons in the MEC reflecting readout of an internally generated learned model of the track.

Functional cell types

Retrohippocampal cortices are widely associated with grid, border and head direction representations (Boccara et al., 2010; Diehl et al., 2017; Hafting et al., 2005b; Hardcastle et al., 2017; Sargolini et al., 2006). However, because ramp firing patterns that we report here are recorded during a task in which recall of the track locations is necessary for reward, whereas firing properties of retrohippocampal neurons are usually evaluated in the absence of location-specific reward contingencies, it's unclear which of the previously described functional cell types ramping neurons correspond to. To address this, in a subset of behavioural sessions after recording from neurons in the location memory task we recorded from the same neurons in an open arena (Figure 3A). We used a hierarchical scheme to classify cells (Diehl et al., 2017), with rank grid > border > spatial > head direction > non-spatial. In this scheme, cells were assigned to the highest ranked category for which their identity score was greater than the 99th percentile of their corresponding shuffled data (see Methods). We find that for most neurons with positional ramping activity in the location estimation task, their firing in the open arena was either spatially selective (n = 229 / 618) or non-spatial (n = 312 / 618)(Figure 3B-E and Supplemental Figure 14). Ramping neurons that had border, grid or head direction firing patterns in the open field were found relatively infrequently (n = 53, 14 and 10 / 618 respectively)(Figure 3F, G). Conversely, neurons with border, grid and head direction firing patterns in the open field were rarely classified as ramping neurons on the linear track (Supplemental Figure 14). We also evaluated the correspondence between the activity of ramping neurons and features used to distinguish principal cells from interneurons. The majority of ramping neurons had firing properties characteristic of principal cells, although a small proportion of neurons with negative ramp slopes had firing properties that are typical of interneurons (Supplemental Figure 15). Therefore, the population of neurons with ramping activity is largely, but not exclusively, non-overlapping with grid, head direction and border cell populations, while many neurons that in an open arena lack well defined spatial firing fields may participate in representation of learned information by generating ramp-like representations that interrupt around recalled reward locations.

Recurrent neural networks learn to estimate location using ramp-like representations

If interruption of ramp-like representations at behaviourally important locations reflects a general neural solution to distance estimation, then similar representations may emerge in artificial neural networks trained to solve location estimation tasks. To investigate this, we trained a network model that received visual input corresponding to a virtual track and that generated stop or go signals to control movement along the track. The model was organised such that visual signals were processed by a three layer convolutional neural network with the output fed into a recurrent neural network that generated movement control signals (Figure 4A). Networks trained with a reinforcement learning algorithm (Schulman et al., 2017) successfully learned the task (Figure 4B). Before training, the agent controlled by the network stopped frequently on the track with no particular spatial stopping pattern (Figure 4C). After training, the agent stopped at the reward zone on non-beaconed (Figure 4D) and probe trials (Figure 4E) indicating that the network solves the task using a distance-based strategy. The agent's stopping patterns are remarkably similar to those of mice before and after learning the task (Tennant et al., 2018).

As in experimental recordings from retrohippocampal neurons, units in the trained recurrent network showed ramping activity that was interrupted at the reward zone (Figure 4F, J). The ramping activity of units in the recurrent network was clearly distinguishable from activity patterns generated by shuffled datasets (Figure 4G-H), and had either a positive or a negative slope (Figure 4F, I-J), with the slope on the region of the track approaching the reward zone not predicting the slope after the reward zone (Figure 4I). Also similar to the experimental data, ramps generated by the trained recurrent network more frequently had a negative rather than a positive slope, although in the recurrent networks negative slopes were typically steeper than positive slopes (Figure 4I). The similarity between activity trajectories of units in recurrent neural networks trained on a distance estimation task and activity generated by neurons in the retrohippocampal cortex is consistent with ramping activity and its interruption providing a general mechanism for readout of location information in tasks when specific locations are associated with rewards.

Discussion

Our results show that retrohippocampal neurons employ ramping neural activity and its interruption at salient locations to generate structured representations of learned environments. These representations can be maintained in the absence of location-specific cues indicating that they result from recall of a spatial memory. As the direction of ramping activity differed between segments of the behavioural track, representations of position may be specific to individual segments, potentially enabling ramp-like codes to combinatorially represent location and local context, for example whether the current location is before or after the reward zone. These representations differ from the global metric provided by grid cells (Hafting et al., 2005a; McNaughton et al., 2006), but are reminiscent of robotic control strategies that segment the environment (Kuipers and Kuipers, 1999). The ramp-like

firing within a track segment is also reminiscent of spatial firing of neurons in the ventral striatum in anticipation of rewarded locations (Carelli and Deadwyler, 1994; Fiorillo et al., 2003; Howe et al., 2013; Lavoie and Mizumori, 1994; van der Meer and Redish, 2011; van der Meer et al., 2010; Schultz et al., 1992). However, retrohippocampal activity differs in the diverse combinations of firing rate slopes approaching and following the reward location (Figure 1), and in their insensitivity to whether reward is delivered (Figure 2). These distinct features of ramp-like firing of retrohippocampal neurons may be suited to representation of the structure of the learned environment and location relative to salient landmarks.

According to current theories for spatial computation, the location task we use here could be solved using path integration by grid cells to drive sequential activation of place cells along the track. However, readout of metric representations of distance to the stopping position from discrete representations is challenging without graded intermediate representations (Bush et al., 2015; Fiete et al., 2008; Stemmler et al., 2015). Our results are consistent with proposed solutions to this problem in which the structure of the environment is stored in the weights of synaptic inputs from place or grid cells onto ramping neurons (Bush et al., 2015; Fiete et al., 2008; Stemmler et al., 2015). Ramping activity could also be generated through recurrent connectivity or intrinsic dynamics of entorhinal neurons (Dhillon and Jones, 2000; Egorov et al., 2002; Peng et al., 2017; Yoshida and Hasselmo, 2009), or be driven by inputs from postrhinal neurons (LaChance et al., 2019). Given that deep layers of entorhinal cortex have extensive projections throughout the telencephalon (Insausti et al., 1997; Sürmeli et al., 2015; Swanson and Köhler, 1986), ramping activity could then be used by motor control neurons receiving input from the retrohippocampal cortices, or as a template for ramping activity of neurons in striatal circuits that receive direct inputs from retrohippocampal neurons. Thus, our results suggest a model for memory recall in which a neural path integrator drives ramping neural activity that reconfigures at salient landmarks and that may be suitable for control of behaviours by downstream brain structures.

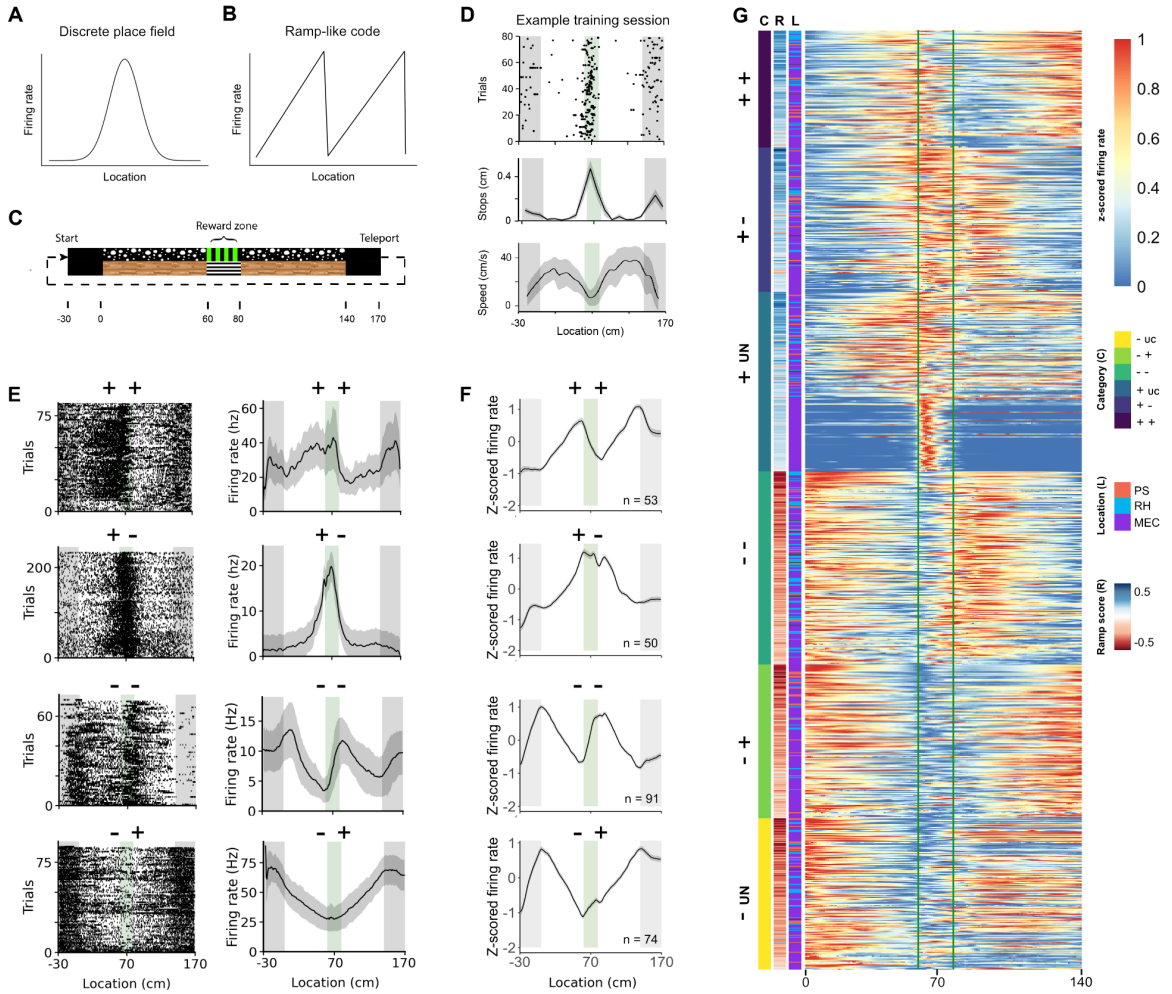


Figure 1. Interruption of ramping activity at a learned reward location

(A, B) Schematic of discrete (A) and continuous (B) firing fields.

(C) Schematic of the virtual track in the beaconed trial configuration. Numbers below the track indicate distance (cm) relative to the start of the track.

(D) Example behaviour of a trained mouse on beaconed trials, showing a raster plot of stopping locations (upper), mean stops (per cm)(middle) and average running speed (lower) each as a function of position on the track. Standard error of the mean is in grey. The vertical green bar indicates the location of the reward zone.

(E) Spike rasters (left) and the corresponding mean firing rate (right) as a function of track position for example neurons. Standard error of the firing rate is in grey. The first and third neurons are recorded from the training session shown in (D). Symbols above the plots indicate classification of the neuron, e.g. + + are neurons with positive slope ramps before and after the reward zone.

(F) Z-scored firing rate as a function of position averaged across neurons classified by their ramp slope before and after the reward zone (symbols above each panel indicate classification).

(G) Normalised firing rates as a function of location along the track (0 - 140 cm) for all neurons that have ramp-like firing in the region of the track before the reward zone. The start and end of the reward zone region is marked with vertical green lines. Neurons are categorized along the y axis by the direction of their slope before and after the reward zone (e.g. + +, - -, etc) and then ordered within each category according to their ramp slope before the reward zone. Columns to the left indicate the ramp score (R), the classification of activity based on fit of the linear model (C) and the location within the retrohippocampal region (L).

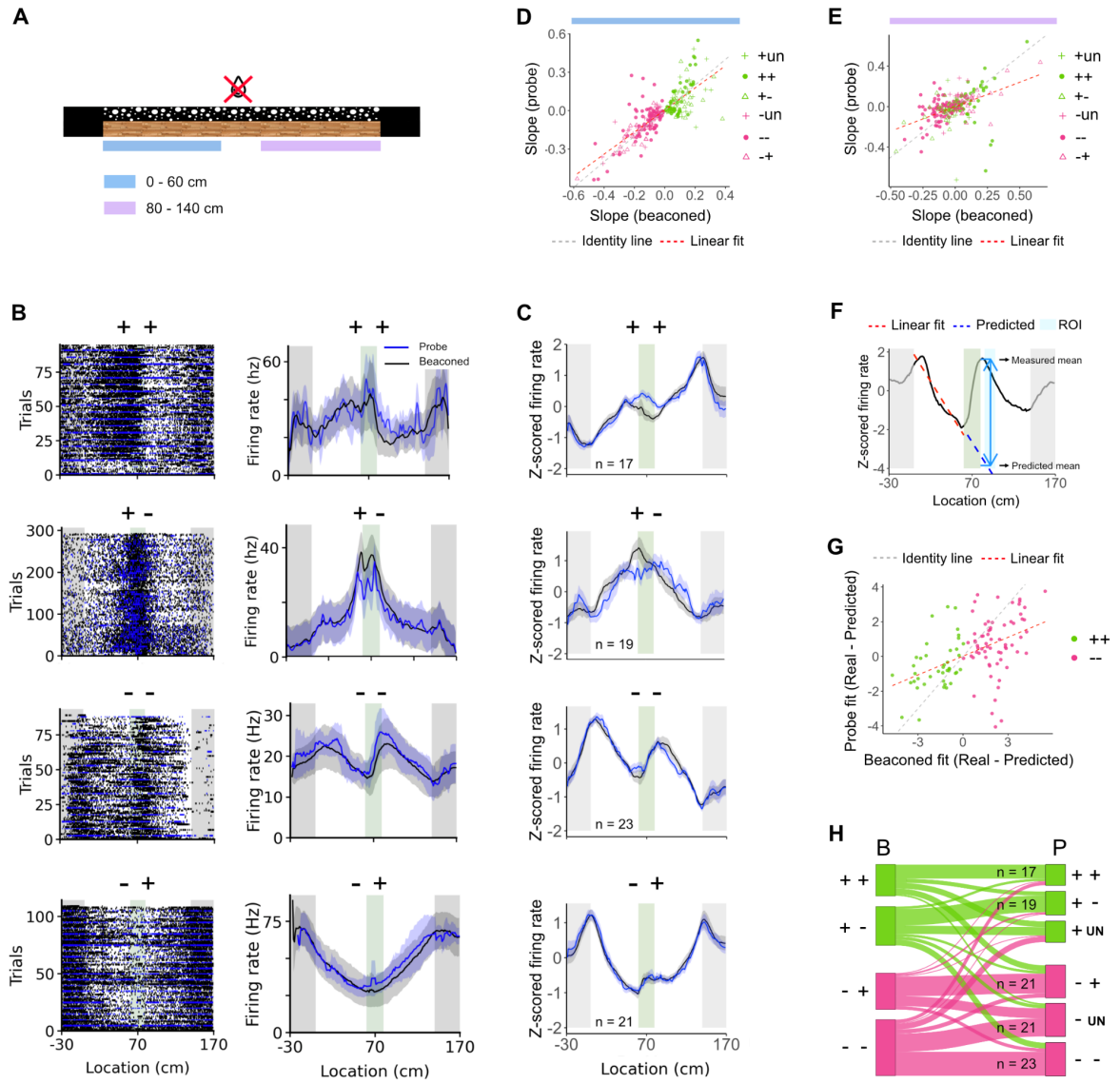


Figure 2. Location cues are not required for reconfiguration of ramping trajectories at the reward zone

(A) Schematic of track configuration. These analyses consider only probe trials; visual cues indicating the reward zone are absent and no reward is delivered. Coloured regions indicate the segments of the track analysed in C and D.

(B, C) Individual examples (B) and population averages (C) of firing rate as a function of track position for neurons with ramps that change direction (+- and -+) or reset (++ and --) between track segments before and after the reward zone on beaconed (black) and probe trials (blue). Solid lines indicate the mean and shaded areas indicate the SEM.

(D, E) Slope of the firing rate before (D) and after (E) the reward zone region during beaconed versus probe trials. The dashed line indicates points for which the slope would be identical on both trial types. Slopes were correlated between beaconed and probe trials in the track segments before ($r^2 = 0.63$, p

$< 2.2 \times 10^{-16}$) and after ($r^2 = 0.21$, $p < 2.08 \times 10^{-15}$) the reward zone. Symbols indicate classification of individual neurons.

(F) Offsets in firing rate were calculated by subtracting the actual firing rate in the first 5 cm of the track following the reward zone with the firing rate predicted by extrapolating the fit in the segment approaching the reward zone.

(G) Firing rate offsets for position encoding neurons classified on beaconed trials as having ++ or -- firing profiles. Offsets were correlated between beaconed and probe trials ($r^2 = 0.22$, $p = 1.63 \times 10^{-7}$).

(H) Classification of neurons based on the slope of their activity on track segments before and after the reward zone on beaconed (B) and probe trials (P). N number is displayed for groups containing 10 or more cells.

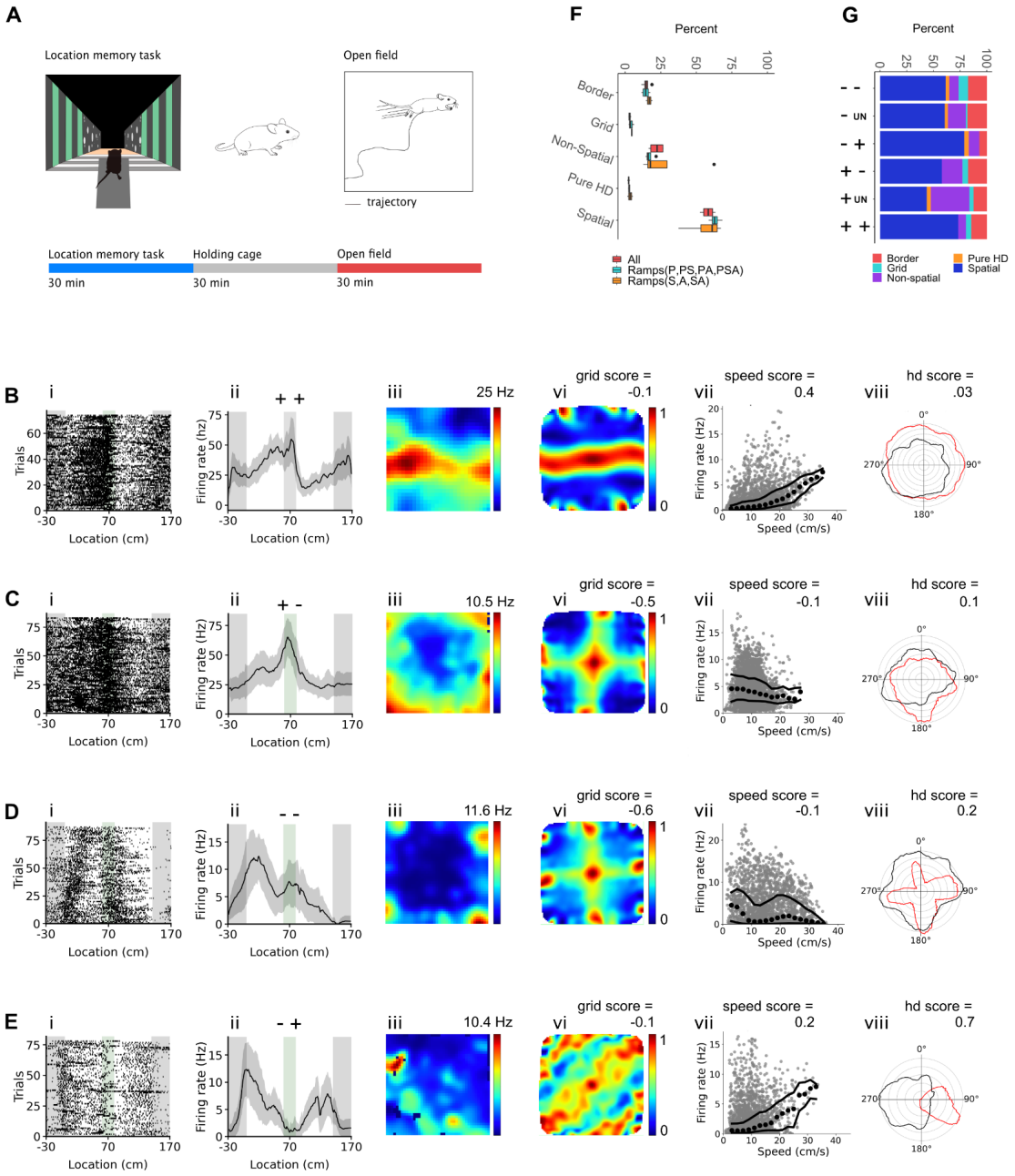


Figure 3. Ramping neurons have diverse spatial firing patterns in an open arena

(A) Animals performed the location estimation task for ~30 minutes, were placed in a holding arena for ~30 minutes and then explored the open arena for ~30 minutes.

(B-E) Example activity in the open arena of cells with ramping activity on the linear track characterised by +ve slopes approaching and following the reward zone (B), +ve slope approaching and -ve slope following the reward zone (C), -ve slopes approaching and following the reward zone (D) and -ve slope approaching and +ve slope following the reward zone (E). Panels show spike rasters (i) and average firing rate (ii) as a function of position on the virtual track, and spatial heat maps of firing rates

(iii), spatial autocorrelograms (vi), firing rate as a function of speed (vii) and polar plots of average firing rate as a function of head direction (red) and movement (black) during exploration of the open arena (viii).

(F) Box and whisker plot showing proportions of neurons classified in the open field as either border, grid, head-direction, spatial or non-spatial. Proportions for all neurons (n = 1499), position (P) dependent ramp cells (n = 618) and non-position (Speed (S) and acceleration (A)) dependent ramp cells (n = 243), as described in Supplemental Figure 8, are shown side by side. (G) Proportion of position-dependent ramp neurons that meet the criteria to be classified in the open field as either border, grid, head-direction, spatial or non-spatial for each of the ramp categories defined in Figure 1 (++, +-, -+, --).

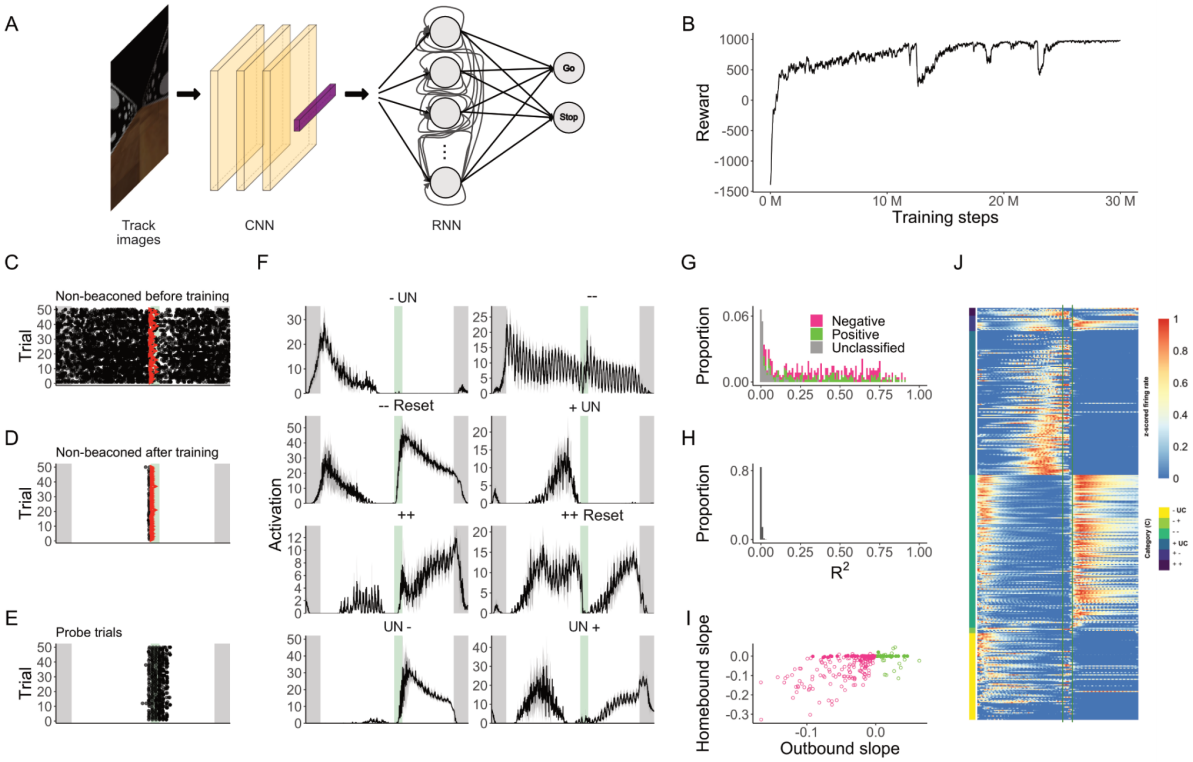


Figure 4. Recurrent networks trained to estimate location generate ramp-like representations

(A) Schematic of the network. Images of the track were played into a convolutional neural network (CNN) that sends outputs to a recurrently connected network (RNN) that sets the activity of movement control neurons. On each simulation step the network can either select GO, which advances the agent, or STOP, which causes the agent to remain at the same location. During training a reinforcement learning algorithm modifies the network weights according to whether the agent stops at the reward zone.

(B) Reward obtained as a function of training step. If the agent stops in a non-goal location, it receives a reward of -1, whereas moving forwards is a neutral action with a reward of 0, and stopping in the goal locations yields a reward of 1000. Dips in the reward occur when the agent tests new (unsuccessful) navigation strategies. After developing a successful strategy during training, the agent typically takes ~ 250 simulation steps to reach the end of the track.

(C-D) Rasters of stopping locations on non-beaconed trials before (C) and after (D) the agent has learned the task.

(E) Rasters of stopping locations on probe trials after the agent has learned the task.

(F) Examples of RNN unit activity as a function of track location after the agent has learned the task.

(G) R^2 values for fits of activity of RNN units as a function of position in the track segment before the reward zone. Units are classified as in Figure 2.

(H) As for (G) but for shuffled data.

(I) Slopes of the firing rate of RNN units after the reward zone as a function of slopes before the reward zone.

(J) Normalised firing rate as a function of position on the track for all units in the RNN.

Materials and methods

Experimental models

All animal experiments were carried out under a project licence granted by the UK Home Office, were approved by the Animal Welfare and Ethical Review Board of the University of Edinburgh School of Medicine and Veterinary Medicine, and conformed with the UK Animals (Scientific Procedures) Act 1986 and the European Directive 86/609/EEC on the protection of animals used for experimental purposes. Experiments used ten male mcos wild type mice 7-14 weeks (mean = 12.6, SD = 0.71) old at the time of the implant surgery. Males were used for consistency in their size which was important for operation of the behavioural apparatus. One week prior to surgery animals were group housed (3-5 mice per cage) in a holding room on a reversed 12 hour on/off light cycle (light from 7 pm to 7 am). After surgery mice were singly housed in cages that contained a running wheel and cardboard igloo for additional enrichment. Prior to surgery and for 2-3 days following surgery mice were given standard laboratory chow and water *ad libitum*.

Experimental methods

Microdrive fabrication

Microdrive fabrication for in vivo electrophysiology recordings was similar to our previous work (Gerlei et al., 2020). 16-channel microdrives consisting of 4 tetrodes were built by threading 90% platinum 10% iridium wire (18µm HML-coated, Neuralynx) to an EIB-16 board (Neuralynx) via an inner cannula (21 gauge 9 mm long). The board was then covered in epoxy and once dried a poor lady frame (Axona) was cemented to the side. An outer cannula (17 gauge 7 mm), sanded at a 45 degree angle on one end to fit the curvature of the skull, was placed around the inner cannula. The outer cannula was secured temporarily using vaseline, allowing it to be lowered during the surgery. Tetrodes were then trimmed to ~3 mm using ceramic scissors (Science Tools, Germany) and gold plated (Non-cyanide gold plating solution, Neuralynx) using a custom built tetrode plater (Matsumoto, 2019) until the impedance was between 150 and 200 kΩ.

Surgical procedures

Before surgery, tips of the tetrodes were washed with ethanol and then with sterile saline. General surgical procedures were carried out as described previously (Sürmeli et al., 2015). Inhalation anaesthesia was induced using 5 % isoflurane / 95 % oxygen, and sustained at 1 – 2 % isoflurane / 98-99 % oxygen with a flow rate of 1 L / minute. An oval incision was made to expose the surface of the skull and a RIVETS head-post (Osborne and Dudman, 2014) attached to the skull and held in place with UV curing resin cement (RelyX™ Unicem, 3M). For electrical grounding, two small craniotomies were drilled on the left hemisphere ~3.4 mm lateral, and ~1 mm rostral relative to Bregma and the center of the intraparietal plate respectively. M1 x 4 mm screws (AccuGroup SFE-M1-4-A2) were then implanted just below the surface of the skull and cemented in place.

The microdrive was attached to an Omnetics to Mill-Max adaptor (Axona, HSADPT-NN1), which was then attached to the stereotaxic frame. The tetrodes were lowered 1.2 - 1.4 mm into the brain, beginning at 3.4 mm lateral from Bregma (right hemisphere) and along the lambdoid suture, and at an angle of -16° in order to target retrohippocampal cortices. The outer cannula was then lowered and sterile Vaseline carefully placed around the rim to seal the craniotomy. The implant was then fixed to the skull with UV curing resin. After the cement hardened, the grounding wires were wrapped around the grounding screws and fixed with silver paint (RS components 101-5621). A layer of resin cement was applied to cover the grounding screws and any holes were filled with dental cement (Simplex Rapid). Mice recovered on a heat mat for ~20 minutes, and had unlimited access to Vetergesic jelly (0.5 mg / kg of body weight buprenorphine in raspberry jelly) for 12 hours after surgery. Mice were given a minimum of 2 days postoperative recovery before handling and habituation to the virtual reality set up.

Handling and habituation

Handling and habituation of mice was similar to that described previously (Gerlei et al., 2020; Tennant et al., 2018). Mice were handled twice a day for 5 - 10 minutes for one week following surgery. Animals were then habituated to the virtual reality setup for 10 and 20 minutes respectively over two consecutive days. After each habituation session animals were given access to soy milk in their cages to familiarise them with the reward and were given access to a larger arena for 5 - 10 minutes of free exploration.

Food restriction

To motivate mice to obtain rewards in the location memory task their access to food was restricted so that their body weight was maintained at ~85% of it’s baseline value (Tennant et al., 2018). From 4-5 days before starting training an allotted food amount, determined on a daily basis to maintain the animal's weight at its target value, was made available to the mice. The baseline weight of the animal was calculated from its weight prior to restriction and normalised to the expected daily growth for the animal's age.

Experimental design overview

An overview of the experimental design is shown in Figure 3A. On experimental days two recording sessions were performed; one in virtual reality and one in an open arena. On these days animals were collected from the holding room 30 - 60 minutes before the first recording session. Before recording, animals were handled for 5 - 10 minutes, weighed and placed for 10-20 minutes in a cage filled with objects and a running wheel. Animals were then head fixed in the virtual reality setup and performed the location memory task for approximately 30 minutes. After the recording session, animals were placed back in the object-filled arena for 30 minutes before a second recording session

in an open field arena for 30 minutes before being returned to their home cage. Tetrodes were typically lowered by 50 - 100 μm after each session. At the end of the experiment animals were sacrificed and the tissue processed for anatomical localisation of the recording location (described below). Animals in which the recording location was outside the retrohippocampal areas were excluded from analysis (n = 2).

Location memory task

Mice were trained to obtain rewards at a marked location on a virtual linear track as described previously (Tennant et al., 2018). During recording sessions animals were head fixed in the virtual reality setup for approximately 30 minutes. Mice were head fixed using a Rivets clamp (Ronal Tool Company, Inc)(Osborne and Dudman, 2014) and ran on a custom made cylindrical treadmill fitted with a rotary encoder (Pewatron). A feeding tube placed in front of the animal dispensed soy milk rewards (5 - 10 μl per reward). Virtual tracks were generated using Blender3D (blender.com). The track had length 2 m, with a 60 cm track zone, a 20 cm reward zone, a 60 cm track zone and a 60 cm black box to separate successive trials. The walls and floor of the track were labelled with a repeating pattern (Figure 1A). On trials in which the reward zone was visible it was indicated by distinct vertical green and black bars. At any point on the track the visible distance ahead of the mouse was 40 cm.

Electrophysiological signals from the 16 channel microdrive were acquired using an Intan headstage connected via an SPI cable (Intan Technologies, RHD2000 6-ft (1.8 m) standard SPI interface cable) attached to an Open Ephys acquisition board (Siegle et al., 2017). Signals were filtered using the board's analogue high and low cut filter (2.5 Hz -7603.8 Hz). Behavioural variables including position, trial number and trial type were calculated in Blender3D at 60 Hz as described previously (Tennant et al., 2018). Using Blender's embedded Python interpreter (Python API), custom scripts wrote each variable (60 Hz, 12-bit precision) to a data acquisition (DAQ) microcontroller (Arduino Due) which in turn sent the signals to the OpenEphys acquisition board via HDMI.

Open arena exploration

Two open arenas were used for recording purposes. In the first, the arena consisted of an 0.8 square metre surface area acrylic box. The second open arena, as described in (Gerlei et al., 2020), consisted of a square metre surface area metal box with removable metal walls, and a metal frame (Frame parts from Kanya UK, C01-1, C20-10, A33-12, B49-75, B48-75, A39-31, ALU3).

For motion and head-direction tracking, a camera (Logitech B525, 1280 x 720 pixels Webcam, RS components 795-0876) was attached to the frame of the open arena. To record motion alongside head-direction data, a red and green ball was attached to the side of the headstage respectively. A custom Bonsai script then tracked the polystyrene balls during recording sessions based on their

colour (Lopes et al., 2015). To synchronize the position and electrophysiology data, an LED was attached to the side of the open arena in the field of view of the camera. Pulses were sent to the LED via an Arduino (Arduino Uno) as well as to Open Ephys acquisition board via the I/O board. The pulses had random generated gaps (20 to 60 seconds) in between them so that the two series could be reconciled during subsequent analysis (Lopes et al., 2015).

Electrophysiological signals were recorded using the custom made 16 channel microdrive described above. An Intan recording headstage with an SPI cable (Intan Technologies, RHD2000 6-ft (1.8 m) Ultra Thin SPI interface cable C3216) attached was used to connect the microdrive to a commutator (SPI cable adapter board, Intan Technologies C3430 and 3D printed holder, custom designed by Patrick Spooner). The commutator was connected to an Open Ephys acquisition board (Siegle et al., 2017) that sent the electrophysiology signal to a computer where the signal was filtered between 2.5 Hz -7603.8 Hz and displayed using the OpenEphys GUI.

Post recording anatomical assessment of tetrode location

To determine tetrode locations we used either Cresyl violet staining (n = 5 mice) or microcomputed tomography (microCT) (n = 5 mice) (Figure S2). Two mice that underwent cresyl violet staining had no visible tetrode tracks and are omitted from Figure S2. For these mice tetrode location was inferred by the presence of theta modulation of the local field potential and the spiking activity of recorded neurons (Supplemental Figure 5). For both methods mice were anaesthetised with isoflurane and given a lethal dose of sodium pentobarbital (Euthatal, Meridal Animal Health, UK).

For cresyl violet staining, we made an electrolytic lesion using 2 x 2 ms, 20 µA current pulses per tetrode. The mice were fixed by transcardial perfusion of ice cold phosphate buffered saline (PBS; Invitrogen) followed by 4 % paraformaldehyde (PFA; Sigma Aldrich) in 0.1 M phosphate buffer (PB; Sigma Aldrich). Brains were removed, left in 4 % PFA in 0.1 M PB overnight and then in 30 % sucrose in PBS for two nights. A freezing microtome was used to cut 50 µm sagittal slices, which were mounted on polarised slides, dried overnight and then placed in xylene before rehydration in varying concentrations of ethanol (100%, 90% then 70% for 1-2 minutes each). Once washed, sections were placed in Cresyl violet for 2 mins before additional washing and dehydration (70% then 90% for 1.5 minutes each followed by two 30 second washes in 100%). Sections were then placed in xylene and mounted with Dibutylphthalate Polystyrene Xylene (DPX). Imaging of sections was performed with a slide scanner (Zeiss Axioscan.Z1). Sections were then manually inspected for the lesion site.

For microCT imaging, we modified a protocol used previously for imaging rat brains (Masís et al., 2018). Animals were perfused with a mixture of PFA and glutaraldehyde and then left in the same solution for 2 nights. Brains were then washed in ddH_20 before being incubated at 4° for two weeks in 2 % osmium tetroxide (2% OsO_4). After osmium staining brains were washed in ddH_20, dehydrated in

ethanol and then embedded in resin. Once the resin had cured, we placed the sample in a micro-CT scanner (Skyscan 1172, Bruker, Kontich, Belgium) which generated a 3D image of the sample. Scanning parameters were: source voltage 54 kV, current 185 µA, exposure 885 ms with a 0.5 mm aluminium filter between the X-ray source and the sample (Xie et al., 2020). After acquiring the scan the dataset was reconstructed (CTan software, v1.13.5.1) and viewed with DataViewer (Bruker).

For the Cresyl violet and microCT methods, tetrodes were localised relative to landmarks in version 2 of the Allen Reference Atlas for the mouse brain (https://mouse.brain-map.org/static/atlas). Tetrodes were assigned either to the MEC (n = 5 mice), the pre/parasubiculum (n = 2), or when the location was ambiguous between these areas to retrohippocampal cortex (n = 3)(Supplemental Figure 2). Theta frequency field potentials and theta phase locking of spike firing (Supplemental Figure 5) were consistent with previous reports for these regions (Boccara et al., 2010; Chrobak and Buzsáki, 1998; Colgin et al., 2009; Kornienko et al., 2018). We did not attempt to differentiate between presubiculum and parasubiculum as in the mouse their borders are not well defined without additional markers and there is disagreement between reference atlases when using landmarks as reference points. Location of all tetrodes in the retrohippocampal region, defined as including the MEC, presubiculum and parasubiculum (e.g. see (Chrobak and Buzsáki, 1994)) was corroborated by electrophysiological recording of a theta field potential and theta modulation of the firing of neurons (cf. (Boccara et al., 2010))(see Supplemental Figure 5 and analyses section below). In contrast, neurons in nearby structures such as the postrhinal cortex do not show theta phase locked firing (LaChance et al., 2019). In one animal, initial recording sessions that did not show a clear theta-modulated local field potential or theta phase locked spike firing were omitted. The animal from which we were unable to localise the tetrodes anatomically, but in which field potential recordings and spike firing showed theta modulation, was assigned to the retrohippocampal group.

Quantification and statistical analyses

Analyses were carried out using Python (version 3.5.1) and R version: 3.3.1 (2016-06-21). All code will be made available at https://github.com/MattNolanLab.

Spike sorting

Spikes were isolated from electrophysiological data using an automated pipeline based on MountainSort (v 0.11.5 and dependencies)(Chung et al., 2017) and similar to that described in (Gerlei et al., 2020). Data pre-processing steps converted Open Ephys files to mda format, filtered signals between 600 Hz - 6000 Hz and then performed spatial whitening over all channels. Events were detected from peaks > 3 standard deviations above baseline and separated by at least 0.33 ms from other events on the same channel. A 10-dimensional feature space was generated using the first 10 principal components of the detected waveforms and spike sorting using the ISO SPLIT algorithm performed on the feature space (Chung et al., 2017). Cluster quality was evaluated using metrics for

isolation, noise-overlap, and peak signal to noise ratio (Chung et al., 2017). Units that had a firing rate > 0.5 Hz, isolation > 0.9, noise overlap < 0.05, and peak signal to noise ratio > 1 were accepted for further analysis.

Classification of ramp-like activity

To classify ramping neurons (Figure 1, Supplemental Figure 7) we generated firing rate maps for each trial by binning the track into 1 cm bins, summing the number of spikes in each bin and dividing by the time the animal spent there. Bins in which the animal's speed was < 3 cm/s were removed. Firing rates were then smoothed with a Gaussian filter (2 cm width, 2 SD, SciPy Python package (Virtanen et al., 2020)) and averaged across trials. For each neuron, we also generated 999 shuffled datasets in which average firing rates were randomly allocated to track bins (Supplemental Figure 7). We then fit linear models to firing rate as a function of position across the regions of the track before (0-60 cm) or after the reward zone (80-140 cm). This was done for each of the 999 shuffled firing rates and the measured dataset for each neuron. The significance (p) values obtained from fitting the linear models were corrected for multiple comparisons using the Benjamini & Hochberg method (Benjamini and Hochberg, 1995). We classified neurons as having ramp-like activity with positive (or negative) slopes if the linear model was significant (corrected $p < 0.01$), the coefficients of the model were outside the 5 - 95 % intervals of the shuffled data, and the slope was positive (or negative). Neurons not meeting these criteria were allocated to the unclassified group. Classification yielded six possible groups with ramp like activity: positive before the reward zone and after the reward zone (+ +), positive before the reward zone and negative after (+ -), positive before the reward zone and unclassified after (+ un), negative before the reward zone and negative after (- -), negative before the reward zone and positive after (- +) or negative before the reward zone and unclassified after (- +).

Ramp score calculation

As a complementary method to identify ramp-like firing we calculated a ramp score based on the correlation between firing rate and the animal's position. This ramp score is related to the score used previously for identification of speed cells (Kropff et al., 2015). For each cell we calculated separate ramp scores for the track segments before and after the reward zone.

We first smoothed the mean firing rate of each cell in each binned position using a Savitzky-Golay filter (window size: 7, polynomial order: 1) and then identified locations before the reward zone at which peaks and troughs in the firing rate were present by the change of the sign of the slope. Using all turning points (peaks and troughs) in the smoothed curve, together with two endpoints (i.e. beginning and end) of the track segment before the reward zone, we construct a set of potential track segments. For example, if there is a single peak at 30 cm, with two endpoints of the zone at 0 cm and 60 cm, then the set of segments will be (0, 30), (30, 60), (0, 60), where each segment is defined by its starting and ending locations.

To avoid bias from location bins with many data points (e.g. when the animal slowed down in the reward zone), we oversample the original unsmoothed firing rate so that all segments have the same number of points. Then in each of the track segments d_i, using the balanced firing rate fr, we calculate the Pearson correlation coefficient r_i between the firing rate and the location loc.

$$r_i = corrcoeff_{d_i}(fr, loc)$$

To calculate the ramp score for each segment we first determine the ratio of the length of the track segment to the whole track,

$$C_i = \frac{d_i}{D}$$

where d_i is the length of the i-th track segment, *D* is the length of the track region of interest (e.g. the regions before or after the reward zone). The ramp score S_i of this segment is then the harmonic mean of the correlation coefficient and the segment ratio:

$$S_i = \frac{2 \cdot |r_i| \cdot C_i}{|r_i| + C_i}$$

A harmonic mean is used to ensure that a high ramp score will mean that the cell has a high correlation coefficient and a high segment ratio.

To determine the ramp score for each cell, the ramp scores of all possible segments were calculated and the final ramp score S_R of the cell is chosen as the ramp score that has the largest magnitude. The sign of the ramp score indicates whether the slope is positive or negative.

$$S_R = sign(r_{s_{max}}) \cdot max_i(S_i)$$

where $r_{s_{max}}$ is the correlation coefficient of the track segment corresponding to the maximum ramp score.

Mixed effect models for position, speed and acceleration

To quantify the influence of position, speed and acceleration on the firing rate of each neuron, we built mixed models to predict firing rate using position, speed and acceleration as fixed effects and trial number as a random effect. For each neuron, all variables were binned into 100 ms blocks and bins when the animal's speed was < 3 cm/s were removed (Góis and Tort, 2018; Tsao et al., 2018). Firing rate and speed was then smoothed with a Gaussian filter (2 cm width, 2 std, SciPy Python package (Virtanen et al., 2020)). The average speed and acceleration of the animal across the whole recording was calculated, and bins in which speed or acceleration was 3 standard deviations (SDs) above the average for that variable were removed. A model with the configuration 'Rate ~ Position + Speed + Acceleration + (1 + Position | Trials) + error' was fit using the lme4 package (version 1.1-12) in R (Bates et al., 2015). The ANOVA function provided by the car package (version 3.0-9) was used to

calculate significance values for each model coefficient (Fox and Weisberg, 2018). To classify neurons as being modulated by position (P), speed (S), acceleration (A), or a combination thereof, we used a significance threshold of 0.01. For example, a neuron with $p < 0.01$ for position and acceleration and ≥ 0.01 for speed coefficients was classified as a PA neuron.

Mixed effect models for evaluating trial outcome

To evaluate possible effects of trial outcome, we used mixed effect models that included position and trial outcome as fixed effects and trial number as a random effect. Possible outcomes were: 'hit' trials in which animals stopped in the reward zone; 'try' trials if the animal did not stop in the reward zone but its speed in this region was within the 5-95% confidence intervals of hit trials; and 'miss' trials when the animal did not stop in the reward zone and its speed was outside the 5-95% confidence intervals for hit trials.

Similar to analyses of position, all spikes that occurred when the animal's speed was < 3 cm/s and bins in which speed or acceleration were > 3 standard deviations (SDs) above the mean removed. A model with the form 'Rates ~ Position * Trial_outcome + (1 + Position | Trials) + error' was fit for each neuron using the lme4 package in R (Bates et al., 2015). Significance values for each model coefficient were then calculated using the ANOVA function provided by the car package (Fox and Weisberg, 2018). We used a significance threshold of 0.01 for the interaction between position and trial outcome to identify neurons with ramp-like activity that depended on the trial outcome. Sessions in which mice had less than 6 trials classified as either run or try were not included in the analyses.

Average speed calculation

The average speed as a function of location on the track for each session was calculated by binning speed into 1 cm location bins for each trial and then smoothing with a Gaussian filter (2 cm width, 3 std, SciPy Python package (Virtanen et al., 2020)). Speed for each location bin was then averaged over trials. The mean speed across all sessions was calculated for each animal for each trial outcome. The mean and standard error across all animals was calculated from the means for each animal.

Calculation of spiking theta index

Theta index was calculated following methods in (Kornienko et al., 2018). Spike times were transformed into instantaneous firing rates. A power spectrum was then estimated using Welch's method. The theta index was defined as

$$T = \frac{\theta - baseline}{\theta + baseline}$$

where θ is the mean theta band power (6-10 Hz) and baseline is the mean power of two adjacent frequency bands (3-5 and 11-13 Hz).

Power spectral analyses of local field potentials

We computed the power spectra across recording sessions for each channel for all raw voltage samples where the mouse was moving (> 3 cm/s). Power was estimated using Welch's method using a 2 second (60000 samples at 30Hz sample rate) Hanning window and averaged across channels.

Spike half width calculation

Spike half width (the width of the spike at half-maximal amplitude) for each neuron was calculated as described in (Clark and Bracci, 2018). Spikes consisted of a positive-negative waveform with the peak aligned to the centre of a 1 ms snippet. For each neuron 50 waveforms were extracted at random and an average waveform generated. Spike amplitude was defined as the difference between the baseline voltage, calculated as the lowest value, and the positive voltage peak, calculated as the highest value. Spike half width was then calculated as the time (ms) the voltage stayed above half the maximal spike amplitude

Classification of putative inhibitory and excitatory cells

Putative interneurons were defined as neurons with mean firing rate higher than 10 Hz, similar to previous studies in entorhinal populations (Ye et al., 2018) (Supplemental 15 D). Excitatory neurons were classified as mean firing rates lower than 10 Hz. Consistent with this, spike half width in cells classified as putative interneurons were typically much shorter than cells classified as excitatory (Supplemental 15 D).

Classification of cell types in the open arena

Neurons were identified across recording sessions in the location memory task and open arena by concatenating the raw voltage channels from each session and then spike sorting. To classify neurons based on their activity in the open arena we followed methods in Diehl et al. (2017) to assign cells into one of five hierarchically organized functional cell types: grid > border > spatial > head direction > non-spatial. Neurons were assigned to the highest cell type in the hierarchy for which their corresponding identity metric was greater than the 99th percentile of the same scores from 1000 shuffled datasets. For example, a cell with grid score above the 99th percentile was classified as a grid cell irrespective of scores in other categories. To generate shuffled spike data, we took each spike and drew a value from a uniform distribution between 20-580 seconds and added this to its timestamp. Spike times that exceeded the recording length were wrapped around to the start of the session. Spike locations were recomputed from the shuffled spike times and spatial scores calculated similar to measured data. Using this method we generated 1000 shuffles per cell and used the 99th percentile of the scores as a threshold for belonging to the five cell types.

We used established scores for grid, border, head direction and spatial stability (Diehl et al., 2017; Sargolini et al., 2006; Solstad et al., 2008, Skaggs et al., 1993). Grid scores were defined as the difference between the minimum correlation coefficient for rate map autocorrelogram rotations of 60 and 120 degrees and the maximum correlation coefficient for autocorrelogram rotations of 30, 90 and 150 degrees (Sargolini et al., 2006). The firing rate map was calculated by binning spikes into 2.5 cm bins and dividing by the total time occupied in each bin and then smoothed with a Gaussian kernel. Autocorrelograms were calculated by sliding the rate map over all x and y bins and calculating a correlation score. Fields were detected in this autocorrelogram by converting it into a binary array using 20% of the maximal correlations as a threshold. If the binary array had more than 7 local maxima, a grid score was calculated. Correlations between the rotated autocorrelograms were then calculated using only a ring containing the 6 local maxima closest to the centre of the binary array and excluding the maximum at the centre. The ring was detected based on the average distance of the 6 fields near the centre of the autocorrelogram (middle border = 1.25* average distance, outer border = 0.25 * average distance).

To calculate border scores (Solstad et al., 2008), firing fields were identified from the firing rate map by collecting neighbouring bins with firing rates greater than 30% the maximal firing rate and covering at least 200 cm^2. For each putative field, the mean firing distance d_m was computed as the average distance to the wall for each bin, divided by the mean firing rate of the field. To obtain a value between 0 and 1, d_m was normalised by the shortest distance from the centre of the arena to the wall. The maximal coverage c_m of a field was measured as the maximum proportion of field bins occupying any of the four edges of the firing rate map. The border score was then given by

$$b = \frac{c_m - d_m}{c_m + d_m}$$

To calculate head direction scores (Sargolini et al., 2006), the head direction recorded at times corresponding to firing events were binned into 360 bins between 0 and 2π and was normalised by the duration of time spent occupying each directional bin. This polar histogram was then smoothed with a rolling sum with a window size of 23 degrees. The head direction score is the length of the mean vector of this polar histogram. To obtain this length, the x and y components are computed by first calculating dx and dy in a unit circle (radius = 1) in steps of 1 degree as

$$dx = \frac{cos(angle)}{radius}, \; dy = \frac{sin(angle)}{radius}$$

The head direction score is then given using the Pythagorean theorem by

$$\sqrt{(\Sigma\, polar\, Histogram \cdot dx)^2 + (\Sigma\, polar\, Histogram \cdot dy)^2}$$

For assessing spatial stability, we used two separate metrics, spatial information and the within-session spatial correlation to identify spatial cells (Diehl et al., 2017). We calculated the spatial information per spike as

$$\sum_{i=1}^{N} Pi(\frac{Ri}{R}) \cdot log_2(\frac{Ri}{R})$$

where i indexes a position bin in the firing rate map, N is the number of bins, P_i is the occupancy probability, R_i is the firing rate in the bin, R is the mean firing rate. The within-session spatial correlation is calculated by computing the Pearson correlation between the firing rate map computed from the first half session and the second half session. Bins that were not visited in both halves were excluded from the calculation.

Neural network model

We trained an artificial neural network to use a reinforcement learning algorithm to solve a task similar to that used for the neuronal recordings. In the artificial task, the track consists of 1010 to 1150 locations (depending on track randomisation), each with a corresponding image. The reinforcement learning agent can either choose to take a step forward or to stop. If the agent chooses to move forwards, the image corresponding to the next location in the track is shown to the network. The optimal behaviour is to keep moving forwards to the hidden reward zone, stop once to acquire the reward, and start moving to the end of the track to proceed to the next trial.

We cropped and re-sampled track images used for the track in the recording experiment to 89x89 pixels and used them as the input to the artificial network. Each image frame of the track was first processed by a three-layer convolutional neural network before being fed into a discrete-time recurrent neural network (RNN) (Elman, 1990). A two-unit linear layer was used to read out the stop/go decision from the RNN. The recurrent states were reset after every 10 trials (which make up a training episode). To make the RNN outputs positive and comparable to neural firing rates, we chose the rectified linear unit (ReLU) as the activation function.

To simulate trial-to-trial variability in movement speed observed in the experimental data, the speed at which the agent moves was randomised at the start of each trial. For example, on 'slow' trials, the agent moved forwards 3 locations every time they chose to step forwards, but on 'fast' trials they moved 7 locations. The possible trial speeds were 3, 4, 5, 6, and 7. To simulate impreciseness of motor control, the number of locations moved forwards was also noisy, for example on the fastest trials there is an 80% chance of moving 7 locations forwards, with 15% probability for moving 7 or 9 (± 1) locations forwards, and 5% probability for 10 or 6 (± 2) locations forwards. Each black box zone was randomly padded by 0 to 70 locations. The artificial track was made up of repeated sequences of

images, meaning it is not possible to know the true location on the track given a single image. These factors also decouple time from location so that the agent adopts a path integration strategy.

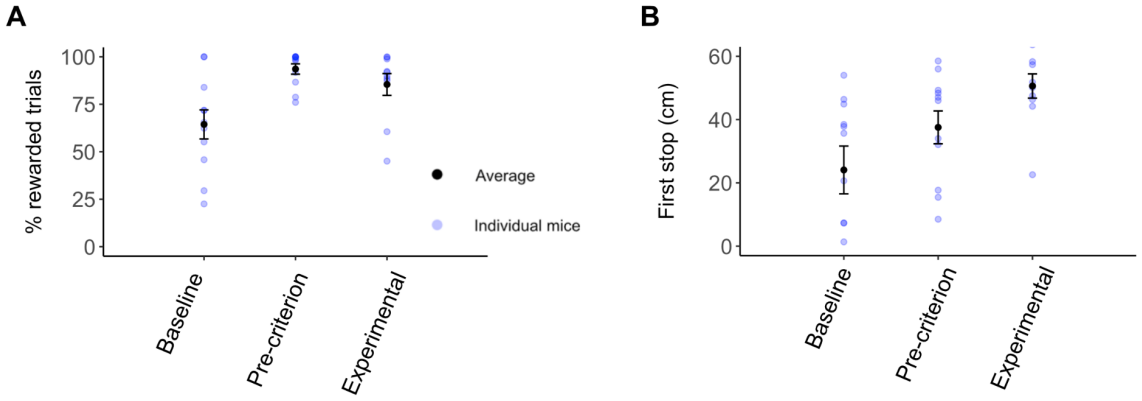
The task given to the agent differed from the task used for experimental recordings in that there was no visual reward zone cue. Training trials were equivalent to 'non-beaconed trials' and included a visual reward indicator in the form of a grey bar that was overlaid at the top of the image input for 5 timesteps if the agent stopped in the reward zone for the first time on a trial. This was designed to be analogous to the immediate feedback the experimental animals receive from a tone and soy milk reward. Probe trials, in which the indicator and reward were absent, were used only for evaluation of the network.

To train the agent to perform the task we used the proximal policy optimisation algorithm (Schulman et al., 2017), which learned an action and value function. During training, stopping in the reward zone for the first time on a given trial resulted in a reward of +100, whereas any other stops resulted in a reward of -1. Moving forwards was a neutral action and had a reward of 0. These rewards were totalled for each trial. The network weights were updated using stochastic-gradient descent to maximise the reward (Kingma and Ba, 2014). To find the optimal hyperparameters for training (such as learning rate), we performed random parameter searches using the Edinburgh Compute and Data Facility (ECDF) (Bergstra and Bengio, 2012), using PyTorch 1.2.0 and cuDNN 10.1.105 (Paszke et al., 2019) and adapted an existing Python implementation of the proximal policy optimisation algorithm (Kostrikov, 2018). The hyperparameters used are in Supplemental Table 1.

Acknowledgements

We thank Richard Morris for feedback on an earlier version of the manuscript, Robert Wallace for assistance with micro-CT imaging, Holly Stevens for technical assistance, and Gülşen Sürmeli, Ian Duguid and members of the Sürmeli, Duguid and Nolan labs for helpful discussions. This work was supported by grants to M.F.N. from the Wellcome Trust (200855/Z/16/Z) and the BBSRC (BB/L010496/1); by grants from the Simons Initiative for the Developing Brain to M.F.N; by a College of Medicine and Veterinary Medicine PhD Studentship, funded by the Thomas Work Fellowship, to K.G., by the Wellcome Trust (108890/Z/15/Z) Translational Neuroscience PhD programme to I.H. and by the Medical Research Council Precision Medicine PhD programme to H.C. This work made use of resources provided by the Edinburgh Compute and Data Facility.

Supplemental figures

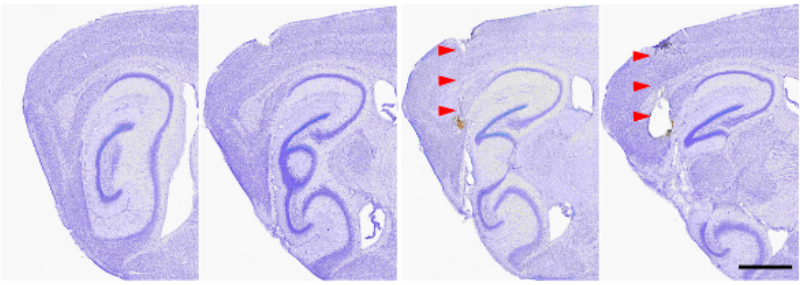


Supplemental Figure 1. Behavioural training

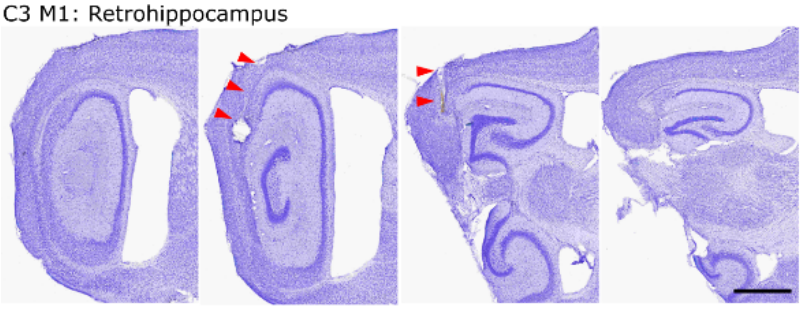
(A) The percentage of rewarded trials for each mouse (blue) and averaged across mice (black, ± SEM) for the first training day (Baseline), the two consecutive days on which mice reached the criterion for introduction of probe trials (Pre-criterion) and the days after the criterion was passed (Experimental). Mice took 13.3 ± 7.1 days to reach the performance criterion. The reward rate depended on training period (Baseline, Pre-criterion or Experimental, $p < 6.72 \times 10^{-7}$, f =15.27; ANOVA). Compared with baseline sessions, the percentage of rewarded trials was larger during pre-criterion (adj $p < 4 \times 10^{-7}$, Tukey's test) and experimental sessions (adj $p < 3 \times 10^{-5}$), whereas pre-criterion and experimental sessions did not differ significantly (adj $p < 6 \times 10^{-2}$).

(B) The mean location at which mice first stopped on each trial for each mouse (blue) and averaged across all mice (black, ± SEM). Categories are as in (A). The first stop location depended on training period (Baseline, Pre-criterion or Experimental, $p < 1.63 \times 10^{-15}$, f = 40.6; ANOVA). Consistent with animals adopting a spatial stopping strategy (Tennant et al., 2018), first stop locations were further along the track during experimental sessions compared with baseline (adj $p < 1 \times 10^{-10}$ Tukey's test) and pre-criterion sessions (adj $p < 1 \times 10^{-10}$ Tukey's test).

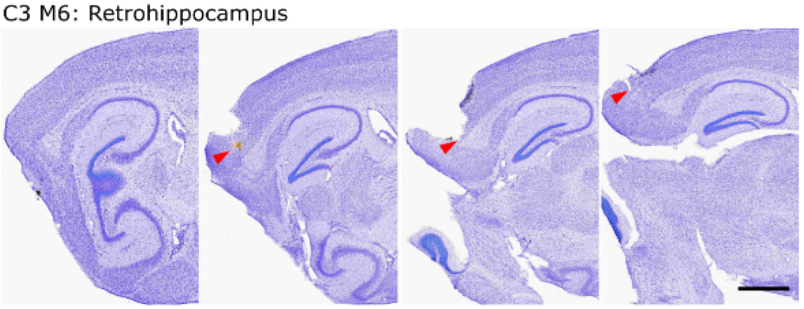
C2 245: Presubiculum



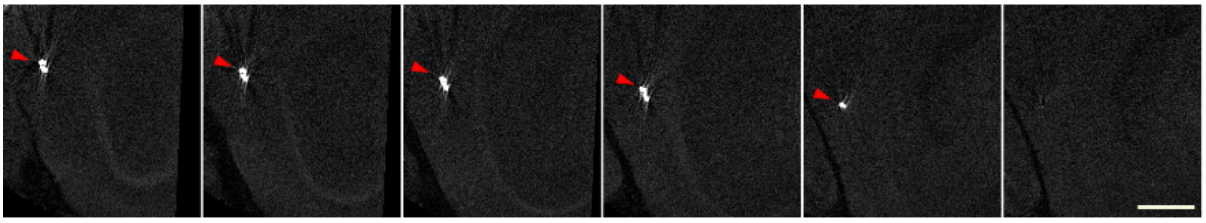
C3 M1: Retrohippocampus



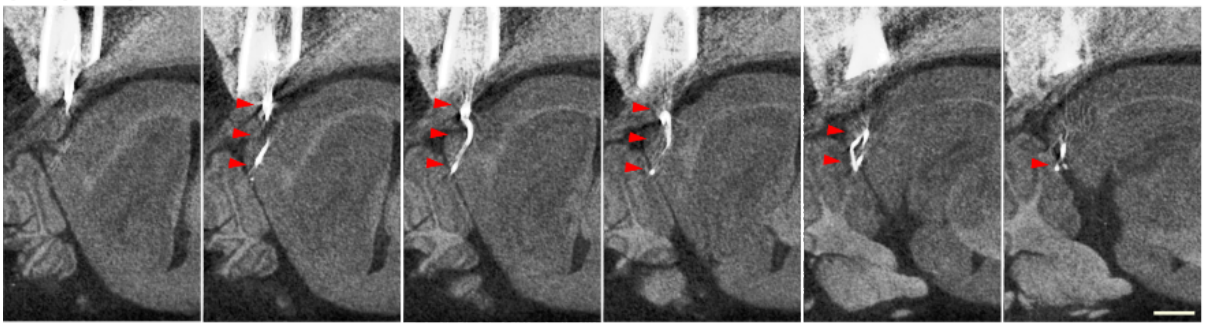
C3 M6: Retrohippocampus

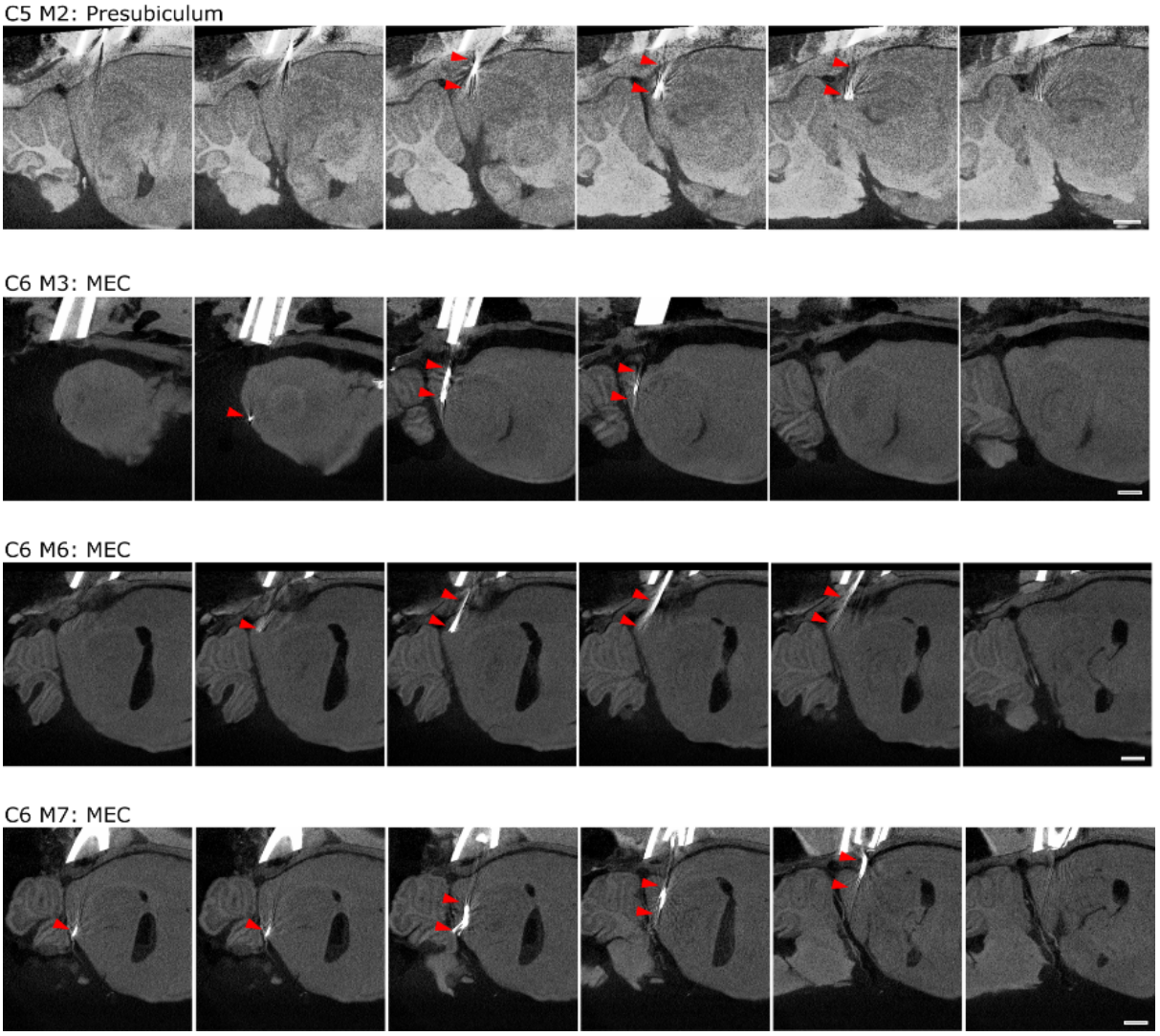


C4 M2: MEC



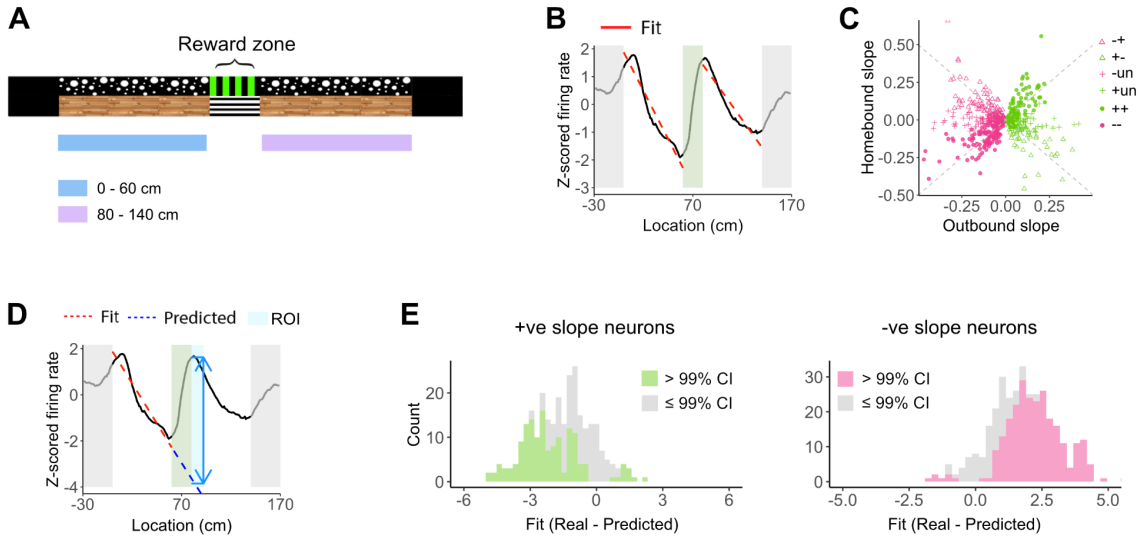
C4 M3: MEC





Supplemental Figure 2. Tetrode localisation

Cresyl violet stained brain sections used for assessment of tetrode locations in three mice and Micro-CT images used for assessment of tetrode locations in six mice. For each animal, sagittal slices are presented lateral to medial from left to right and the classification of the tetrodes target is shown at the top left. This classification is based on the terminal location of the tetrode and the distance travelled during the experiment (see Methods). One mouse had no visible tetrode tracks in any slice and is not shown. In all images red triangles point to the putative tetrode tracks. Scale bar denotes 1mm.



Supplemental Figure 3. Ramp-like activity is modified between the regions of the track before and after the reward zone

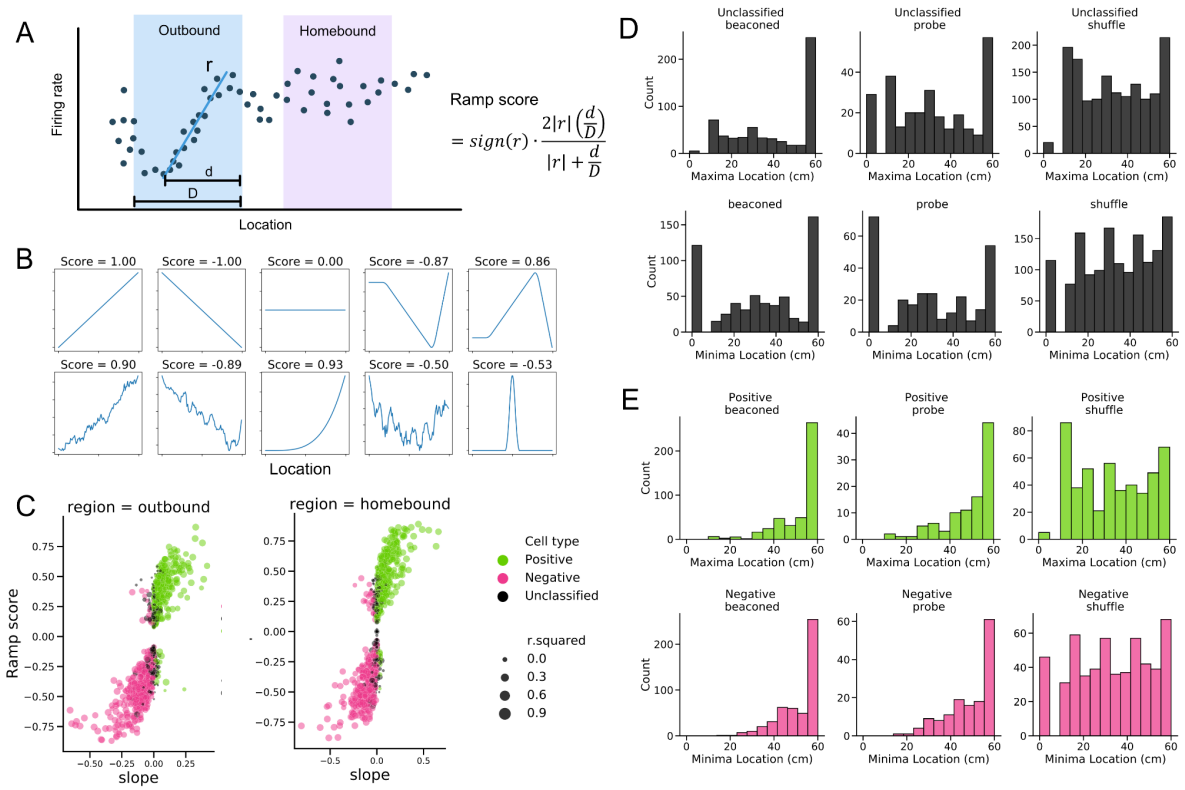
(A) Schematic indicating the regions of the track used for analyses. Coloured regions indicate the segments of the track used for analysis.

(B) Spiking activity as a function of location was fit separately for the track segments before the reward zone (0 - 60 cm) and after the reward zone (80 - 140 cm)..

(C) Slope of the fit to the firing rate on the track segment before the reward zone as a function of slope on the track segment after the reward zone. Dashed lines indicate points for which the slope would be identical on track segments before and after the reward zone. Symbols indicate classification of individual neurons.

(D) Offsets in firing rate were calculated by subtracting the actual firing rate in the first 5 cm after the reward zone with the firing rate predicted by extrapolating the fit from the track segment before the reward zone (see Figure 2).

(E) Distribution of offsets as calculated in (D) for neurons classified on beaconed trials as having ++ (left) or -- (right) activity patterns.



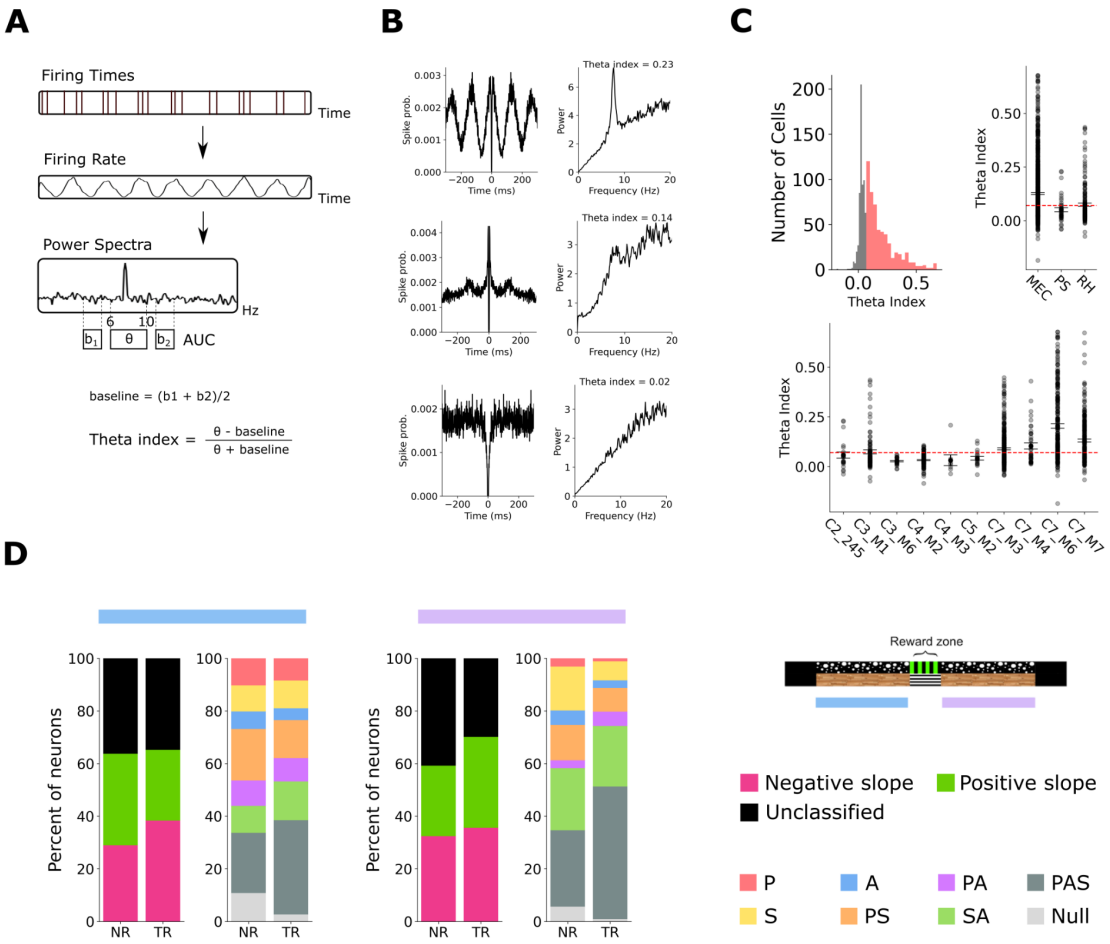
Supplemental Figure 4. Evaluation of ramping activity using 'ramp scores'

(A) Schematics showing the calculation of the ramp score. *r* is the correlation coefficient between the firing rate and the location. *d* is the length of the span of the ramp-like region. *D* is the length of the track region under consideration. Coloured regions indicate the segments of the track used for analysis.

(B) Examples of ramp scores calculated for idealised reference waveforms.

(C) Ramp scores plotted as a function of the slope of linear fits of firing rate to location for all neurons.

(D-E) Locations of maxima and minima for unclassified neurons in the track segment before the reward zone (D) and neurons classified as having positive or negative slopes in the track segment before the reward zone (E).



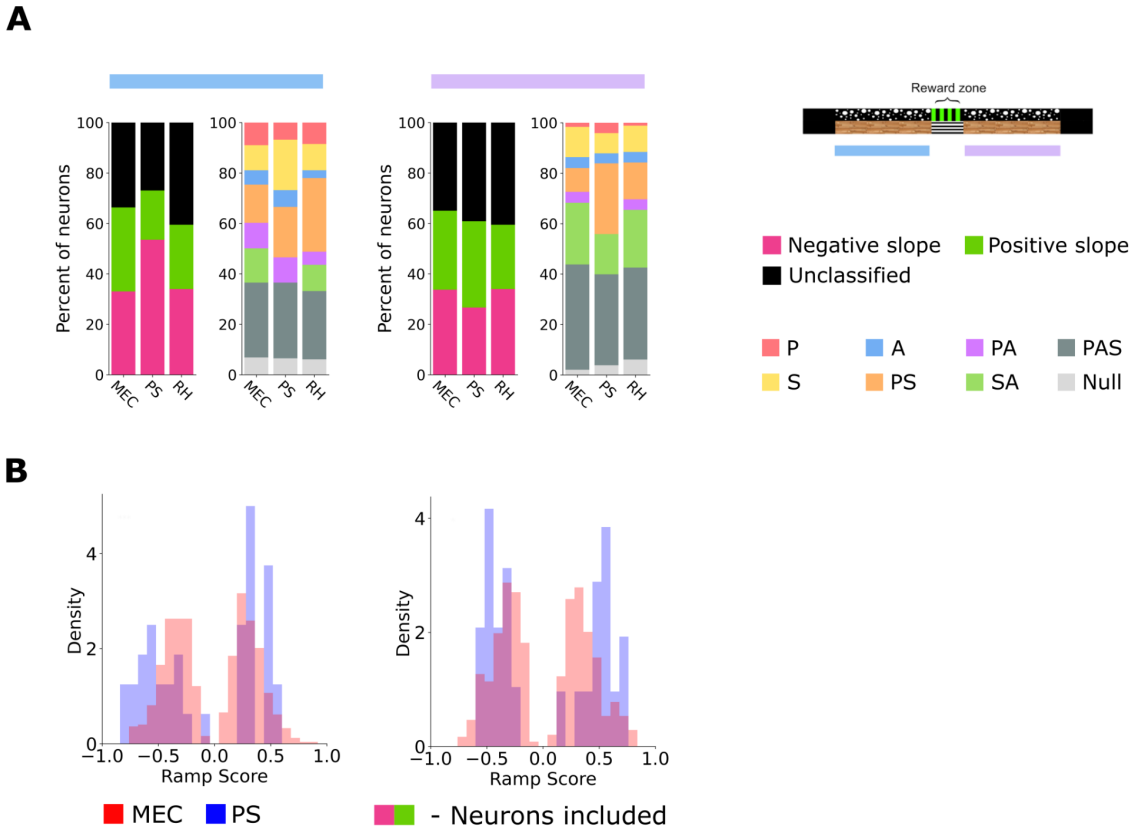
Supplemental Figure 5. Theta modulation of neuronal firing

(A) Schematic illustrating the procedure for calculation of the theta index (see (Kornienko et al., 2018)). Firing indices were transformed into instantaneous firing rates from which a power spectrum was estimated using Welch's method. The theta index was defined as (θ-baseline)/(θ+baseline), where θ is the mean power in the theta band (6-10 Hz) and baseline (b) is the mean power of two adjacent frequency bands 3-5 Hz (b_1) and 11-13 Hz (b_2).

(B) Autocorrelograms (left) and power spectra (right) for three simultaneously recorded neurons that exemplify strong (top), moderate (middle) and low (bottom) theta indices.

(C) Histogram of the theta index across all cells (top left) and scatter plots showing the theta index grouped by tetrode location (top right) and mouse (bottom). Cells were classified as having theta rhythmic firing (red in the histogram) when their theta index was > 0.07 (see (Kornienko et al., 2018)).

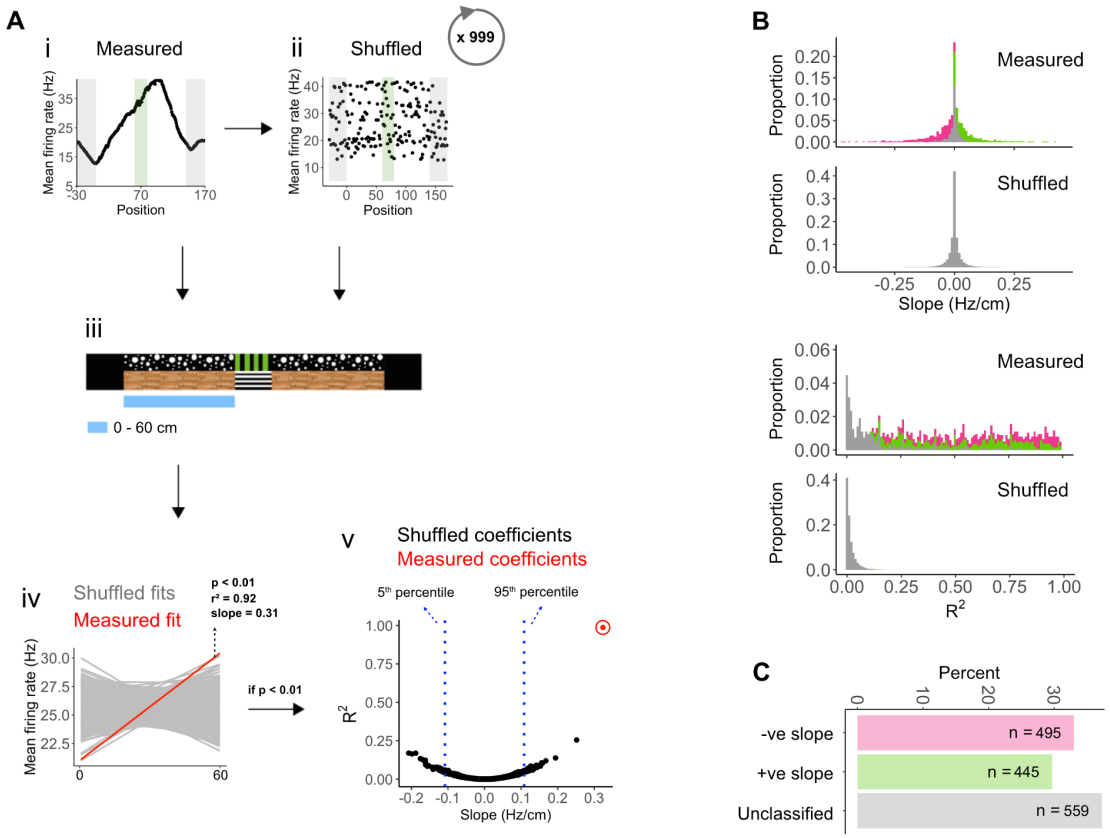
(D) Proportions of non-rhythmic (NR) cells and theta rhythmic (TR) cells classified as having positive or negative slopes (see Supplemental Figure 7) and among these neurons the proportions classified as having firing that correlates with position (P), speed (S) and acceleration (A) (see Supplemental Figure 8).



Supplemental Figure 6. Comparison of ramp scores and slopes in medial entorhinal cortex versus pre/parasubiculum

(A) Proportions of cells classified as having positive or negative slopes in the regions of the track before (left) and after (right) the reward zone (see Supplemental Figure 7) according to their locations in the medial entorhinal cortex (MEC), pre/parasubiculum (PS) or retrohippocampal cortex (RH). RH is used for neurons for which anatomical location in either MEC or PS could not be accurately determined, among these neurons the proportions classified as having firing that correlates with position (P), speed (S) and acceleration (A) (see Supplemental Figure 8).

(B) Distribution of ramp scores for neurons from MEC and pre/parasubiculum that were classified as having positive or negative slopes. Ramp score distributions in MEC and pre/parasubiculum were compared separately for neurons with negative and positive slopes for the track segments before (-ve neurons p=0.0008, +ve neurons p=0.4, Kolmogorov–Smirnov test) and and after the reward zone (-ve neurons p=0.06, +ve neurons p=0.01, Kolmogorov–Smirnov test).

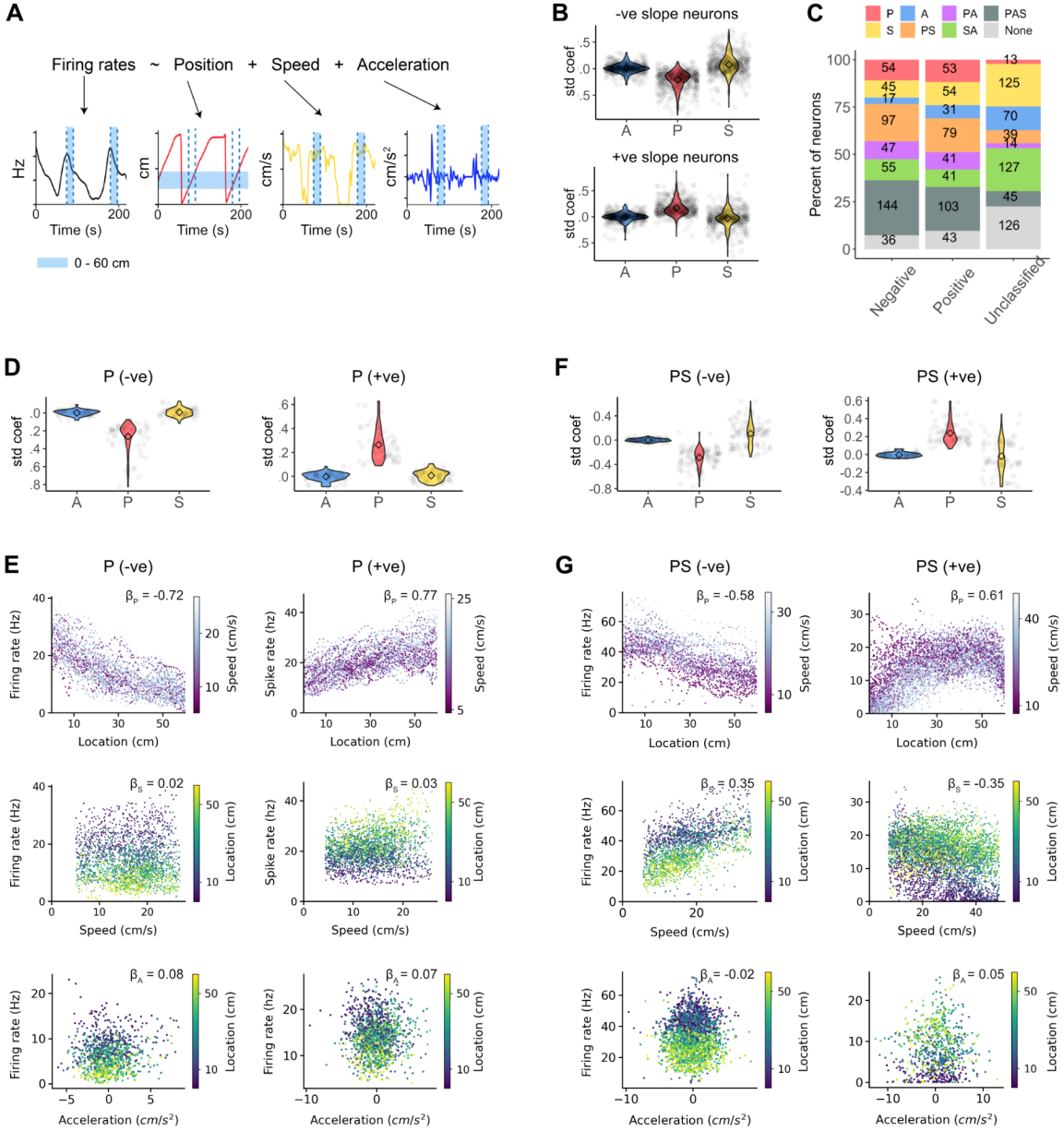


Supplemental Figure 7. **Quantification of ramp-like activity in shuffled data**

(A) Shuffling procedure. Recorded (i) and a single set of shuffled firing rates (ii) plotted as a function of location (iii). Linear fits were calculated for the average recorded firing rate as a function of position on the track segment before the reward zone (0-60 cm, blue bar) and for the corresponding shuffled data (999 shuffled datasets per neuron) (iv). Neurons were classified as ramping if the fit was significant ($p < 0.01$), and as having a positive ramp if the slope of the fit to their recorded data was above the 95th percentile (blue dotted line) of slopes in the shuffled dataset (v) and a negative ramp if the slope was below the 5th percentile.

(B) The distribution of slopes (upper) and R^2 (lower) of the model fit for each neuron. Neurons are colour-coded as having a positive slope (green), negative slope (pink) or unclassified (grey). Distribution of slope and R^2 of the model fit for each shuffled dataset (n=1497501) is plotted in black below the corresponding experimental data.

(C) Proportions of neurons classified in each group.



Supplemental Figure 8. Differential influence of position, speed and acceleration on ramp-like firing

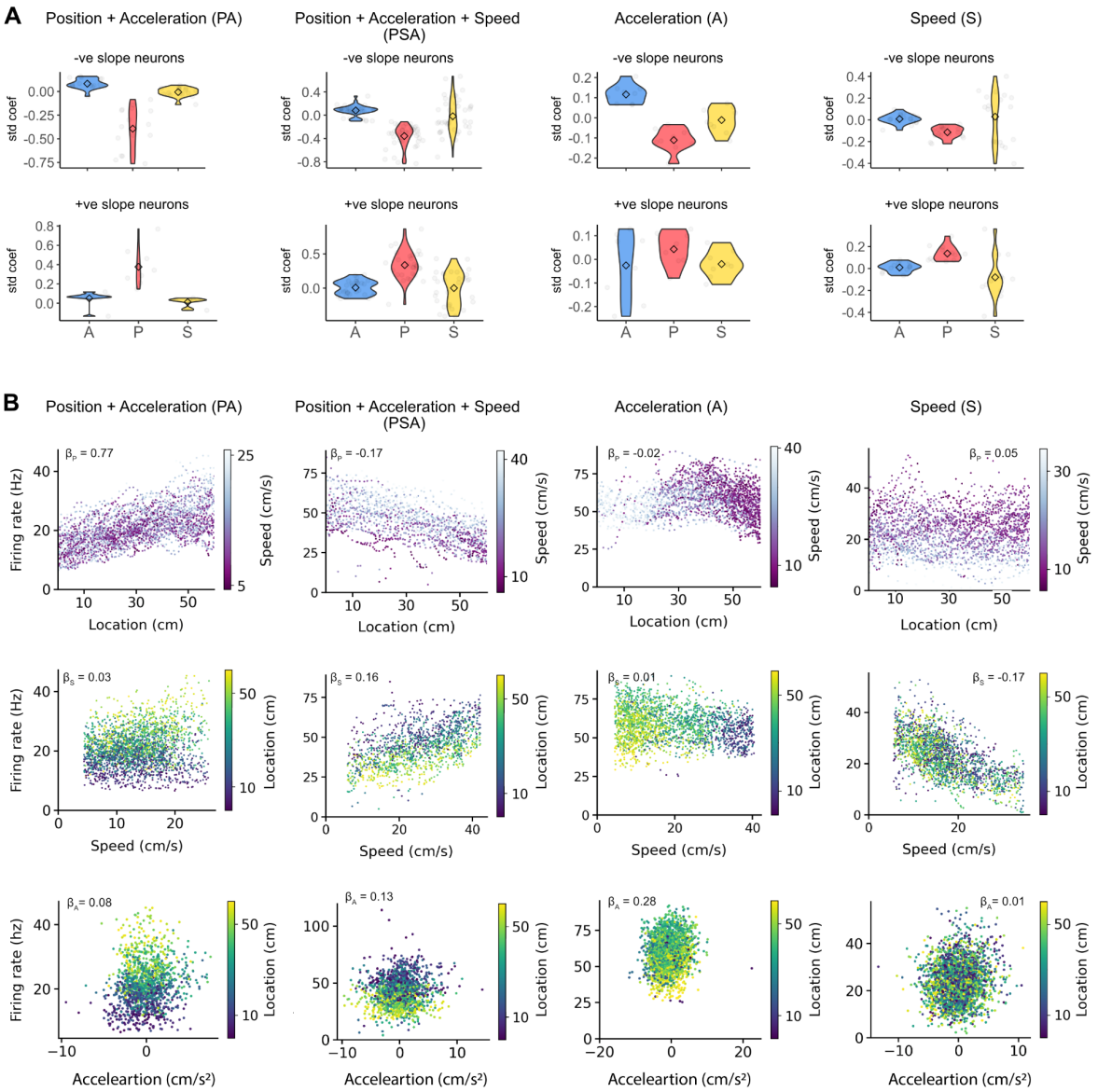
(A) Models that include speed (S), acceleration (A) and position (P) as fixed effects were fit to the firing rate in the track region approaching the reward zone (0-60 cm, see Figure 1C). Trial number was included as a random effect.

(B) Standardised coefficients obtained from the model fits for negative (upper) and positive (lower) slope neurons. Coefficients for speed, acceleration and position differed for neurons with negative and positive slopes (+ve slope: $p < 2 \times 10^{-16}$, f = 219.6133; -ve slope: $p < 2 \times 10^{-16}$, f = 397.4257.5; ANOVA). For neurons in negative and positive slope groups the position coefficients were larger than those for acceleration (+ve slope: adj p = $p < 1 \times 10^{-7}$; -ve slope: adj $p < 1 \times 10^{-7}$; Tukey's test) and speed (+ve slope: adj p = $p < 1 \times 10^{-7}$; -ve slope: adj $p < 1 \times 10^{-7}$ Tukey's test).

(C) Proportions of neurons for which each combination of P, A and S coefficients were significant at a threshold of $p < 0.01$ (Wald's chi squared test) for neurons classified as having a positive slope, negative slope or as unclassified.

(D, F) Distribution of standardised coefficients for neurons classified as having negative slopes (left) or positive slopes (right) and encoding either position alone (D) or a combination of position and speed (F).

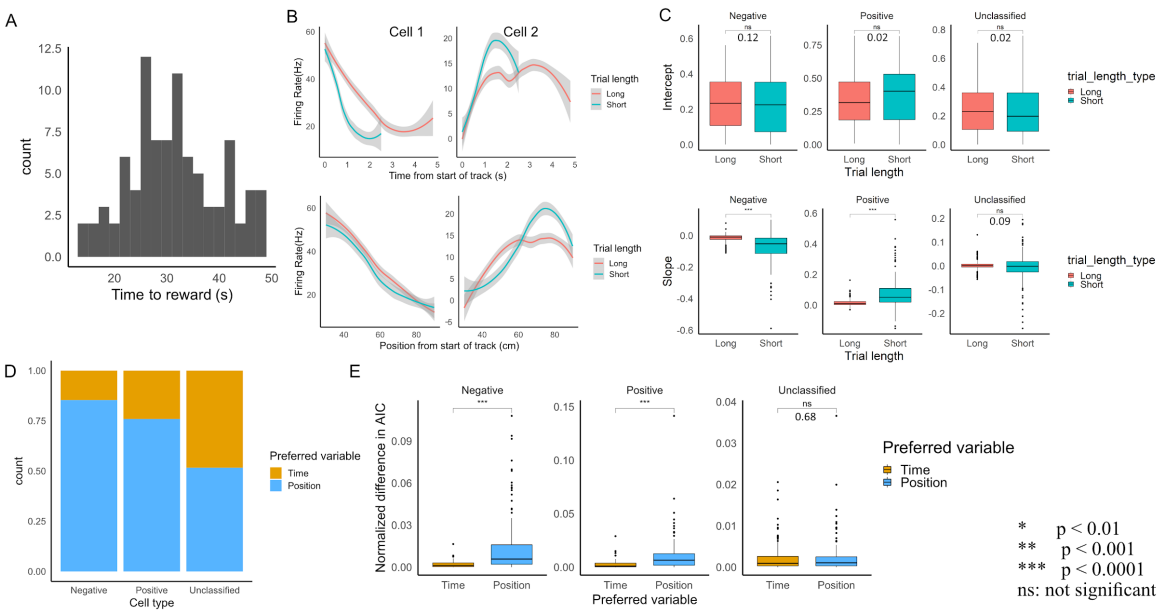
(E, G) Examples of firing rate as a function of position (upper), running speed (middle) and acceleration (lower) for neurons corresponding to the groups in (D and F). Firing rates as a function of position are color coded by speed (cm/s) while firing rates as a function of speed or acceleration are color coded by position along the track (0 - 60 cm).



Supplemental figure 9. Firing rate dynamics of neurons with activity modulated by speed and acceleration

(A) Distribution of standardised coefficients for neurons classified as having negative slopes or positive slopes and encoding position and acceleration (PA), position, speed and acceleration (PSA), acceleration alone (A) or speed alone (S).

(B) Examples of firing rate as a function of position (upper), running speed (middle) and acceleration (lower) for neurons corresponding to the groups in (A). For each neuron, the standardised coefficient is shown in the top right or left hand corner. Fit coefficients for the parameter on the x-axis are shown on each plot. Color coding is as in Supplemental Figure 8E, G.



Supplemental figure 10. Evaluation of time versus position encoding

(A) Distribution of time to reach the reward zone in a representative session.

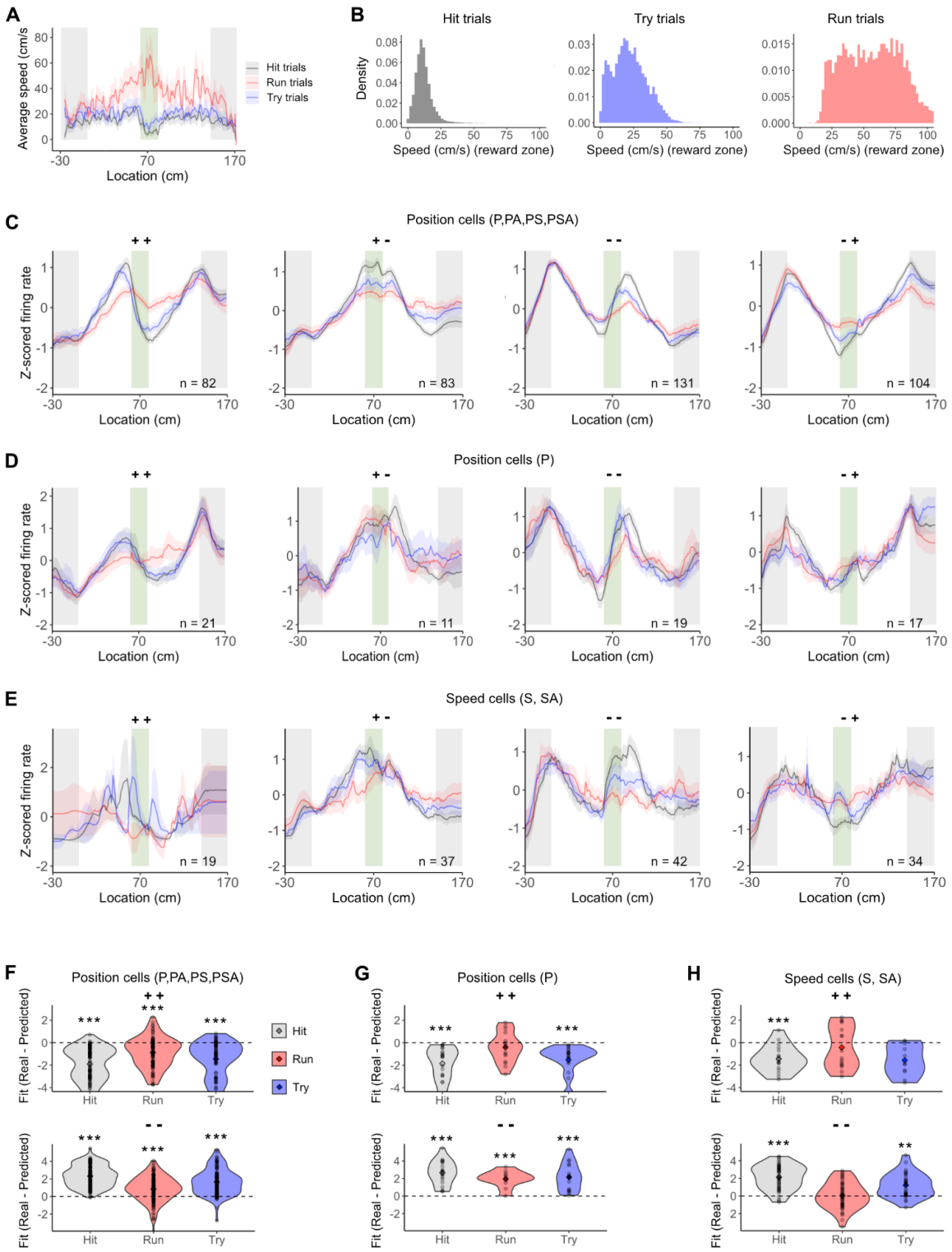
(B) Mean firing rate of two ramping neurons in the track segment approaching the reward zone (track locations 0-60 cm) for short and long trials plotted as a function of time (upper) or position on the track (lower), from the same behavioral session as (A). Short trials are trials with duration less than the 25th percentile of all the trials in a session, while long trials are those with duration greater than the 75th percentile. A similar relationship between firing rate and distance on long and short trials is consistent with position encoding, whereas a similar relationship between firing rate and time is consistent with time encoding.

(C) Comparison of the intercept (upper) and slope (lower) on short and long trials, for cells classified as having positive or negative slopes, when firing rate on the track segment approaching the reward zone is fit as a function of time. Significant differences between the slopes of short versus long trials are inconsistent with neurons encoding time (paired Wilcoxon signed-rank test). Values for significance estimates are shown in the figure. Test statistic top row, left to right: V=13371, V=4383, V=23232, bottom row: V=21704, V=1305, V=22255

(D) Proportion of cells encoding time or position based on comparison of Akaike Information Criterion (AIC) values obtained from mixed effect models fit to firing rate with either time or position as the fixed effect and trial number as the random effect. The model with the more negative AIC value was assumed to provide a better account of the data. In general the activity of neurons classified as having positive or negative slopes in the track segment following the reward zone were better predicted by models in which position rather than time is a fixed effect.

(E) Comparison of the normalized difference in AIC between the position model and time model for different cell types. The difference is normalized by the sum of the AIC of position and time model. A larger difference is indicative of a model providing a better explanation for the data. Whereas position coding models were often strongly favoured, time encoding models were not (independent Wilcoxon

signed-rank test, p-value for significance estimates are shown in the figure. Test statistic left to right: W=1351, W=1024, W=9616).



Supplemental Figure 11. Relationship between task performance and ramp-like activity

(A) Population average of mean running speed as a function of position for each trial outcome across all animals (n=10). Shaded regions correspond to the standard error of the mean across animals. Note that given the reversal of the speed trajectory between hit and try groups (reduced in the reward

zone) compared with the run group (higher in the reward zone), neurons with activity determined by running speed should show a similar reversal.

(B) Distribution of trial level mean running speeds within the reward zone for hit, try and run trial outcomes. The histograms are calculated from all cells (n = 1499) and all trials with a given outcome (hit: n=168890; try: n=73485; run: n=46041).

(C) Population averaged firing rate of position modulated neurons (P, PA, PS, PSA groups) as a function of position for +-, -+, ++ and -- groups on hit, try and run through trials. Shaded regions correspond to the standard error of the mean across cells. Switches in ramp polarity (+- and -+ groups) and resetting of firing (++ and -- groups) are maintained on all trial types.

(D) Population averaged firing rates plotted as in (C) but exclusively for ++ and -- neurons modulated solely by position and not in combination with speed or acceleration (P group). After removing neurons with modulation by speed and acceleration, resetting of activity is similar in the -- group, but in the ++ group appears attenuated on run trials.

(E) Population averaged firing rate as a function of position on hit, try and run through trials for speed modulated neurons (S, SA groups). Data includes all neurons from the S and SA groups irrespective of their slope.

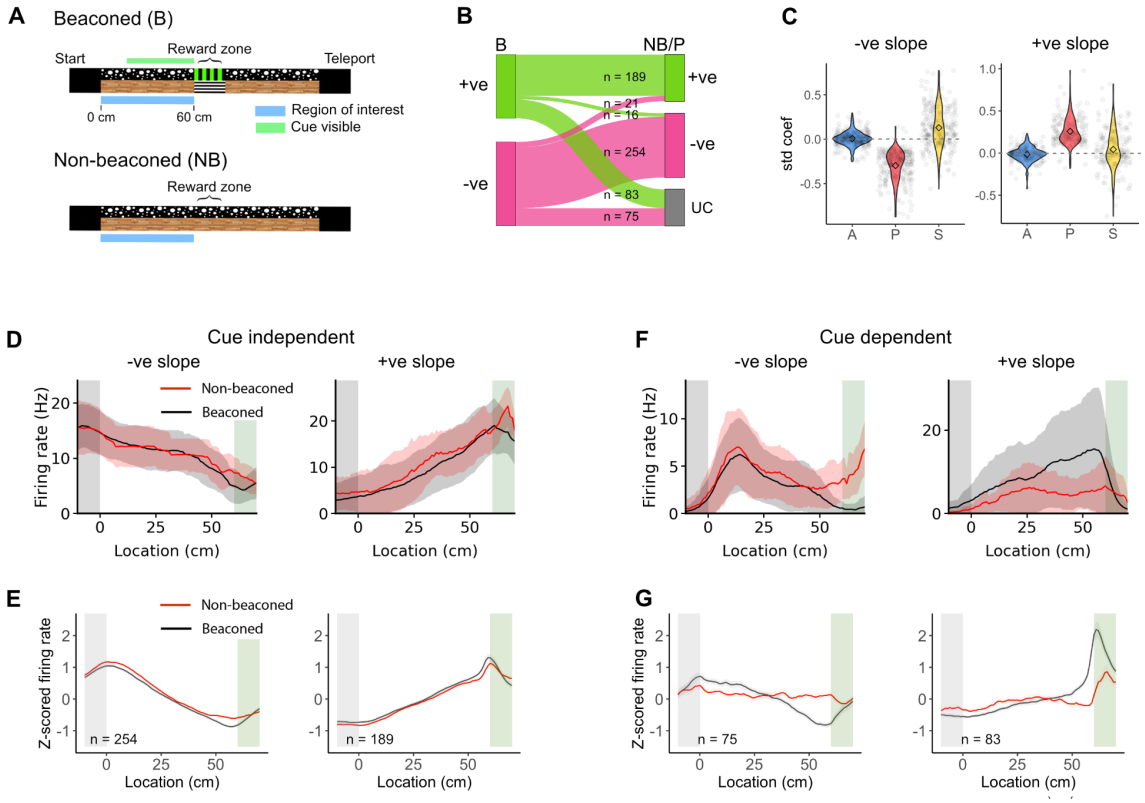
(F-H) Firing rate offsets in the first 5 cm of the track after the reward zone on run, try and hit trials for ++ and -- neurons modulated by position (P, PS, PA, PSA groups)(F), ++ and -- neurons modulated solely by position (P group)(G) and ++ and -- speed modulated neurons (S and SA groups)(H).

For position neurons in general (P, PS, PA, PSA groups), the amplitude of the firing rate offset at the reward zone depended on trial type (+ve slope: $p = 1.64 \times 10^{-5}$, f = 11.53; -ve slope: $p = 2 \times 10^{-16}$, f = 16.7; ANOVA). Nevertheless, offsets for both ++ and -- position neurons (P, PA, PS, PSA) were significantly above zero (++ neurons: hit: t = -13.29, $p = 2.2 \times 10^{-16}$; run: t = -5.75, $p = 1.9 \times 10^{-7}$; try: t = -8.02, $p = 5.09 \times 10^{-11}$; -- neurons: hit: t = 21.81, $p = 2.2 \times 10^{-16}$; run: t = -5.39, $p = 3.97 \times 10^{-7}$; try: t = 10.96, $p = 2.2 \times 10^{-16}$), indicating that position modulation is maintained even though the speed trajectory is reversed.

For position neurons that were not also modulated by speed or acceleration (P group), those with negative slopes (-- group) did not differ in their offsets between hit, run and try trials (ve slope: p = 0.29, f = 1.24; ANOVA), whereas those with positive slopes (++ group) did (+ve slope: p = 0.0065, f = 5.55). For ++ neurons, offsets were significantly above 0 in hit (t = -5.83, $p = 1.1 \times 10^{-5}$, One Sample t-test) and try trials (t = -5, $p = 9.6 \times 10^{-5}$, One Sample t-test) but not run (t = -1.31, p = 0.21, One Sample t-test) trials. For -- neurons, offsets were significantly above 0 across all trial categories (hit: t = 9.38, $p = 2.35 \times 10^{-8}$; run: t = 7, $p = 3.69 \times 10^{-5}$; try: t = 6.17 , $p = 1.33 \times 10^{-5,}$ One Sample t-test). Thus, resetting of firing rate trajectories is maintained on all trial types for pure position neurons with negative slopes before and after the reward zone (-- group), indicating that in these neurons resetting activity does not depend on task success. For neurons with positive slopes (++ group), resetting

activity is maintained on hit and try trials. The apparent loss of resetting activity in this group on run trials may reflect increased trial to trial variation in the location at which resetting occurs.

For neurons modulated by speed (S and SA groups) firing rate offsets also depended on trial type (+ve slope: $p = 0.004$, $f = 5.97$; -ve slope: $p = 9.9 \times 10^{-10}$, $f = 24.83$; ANOVA). Offsets for ++ speed modulated neurons (S, SA groups) were significantly below 0 in hit trials ($t = 7$, $p = 7.47 \times 10^{-6}$, One Sample t-test) but not run ($t = 0.37$, $p = 0.71$, One Sample t-test) or try ($t = -1.76$, $p = 0.12$, One Sample t-test) trials., and for -- speed modulated neurons were above zero in hit and try but not run trials (hit : $t = 11.24$, $p = 4.21 \times 10^{-14}$; run: $t = -1.04$, $p = 0.3$; try : $t = 4.49$, $p = 0.0002$).



Supplemental Figure 12. Differential effects of reward zone cues on preceding ramping activity

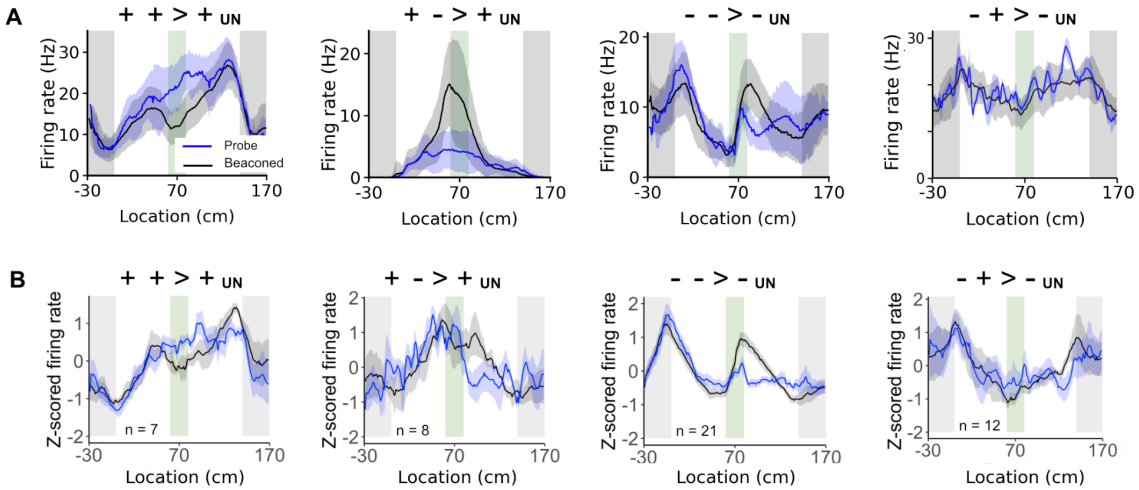
(A) Schematic of the track configuration for beaconed and non-beaconed/probe trials. The green bar indicates the region of the track on which the upcoming reward zone cue is visible.

(B) Classification of the firing rate slope in the track segment approaching the reward zone on beaconed (left) and non-beaconed/probe trials (right).

(C) Standardised coefficients on non-beaconed/probe trials calculated as in Supplemental Figure 8 for neurons that were classified as having negative (left) or positive (right) slopes on beaconed and on non-beaconed/probe trials. (standardized coefficients, ANOVA $p = 2 \times 10^{-16}$, f=52.28; P vs A, adj $p = 5.48 \times 10^{-5}$; P vs S, adj $p = 4.06 \times 10^{-14}$; Tukey's test).

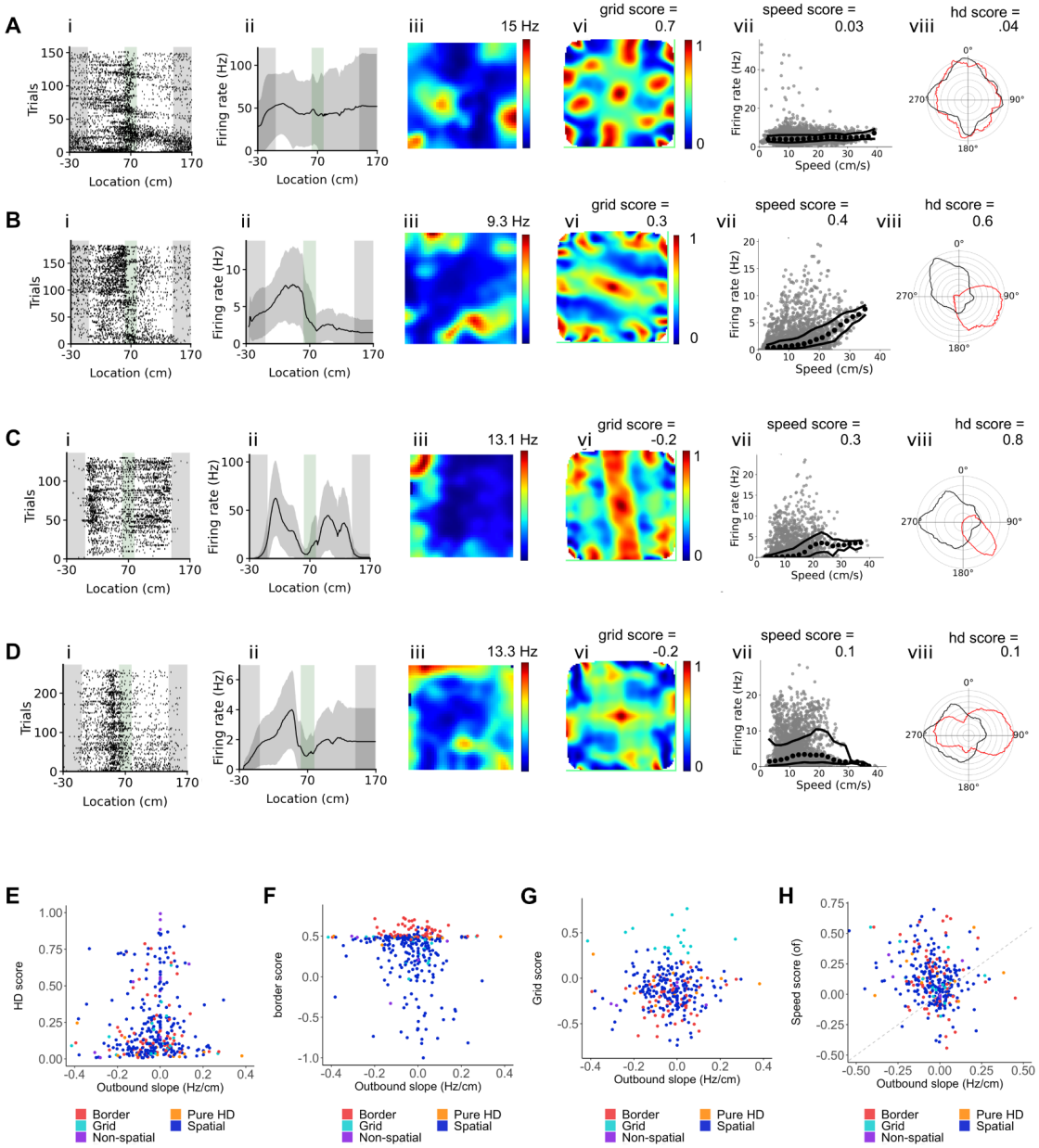
(D, F) Mean firing rate as a function of position for example neurons with classification that was independent from (D) or that was modified by the presence of the reward zone cue (F).

(E, G) Population average firing rate as a function of location for neurons with slope classification that was independent from (E) or that was modified by the presence of the reward zone cue (G), and grouped according to whether the slope on beaconed trials was negative (left) or positive (right). In (D-F) the SEM is indicated in grey.



Supplemental Figure 13. Effects of reward zone cues on ramping activity following the reward zone

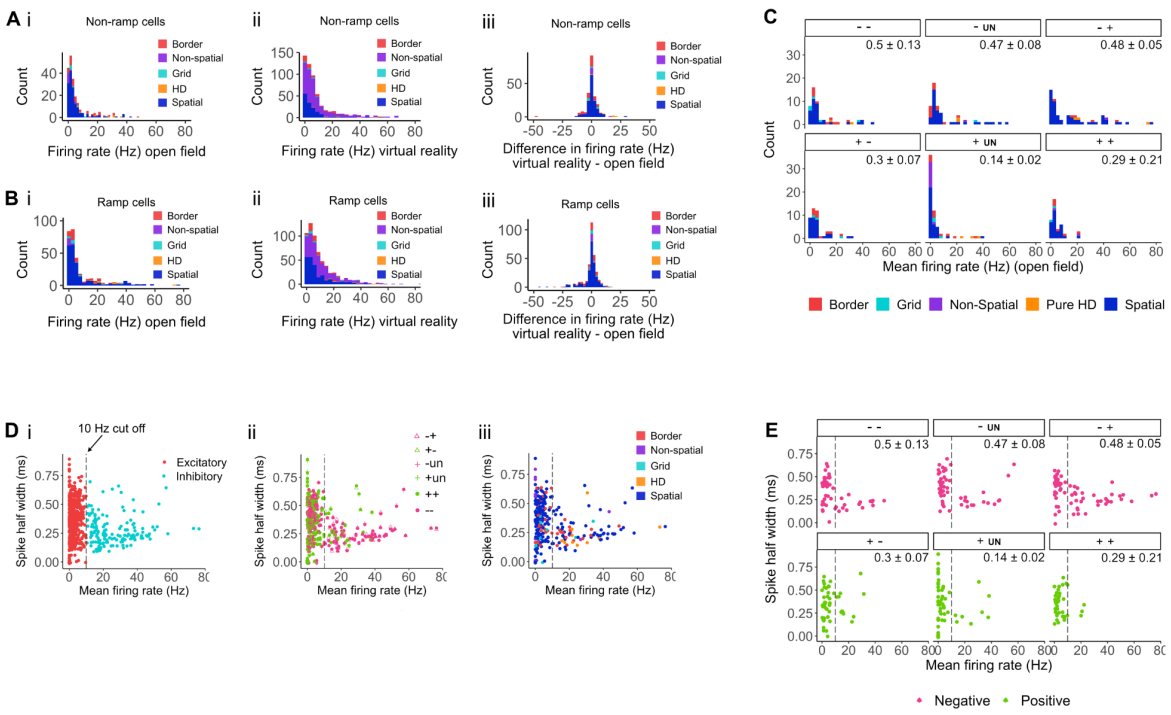
(A-B) Examples (A) and averages (B) of firing rates of ramp cells whose slope on the track segment after the reward zone differs between beaconed and probe trials. Firing rate during beaconed trials is shown in black and probe trials in blue. Shaded regions correspond to the standard error of the mean across cells in B, and across trials in A. Plot titles indicate the classifications on beaconed trials (left) and probe trials (right) i.e. + + > + un indicates a positive ramp both before and after the reward zone on beaconed trials and a positive ramp before the reward zone and unclassified (un) after the reward zone on probe trials.



Supplemental Figure 14. Firing in the location estimation task of neurons classified in the open arena as grid, head direction and border cells

(A-D) Examples of firing along the linear track (i-ii) and in the open arena (iii-iv), firing rate as a function of speed (vii) and polar plots of average firing rate (red) and movement (black) as a function of head direction of neurons classified on the basis of their activity in the open arena (see Methods) as a grid cell (A), a head direction cell (B), a border cell with conjunctive modulation by head direction score (C) and a pure border cell (D).

(E-H) Distribution among position encoding neurons (n = 618) of head direction scores (E), border scores (F). Grid scores (G) and speed scores (H) in the open field against the slope (Hz/cm) of firing rate along the linear track for position encoding ramp neurons.



Supplemental Figure 15. Firing properties of ramp cells in the open arena

(A, B) Stacked histograms showing the distribution of mean firing rates (Hz) of non-ramp cells (A) and ramp cells (as classified in Figure 1) (B) in the open field (i), in the location memory task (ii) and the difference in mean firing rate between the two (open field Hz - location memory task Hz) (iii).

(C) Facet of neurons in (B)(ii) showing histograms of firing rate distributions for each of the ramp categories described in Figure 1. The mean and standard deviation across mice (n = 10) are top right.

(D) Scatter spike half width as a function of firing rate for all position encoding ramp neurons (n = 618), colour coded according to classification of putative interneurons (rate > 10 Hz) and excitatory cells (rate < 10 Hz)(i), ramp trajectory approaching and following the reward zone (ii) and firing in the open arena (iii).

(E) Facet of (D)(ii) showing histograms of firing rate distributions for each of the ramp categories described in Figure 1. Top right corner displays the mean and standard deviation of the firing rate across mice (n = 10).

Supplemental Table 1: Model hyperparameters

Hyperparameters used in the recurrent neural network model

Parameter	Value
Learning rate	1.15e-04
Entropy coefficient	0.01
Epsilon	1e-05
Number of minibatches	4
Number of steps	5000
Number of epochs	14
Weight decay	5e-07
Gamma	0.99
Number of neurons in RNN layer	512

References

Bates, D., Mächler, M., Bolker, B.M., and Walker, S.C. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Softw. *67*, 1–48.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statisitical Society, Series B *57*, 289–300.

Bergstra, J., and Bengio, Y. (2012). Random search for hyper-parameter optimization. J. Mach. Learn. Res. *13*, 281–305.

Boccara, C.N., Sargolini, F., Thoresen, V.H., Solstad, T., Witter, M.P., Moser, E.I., and Moser, M.-B. (2010). Grid cells in pre- and parasubiculum. Nat. Neurosci. *13*, 987–994.

Burgess, N., and O'Keefe, J. (1996). Neuronal computations underlying the firing of place cells and their role in navigation. Hippocampus *6*, 749–762.

Bush, D., Barry, C., Manson, D., and Burgess, N. (2015). Using Grid Cells for Navigation. Neuron *87*, 507–520.

Carelli, R.M., and Deadwyler, S.A. (1994). A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. J. Neurosci. *14*, 7735–7746.

Chen, G., King, J. a., Burgess, N., and O'Keefe, J. (2013). How vision and movement combine in the hippocampal place code. Proc. Natl. Acad. Sci. U. S. A. *110*, 378–383.

Chrobak, J.J., and Buzsáki, G. (1994). Selective activation of deep layer (V-VI) retrohippocampal cortical neurons during hippocampal sharp waves in the behaving rat. J. Neurosci. *14*, 6160–6170.

Chrobak, J.J., and Buzsáki, G. (1998). Gamma oscillations in the entorhinal cortex of the freely behaving rat. J. Neurosci. *18*, 388–398.

Chung, J.E., Magland, J.F., Barnett, A.H., Tolosa, V.M., Tooker, A.C., Lee, K.Y., Shah, K.G., Felix, S.H., Frank, L.M., and Greengard, L.F. (2017). A Fully Automated Approach to Spike Sorting. Neuron *95*, 1381–1394.e6.

Clark, M., and Bracci, E. (2018). Dichotomous Dopaminergic Control of Ventral Pallidum Neurons. Front. Cell. Neurosci. *12*, 260.

Colgin, L.L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., Moser, M.-B., and Moser, E.I. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. Nature *462*, 353–357.

Dhillon, A., and Jones, R.S. (2000). Laminar differences in recurrent excitatory transmission in the rat entorhinal cortex in vitro. Neuroscience *99*, 413–422.

Diehl, G.W., Hon, O.J., Leutgeb, S., and Leutgeb, J.K. (2017). Grid and Nongrid Cells in Medial Entorhinal Cortex Represent Spatial Location and Environmental Features with Complementary Coding Schemes. Neuron *94*, 83–92.e6.

Domnisoru, C., Kinkhabwala, A. a., and Tank, D.W. (2013). Membrane potential dynamics of grid cells. Nature 1–6.

Egorov, A.V., Hamam, B.N., Fransén, E., Hasselmo, M.E., and Alonso, A.A. (2002). Graded persistent activity in entorhinal cortex neurons. Nature *420*, 173–178.

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. Nat. Rev. Neurosci. *1*,

41–50.

Eichenbaum, H. (2017). The role of the hippocampus in navigation is memory. J. Neurophysiol. *117*, 1785–1796.

Elman, J.L. (1990). Finding Structure in Time. Cogn. Sci. *14*, 179–211.

Fiete, I.R., Burak, Y., and Brookings, T. (2008). What grid cells convey about rat location. J. Neurosci. *28*, 6858–6871.

Fiorillo, C.D., Tobler, P.N., and Schultz, W. (2003). Discrete Codingof Reward Probabilityand Uncertaintyby Dopamine Neurons. Science *299*.

Foster, D.J., Morris, R.G., and Dayan, P. (2000). A model of hippocampally dependent navigation, using the temporal difference learning rule. Hippocampus *10*, 1–16.

Fox, J., and Weisberg, S. (2018). An R Companion to Applied Regression (SAGE Publications).

Gerlei, K., Passlack, J., Hawes, I., Vandrey, B., Stevens, H., Papastathopoulos, I., and Nolan, M.F. (2020). Grid cells are modulated by local head direction. Nat. Commun. *11*, 4228.

Gershman, S.J. (2014). Dopamine ramps are a consequence of reward prediction errors. Neural Comput. *26*, 467–471.

Góis, Z.H.T.D., and Tort, A.B.L. (2018). Characterizing Speed Cells in the Rat Hippocampus. Cell Rep. *25*, 1872–1884.e4.

Gustafson, N.J., and Daw, N.D. (2011). Grid Cells, Place Cells, and Geodesic Generalization for Spatial Reinforcement Learning. PLoS Comput. Biol. *7*, e1002235–e1002235.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E.I. (2005a). Microstructure of a spatial map in the entorhinal cortex. Nature *436*, 801–806.

Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., and Moser, E.I. (2005b). Microstructure of a spatial map in the entorhinal cortex. Nature *436*, 801–806.

Hardcastle, K., Maheswaranathan, N., Ganguli, S., Giocomo, L.M., Hardcastle, K., Maheswaranathan, N., Ganguli, S., and Giocomo, L.M. (2017). A Multiplexed , Heterogeneous , and Adaptive Code for Navigation in Medial Entorhinal Cortex. Neuron 1–13.

Harvey, C.D., Collman, F., Dombeck, D. a., and Tank, D.W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. Nature *461*, 941–946.

Howe, M.W., Tierney, P.L., Sandberg, S.G., Phillips, P.E.M., and Graybiel, A.M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature *500*, 575–579.

Insausti, R., Herrero, M.T., and Witter, M.P. (1997). Entorhinal cortex of the rat: cytoarchitectonic subdivisions and the origin and distribution of cortical efferents. Hippocampus *7*, 146–183.

Kingma, D.P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization.

Kornienko, O., Latuske, P., Bassler, M., Kohler, L., and Allen, K. (2018). Non-rhythmic head-direction cells in the parahippocampal region are not constrained by attractor network dynamics. Elife *7*.

Kostrikov, I. (2018). Pytorch implementations of reinforcement learning algorithms.

Kropff, E., Carmichael, J.E., Moser, M.-B., and Moser, E.I. (2015). Speed cells in the medial entorhinal cortex. Nature *523*, 419–424.

Kuipers, B., and Kuipers, B. (1999). The Spatial Semantic Hierarchy {(v.2)}. Artif. Intell. *119*, 191–233.

LaChance, P.A., Todd, T.P., and Taube, J.S. (2019). A sense of space in postrhinal cortex. Science

365, eaax4192.

Lavoie, A.M., and Mizumori, S.J. (1994). Spatial, movement- and reward-sensitive discharge by medial ventral striatum neurons of rats. Brain Res. *638*, 157–168.

Lopes, G., Bonacchi, N., Frazão, J., Neto, J.P., Atallah, B.V., Soares, S., Moreira, L., Matias, S., Itskov, P.M., Correia, P.A., et al. (2015). Bonsai: an event-based framework for processing and controlling data streams. Front. Neuroinform. *9*, 7.

Masís, J., Mankus, D., Wolff, S.B.E., Guitchounts, G., Joesch, M., and Cox, D.D. (2018). A micro-CT-based method for quantitative brain lesion characterization and electrode localization. Sci. Rep. *8*, 5184.

Matsumoto, J. (2019). Matsumotoj/tetroplater: The first release. (http://dx.doi.org/10.5281/zenodo.3334858).

McClelland, J.L., McNaughton, B.L., and O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol. Rev. *102*, 419–457.

McNaughton, B.L., Battaglia, F.P., Jensen, O., Moser, E.I., and Moser, M.-B. (2006). Path integration and the neural basis of the "cognitive map." Nat. Rev. Neurosci. *7*, 663–678.

van der Meer, M.A.A., and Redish, A.D. (2011). Theta phase precession in rat ventral striatum links place and reward information. J. Neurosci. *31*, 2843–2854.

van der Meer, M.A.A., Johnson, A., Schmitzer-Torbert, N.C., and Redish, A.D. (2010). Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. Neuron *67*, 25–32.

Moser, E.I., Moser, M.-B., and McNaughton, B.L. (2017). Spatial representation in the hippocampal formation: a history. Nat. Neurosci. *20*, 1448–1464.

O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. *34*, 171–175.

O'Keefe, J., and Nadel, L. (1978). The hippocampus as a cognitive map (Oxford: Clarendon Press).

O'Reilly, R.C., and Rudy, J.W. (2001). Conjunctive representations in learning and memory: principles of cortical and hippocampal function. Psychol. Rev. *108*, 311–345.

Osborne, J.E., and Dudman, J.T. (2014). RIVETS: a mechanical system for in vivo and in vitro electrophysiology and imaging. PLoS One *9*, e89007.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Advances in Neural Information Processing Systems, H. Wallach, H. Larochelle, A. Beygelzimer, F. d\textquotesingle Alché-Buc, E. Fox, and R. Garnett, eds. (Curran Associates, Inc.), pp. 8026–8037.

Peng, Y., Barreda Tomás, F.J., Klisch, C., Vida, I., and Geiger, J.R.P. (2017). Layer-Specific Organization of Local Excitatory and Inhibitory Synaptic Connectivity in the Rat Presubiculum. Cereb. Cortex *27*, 2435–2452.

Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B.L., Witter, M.P., Moser, M.-B., and Moser, E.I. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. Science *312*, 758–762.

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal Policy Optimization Algorithms.

Schultz, W., Apicella, P., Scarnati, E., and Ljungberg, T. (1992). Neuronal activity in monkey ventral

striatum related to the expectation of reward. J. Neurosci. *12*, 4595–4610.

Siegle, J.H., López, A.C., Patel, Y.A., Abramov, K., Ohayon, S., and Voigts, J. (2017). Open Ephys: an open-source, plugin-based platform for multichannel electrophysiology. J. Neural Eng. *14*, 045003.

Skaggs, W.E., McNaughton, B.L., and Gothard, K.M. (1993). An information-theoretic approach to deciphering the hippocampal code. In Advances in Neural Information Processing Systems, pp. 1030–1037.

Solstad., T., Boccara, C.N., Kropff. E., Moser, M.B., Moser, E.I. (2008). Representation of Geometric Borders in the Entorhinal Cortex. Science, 322, 1865-1868.

Stachenfeld, K.L., Botvinick, M.M., and Gershman, S.J. (2017). The hippocampus as a predictive map. Nat. Neurosci. *20*, 1643–1653.

Stemmler, M., Mathis, A., and Herz, A.V.M. (2015). Connecting multiple spatial scales to decode the population activity of grid cells. Science Advances *1*, e1500816–e1500816.

Sürmeli, G., Marcu, D.C., McClure, C., Garden, D.L.F., Pastoll, H., and Nolan, M.F. (2015). Molecularly Defined Circuitry Reveals Input-Output Segregation in Deep Layers of the Medial Entorhinal Cortex. Neuron *88*, 1040–1053.

Swanson, L.W., and Köhler, C. (1986). Anatomical evidence for direct projections from the entorhinal area to the entire cortical mantle in the rat. J. Neurosci. *6*, 3010–3023.

Tennant, S.A., Fischer, L., Garden, D.L.F., Gerlei, K.Z., Martinez-Gonzalez, C., McClure, C., Wood, E.R., and Nolan, M.F. (2018). Stellate Cells in the Medial Entorhinal Cortex Are Required for Spatial Learning. Cell Rep. *22*, 1313–1324.

Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J.J., Moser, M.-B., and Moser, E.I. (2018). Integrating time from experience in the lateral entorhinal cortex. Nature *561*, 57–62.

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods *17*, 261–272.

Whittington, J.C.R., Muller, T.H., Mark, S., Barry, C., Burgess, N., and Behrens, T.E.J. (2019). The Tolman-Eichenbaum Machine: Unifying space and relational memory through generalisation in the hippocampal formation.

Xie, S., Wallace, R.J., and Pankaj, P. (2020). Time-dependent behaviour of demineralised trabecular bone - Experimental investigation and development of a constitutive model. J. Mech. Behav. Biomed. Mater. *109*, 103751.

Ye, J., Witter, M.P., Moser, M.-B., and Moser, E.I. (2018). Entorhinal fast-spiking speed cells project to the hippocampus. Proc. Natl. Acad. Sci. U. S. A. *115*, E1627–E1636.

Yoshida, M., and Hasselmo, M.E. (2009). Persistent firing supported by an intrinsic cellular mechanism in a component of the head direction system. J. Neurosci. *29*, 4945–4952.