The Primary Function of MTL is Memory, not Navigation: Grid Cells are Non-spatial (*what*) and Place Cells are Memories (*what* and *where*) that Cause Grid Fields through Retrieval

A theory and neurocomputational model are presented that explain grid cell responses as the byproduct of equally dissimilar hippocampal memories. On this account, place and grid cells are not best understood as providing a navigational system. Instead, place cells represent memories that are conjunctions of both spatial and non-spatial attributes, and grid cells primarily represent the non-spatial attributes (e.g., odors, surface texture, etc.) found throughout the two-dimensional recording enclosure. Place cells support memories of the locations where non-spatial attributes can be found (e.g., positions with a particular odor), which are arranged in a hexagonal lattice owing to memory encoding and consolidation processes (pattern separation) as applied to situations in which the non-spatial attributes are found at all locations of a two-dimensional surface. Grid cells exhibit their spatial firing pattern owing to feedback from hippocampal place cells (i.e., a hexagonal pattern of remembered locations for the non-spatial attribute represented by a grid cell). The model explains: 1) grid fields that appear to be centered outside the box; 2) the toroidal nature of grid field representations; 3) grid field alignment with the enclosure borders; 4) modules in which grid cells have the same orientation and spacing but different phases; 5) head direction conjunctive grid cells that become simple head direction cells in the absence of hippocampal feedback; 6) the instant existence of grid fields in a novel environment; 7) the slower learning of place cells; 8) the manner in which head direction sensitivity of place cells changes near borders and in narrow passages; 9) the kinds of changes that underlie remapping of place cells; and 10) grid-like responses for two-dimensional coordinate systems other than navigation.

David E. Huber
University of Colorado, Boulder
Department of Psychology and Neuroscience
e-mail: david.huber@colorado.edu

Introduction

According to the human neuropsychological literature and primate studies, the primary function of the medial temporal lobe (MTL) is to encode and retrieve episodic memories (Corkin, 1984; Squire & Zola-Morgan, 1991). However, when studying rodent behavior, the MTL is often associated with spatial navigation (Moser et al., 2008; O'Keefe, 1976). There are some studies reporting spatial navigation cells in human MTL (Doeller et al., 2010; e.g., Jacobs et al., 2013) and some studies reporting non-spatial cell types in rodent MTL (e.g., Aronov et al., 2017), but these findings are exceptions to the general trends in the literature. It is possible that these general trends are accurate, with the MTL playing different roles in different species, as might be the case if primitive spatial networks were coopted for representing concepts and memories in humans (Buzsáki & Moser, 2013; Milivojevic & Doeller, 2013). Alternatively, these general trends may reflect researchers' use of different behavioral tasks for each species out of convenience – it is difficult to ask rodents to recall events, and it is typically not possible to study humans wandering in large open spaces while recording brain activity. However, if the MTL is fundamentally a memory structure designed to remember *what* happened and *where* it happened, it may be that different behavioral tasks appear to identify different functions for the MTL because they focus on either the *what* or the *where* of the situation, rather than the conjunction of these attributes. Supporting this view, there is ample evidence that spatial location is a powerful episodic retrieval cue (Godden & Baddeley, 1975; Roediger, 1980).

Before presenting this theory in greater detail, I give an example in which a memory conjunction can produce a place cell response, depending on how one analyzes the situation. On the day that I turned 18, I was driving to my parents' home in eastern Massachusetts. Without warning, the front-right ball joint broke and the car tipped onto the road surface. Miraculously, the car safely drifted to a stop in the breakdown lane and I was unharmed. This is a vivid memory of a specific conjunction of attributes, including the exact place on Route 2 where the accident occurred. In the decades since, I have thought of this episode at various times. Of particular relevance to this memory account of hippocampal place cells, I *always* think of this episode when passing that location when traveling along Route 2. If you recorded from one of the hippocampal cells that represents this memory, that cell would appear to be a place cell for a specific location along Route 2. Furthermore, if you put an electrode in one of my cells that represents the concept of car accident, it would systematically respond at that location as well as at other locations of notable accidents (e.g., not one place, but multiple places, similar to a grid cell).

The proposal that hippocampus represents the multimodal conjunctions that define an episode is not new (Marr et al., 1991; Sutherland & Rudy, 1989). This view of the hippocampus is consistent with "feature in place" results (O'Keefe & Krupic, 2021) in which hippocampal cells respond to the conjunction of a non-spatial attribute affixed to a specific location, rather than responding more generically to any instance of a non-spatial attribute. In other words, the what/where conjunction is unique. However, this memory conjunction view of the MTL must be reconciled with the rodent electrophysiology finding that *most* cells in MTL appear to have receptive fields related to some aspect of spatial navigation (Boccara et al., 2010; Grieves & Jeffery, 2017). In brief, if the majority of the cells in the hippocampus and medial entorhinal cortex (mEC) are place cells (O'Keefe, 1976), grid cells (Hafting et al., 2005), head direction cells (Taube et al., 1990), conjunctive grid cells that are also sensitive to head direction (Sargolini et al., 2006), border/boundary cells (Lever et al., 2009; Solstad et al., 2008), and

object-vector cells (Høydal et al., 2019), how can it be that the MTL is primarily a memory structure?

The paucity of non-spatial cells in MTL could be explained if grid cells have been mischaracterized as spatial. On this account, each grid cell represents some non-spatial attribute (e.g., a particular food) and place cells mark where the non-spatial attribute can be found (e.g., memory of where to find that food). Because the non-spatial properties are not manipulated in a typical rodent navigation study, I give them the generic label K, and future studies are needed to determine the bottom-up non-spatial receptive fields that specify the non-spatial attributes K (e.g., see Aronov et al., 2017 for an example in which some grid cells were also found to be sound-frequency cells). On this account, each place cell represents the conjunction of one or more non-spatial attributes and a specific location where those attributes can be found (which might be everywhere within the enclosure), and each grid cells represent the presence of a particular non-spatial attribute.

In naturalistic situations, different attributes (i.e., different K) are positioned in specific places in the environment (e.g., a large tree that serves as a landmark). However, rodent navigation studies often use homogenous environments in which the non-spatial attributes are found everywhere. In this case, any attribute found in the environment (e.g., a particular odor), will be one that is found everywhere in the enclosure. If the non-spatial attribute is found everywhere in the enclosure, the memories of where that attribute can be found will uniformly tile the space of the enclosure. Furthermore, grid cells representing that attribute will be more strongly active in the place field centers of the memories that tiles the space, owing to feedback from place cells – i.e., there is a place cell that is the cause of each separate grid field and the corresponding place cell becomes active owing to memory retrieval of the properties associated with that place.

In a typical rodent grid cell experiment, the *what* of the situation includes many possible factors (i.e., many K) found throughout a two-dimensional navigation surface – e.g., odor, lighting, temperature, surface texture, etc. – as well as many episodic factors that are relatively constant during the recording session – e.g., morning/evening, human experimenter, state of satiety, etc. Given the sameness of the non-spatial factors during the recording session, the only thing that differentiates one memory from another is *where* each memory occurred. This reduces the array of memories to a two-dimensional plane, resulting in a discrete grid of memories, with each place cell capturing a different memory in a different location, and each grid cell modulating its activity as the animal traverses the enclosure, with each location triggering an associated memory.

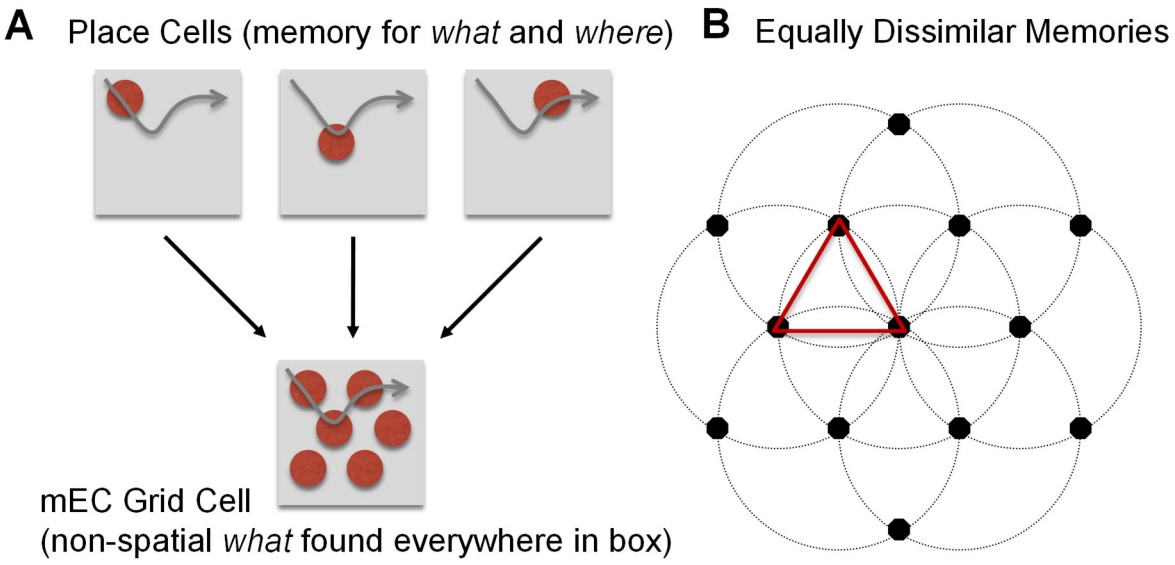

Figure 1. Proposed relationship between grid cells and place cells. Each place cell captures the conjunction of *what* happened and *where* it happened. (**A**) When revisiting a position, the memory associated with that position is retrieved, providing feedback to the non-spatial *what* cells (e.g., the odor of the box), which reside in medial entorhinal cortex (mEC). Because the non-spatial attribute is common to all positions (e.g., it is the same odor everywhere), and because various mechanisms ensure that cells avoid firing constantly at their maximum possible rate during the entire recording session, grid cells fire preferentially in the locations of each memory. (**B**) If memories are formed whenever the current situation is sufficiently different from prior situations, this results in a hexagonally arranged grid of place cells that tile the two-dimensional surface because the only thing that varies in a typical navigation experiment is X/Y location; each memory is formed whenever prior memories are sufficiently far away (the circles represent a fixed dissimilarity between memories).

Figure 1 shows the core hypothesis of this proposal, which assumes that the bottom-up receptive fields of grid cells in medial entorhinal cortex (mEC) are non-spatial attributes found throughout the two-dimensional surface. For instance, all positions in the enclosure might have the same odor. Because the odor is found everywhere in the box, the mEC cell that detects that odor fires preferentially at the positions where the cell also receives feedback from place cells representing associated memories (***Figure 1A***). More specifically, an mEC grid cell experiences constant bottom-up excitatory input combined with top-down memory feedback that is stronger in particular remembered locations. Rather than firing at a constant high rate regardless of location, the cell adopts an adjusted firing threshold owing to inhibitory interneurons and divisive normalization (Bhatia et al., 2019; Carandini & Heeger, 2012; Olsen et al., 2010) such that the cell primarily modulates its activity as a function of top-down memory feedback (because the attribute is found everywhere, the only source of variability in the cell's response is the magnitude of memory feedback). The memories providing this feedback are encoded by hippocampal place cells, which reflect the conjunction of *what* happened (the non-spatial attribute of a particular grid cell) and *where* it happened (one position versus another position in the box). As the animal navigates (gray curved arrow), the current location cues prior memories that occurred in that location, which might be memories that were created just seconds or minutes ago ("here I am again in the middle of the box, which still has that same odor"). Thus,

the finding that the mEC contains a high percentage of grid cells (Boccara et al., 2010) might be an artifact of an experimental paradigm in which nothing of interest varies aside from position.

Memory encoding is more likely to occur in novel situations (Tulving et al., 1996) whereas retrieval occurs in familiar situations (Howard et al., 2005), where "familiar" includes situations that might have occurred just a few seconds ago. Combining these memory principles leads to a regularly spaced array of remembered positions (i.e., equally dissimilar memories) that is created "on-the-fly" when navigating a novel two-dimensional plane that is unvarying in terms of its non-spatial attributes (***Figure 1B***). When first positioned in the enclosure, the animal creates a memory of that position conjoined with the non-spatial attributes of the recording session. This might be in the form of separate memories for each attribute K (e.g., particular odor), or it might be a complex multidimensional memory (e.g., the combination of odor, surface texture, etc.). Once the animal wanders sufficiently far from its initial position, the spatial attributes of the current situation mismatch recently-encoded memories and the animal encodes a new memory. This memory formation process continues until the animal has fully explored the box. When not forming new memories, the animal retrieves recently-encoded memories – in other words, the animal is either creating a memory of the non-spatial attributes associated with a new position or remembering what exists at recently visited positions, with the current position cueing memory retrieval. Because the non-spatial attributes are constant throughout the two-dimensional surface, this results in an array of discrete memory locations that is approximately hexagonal (as explained in the Model Methods, an "online" memory consolidation process employing pattern separation rapidly turns an approximately hexagonal array into one that is precisely hexagonal). Because this array of memories is created on-the-fly, and because there is constant feedback from newly created memories, the hexagonal layout of memories will appear to exist instantly upon entry into the enclosure, as if it were a pre-existing representation designed to aid navigation. In summary, the grid array of memories is rapidly created as the animal explores a novel environment that is devoid of landmarks.

The foregoing account explains the core theory of this proposal. Providing an overview of the findings and model predictions, ***Box 1*** shows a list of results that are explained by this memory model as well as other models, a list of results that are uniquely explained by this model, and a list of predictions made by this model. Items on this list will be addressed in the model methods, simulation results, and discussion.

Box 1

Results Explained by this Memory Model and Other Models
(no single other model explains all of these)

- grid fields can be centered outside the box
- grid is aligned with the walls of the box
- the immediate existence of the grid pattern
- the slower learning of place fields
- dependency of grid cells on hippocampal feedback
- grid cell modules
- grid pattern is disrupted in narrow passages
- partial versus global remapping of place cells
- place cell head direction sensitivity depends on distance to place center
- hexagonal grid patterns for non-spatial representations

Results Uniquely Explained by this Memory Model

- some grid cells become head direction cells w/o hippocampal feedback
- grid cells that are also sensitive to sound frequency
- the population code of grid cells lies on a torus
- place cell head direction sensitivity increases in narrow passages
- the absence of 3D grid cells for bats

Unique Predictions of this Memory Model, yet to be Tested

- sets of place cells (those representing the same properties) are arrayed in a grid
- all grid cells have non-spatial bottom-up receptive fields
- all grid cells are conjunctive with one or more non-spatial property
- all grid cells revert to bottom-up receptive field in the absence of hippocampal feedback
- place cells centered at the same location have complementary head direction sensitivity
- stabilization of place fields depends on learning the borders of the enclosure
- separate regions containing different non-spatial properties produce different grids
- memories rapidly consolidate to become neither too similar nor too dissimilar

Model Methods

In developing a model implementation of the theory outlined in the introduction, auxiliary assumptions (e.g., the shape of neural tuning functions, the manner in which X/Y locations are perceived, the mechanism for memory consolidation, etc.) are needed to address the findings listed in ***Box 1***.

Combinations of Border Cells Specify Position

The model is a bidirectional two-layer network (***Figure 2***) in which border cells (Lever et al., 2009; Solstad et al., 2008) – a.k.a., boundary cells – are the primary spatial input to hippocampus that determines place coding. Three border cell dimensions capture values along three allocentric directions of the two-dimensional surface. These position responses are combined with other entorhinal inputs (head direction and non-spatial grid cells) to form a multi-dimensional egocentric memory of where the non-spatial attributes occurred and which way the animal was oriented at the time of memory formation. When a new memory is formed, a

hippocampal place cell (or more plausibly a population of cells) is recruited and the weights connecting the place cell to the entorhinal inputs are set equal to the current input values (Grossberg & Grossberg, 1982). This assumption is in keeping with the "hippocampal indexing theory" of episodic memory (Teyler & Rudy, 2007), and is a natural outcome of Hebbian learning (Hebb, 2005).

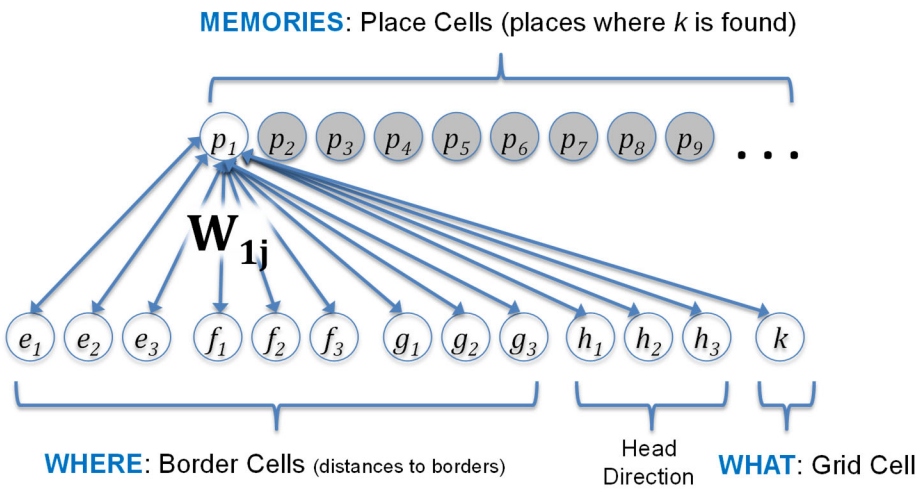

Figure 2. Assumed connectivity between place cells in hippocampus and border cells, head direction cells, and a grid cell in medial entorhinal cortex. When a memory is formed, a place cell (e.g., p_1) is recruited and the weights between the entorhinal inputs and the place cells are set equal to the inputs (e.g., W_{1j} for weights to p_1), with j ranging over the 13 entorhinal inputs to hippocampus. These weights are bidirectional, with feedback supporting memory recall. Because feedback modulates the response of the grid cell, this produces a higher firing rate at positions where the non-spatial attribute is remembered. Code for the model can be found at: https://github.com/dhuber1968/GridCellMemoryModel.

These assumptions regarding border cells are based on the boundary vector cell (BVC) model of Barry et al. (2006). As in the BVC model, combinations of border cells encode where each memory occurred in the real-world X/Y plane. For the model shown in ***Figure 2***, this population code of border cells is additionally combined with the response of head direction cells (Taube et al., 1990) to indicate where the animal was positioned and in which direction the animal was looking/headed (i.e., an egocentric memory). Thus, when including head direction, there is a three-dimensional space in which memories may vary: X, Y, and head direction.

Similar to the BVC model, each border cell receptive field is defined relative to a particular border, with a preferred allocentric distance to that border (e.g., 10 centimeters East of a border). However, unlike the BVC model, not all border cells are assumed to be parallel to a border. Instead, it is assumed that the border cells are pre-arranged into three non-orthogonal dimensions, tipped 60 degrees apart (as explained below, this arrangement provides a basis set that naturally calculates Euclidean distances between current position and remembered positions). When the animal first enters a square enclosure, one of the three dimensions aligns with the most salient wall, with the other two dimensions tipped at an angle relative to the borders.

Unlike the BVC model, the boundary cell representation is sparsely populated using a basis set of three cells for each of the three dimensions (i.e., 9 cells in total), such that for each of the three non-orthogonal orientations, one cell captures one border, another the opposite border, and the third cell captures positions between the opposing borders (Solstad et al., 2008). However, this is not a core assumption, and it is possible to configure the model with border cell configurations that contain two opponent border cells per dimension, without needing to assume that any cells prefer positions between the borders (with the current parameters, the model predicts there will be two border cells for each between-border cell). Similarly, it is possible to configure the model with more than 3 cells for each dimension (i.e., multiple cells representing positions between the borders).

The assumed representational space is allocentric in the sense of aligning with the most salient characteristic of either the interior of the enclosure (e.g., a straight wall), the exterior to the enclosure (e.g., an external cue card), or some combination of the interior and exterior (e.g., an external cue card by virtue of its alignment with a straight wall). As such, the model can explain the finding that border cell representations rotate with rotation of the enclosure (but stay fixed relative to the experimental room) in the case of a circular enclosure that lacks any salient characteristics (Hafting et al., 2005). At the same time, the model can explain the finding that border cell representations do not rotate for a square environment, unless the rotation is 90, 180, or 270 degrees such that a different straight wall aligns with a salient aspect of the exterior (Krupic et al., 2015; Savelli et al., 2017).

Circular Basis Set for Each Dimension

In the model, basis sets of three neural tuning functions (i.e., three cells, each with a neural tuning function centered on a different preferred input) provide an unbiased representation of each dimension (i.e., one basis set for each dimension). An unbiased representation is defined as one that can capture any position along the dimension with equivalent precision. For instance, consider the dimension of color, with the three types of color photoreceptors providing a basis set to identify all visible hues for human color vision. However, color discrimination/precision is not unbiased as a function of hue because the preferred wavelength of the short-wavelength photoreceptors is somewhat separated from the other two preferred wavelengths, which lie closer together (Solomon & Lennie, 2007). In contrast to the biased representation of color, which is constrained by the photochemical properties of photoreceptors, the model assumes that the preferred positions of different border cells (or head direction cells) in the basis set are placed at regularly spaced intervals.

Because memory formation sets the weights equal to entorhinal input responses, the important metric for assessing bias is the sum of squares of the entorhinal input. More specifically, memory retrieval strength is equal to the sum of squares of the input when revisiting a circumstance that precisely matches the circumstance of memory formation (i.e., the input response values will be equal to the weight values in such a circumstance). If this sum of squares of the entorhinal input is not constant with location, then certain pre-defined positions would have the capacity to produce stronger memories (e.g., a larger sum of squares) as compared to other pre-defined positions. Such a situation would bias memory retrieval for some locations as compared to other locations, regardless of prior experience. Thus, beyond equal spacing of preferred stimuli, an unbiased representation also requires a particular shape for the tuning curve

(e.g., how quickly firing rate falls off with dissimilarity from preferring input) to provide a constant sum of squares. This is achieved by assuming a sine wave shape for the tuning curves.

Because head direction is a circular dimension, it was assumed that all dimensions are circular. This assumption was made to keep the model relatively simple. More specifically, the model requires a weight normalization process to ensure that the pattern of weights for each dimension corresponds to a possible input value along that dimension, and the assumption that all dimensions are circular avoids the need to adopt different weight normalization processes for some dimensions as compared to other dimensions, such as would occur if the entorhinal input was a mixture of circular and linear dimensions.

Each basis set that represents a dimension contains three equally spaced sine wave neural tuning functions. ***Equation 1*** is the neural response, r, of neuron, i, at dimension value, d ($-1 < d < +1$). The preferred value, p_i, for the tuning function produces the largest response. The three preferred values for the three simulated neurons in the basis set are set to -2/3, 0, and +2/3, which span the range from -1 to +1. The value of 1 is added to the cosine so that the neural tuning functions are purely positive.

$$r_i(d) = \frac{\sqrt{2}}{3}\{1 + \text{cosine}[\pi(d - p_i)]\} \quad (1)$$

Setting the constant equal to the square root of 2 divided by 3 ensures that the sum of squares of the three sine waves is 1 across all values of the dimension (see ***Figure 3A*** for an example of the three sine waves). Furthermore, the sum of the three sine waves is a constant square root of 2, providing a constant level of input activity across the entorhinal inputs.

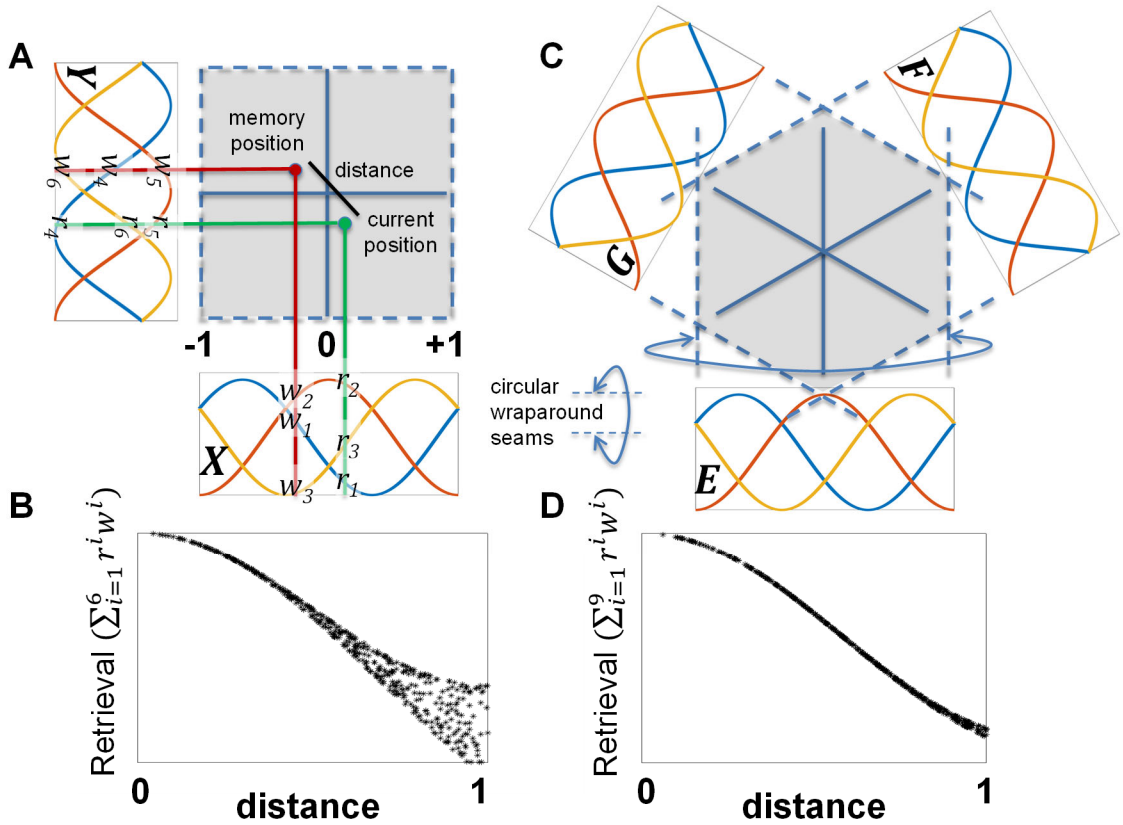

Figure 3. Comparison between memory retrieval with two orthogonal dimensions (*X/Y*) versus three non-orthogonal dimensions (*E/F/G*) that are 60 degrees apart. Each dimension is represented by a circular basis set with three equally spaced sine waves with a period of 2. When a place cell is learned, the weights connecting each sine wave input and the place cells are set equal to the input values. (**A**) In the case of orthogonal *X/Y* dimensions, this results in a pattern of 6 weight values (w_1 – w_6) across the two dimensions, as shown by the intersection of the red lines emanating from the position where the memory was formed (the red dot) and the three sine waves for each dimension. After memory formation, the current position (green dot) reactivates the memory based on the 6 current position response values (r_1 – r_6), summing the multiplication of the response values and the weight values. (**B**) The graph shows the result of randomly sampling 1,000 different memory positions and retrieval positions, plotting retrieval strength as a function of Euclidean distance for each pair of positions. Retrieval strength is variable with Euclidean distance because the sum across the two orthogonal dimensions is a city-block metric (e.g., the same Euclidean distance can map onto multiple city-block distances). (**C**) To capture Euclidean distances, three non-orthogonal dimensions (*E/F/G*) were used. (**D**) This produces a retrieval function that is approximately monotonic with Euclidean distance.

Memory Retrieval Strength is Proportional to Euclidean Distance with Three Non-orthogonal Dimensions

Rather than a two-dimensional torus defined by orthogonal *X/Y* dimensions (***Figure 3A***), the model uses three non-orthogonal *E/F/G* circular dimensions (***Figure 3C***), which define a space that is a hexagonally connected 3-dimensional torus. With three non-orthogonal dimensions, the similarity between memories is approximately Euclidean (***Figure 3D***) whereas with two orthogonal dimensions, the similarity between memories follows a city-block metric. As a result, orthogonal dimensions produce variable memory retrieval strength considering that the same Euclidean distance corresponds to multiple city-block distances (***Figure 3B***).

Memory Encoding and Retrieval Produces an Approximate Grid

Model simulations start the animal at a random location in the enclosure and at each time step, the animal adopts a new random goal direction as compared to the last time step. The simulated animal moves toward the new goal direction, but momentum dictates continuity across time steps. This results in a random curved path that eventually visits all positions with all head directions. An example path with 1,000 timesteps is shown in ***Figure 4C***. A simulated recording session involves 10,000 timesteps. Using the "Grid Score" measure developed by Solstad et al. (2008), even the first 1,000 time steps produce a fairly accurate hexagonal grid of memory positions. The Grid Score measure compares the 2D spatial autocorrelation of the spike rate map to one that is rotated by 60 or 120 degrees (sixfold symmetry for a hexagonal grid) versus one that is rotated 30, 90, or 150 degrees (fourfold symmetry for a square grid). Thus, a grid that is regular but square rather than hexagonal produces a negative Grid Scores. A Grid Score of zero can occur either if the spatial autocorrelation fails to exhibit any systematic grid or if the grid is halfway between square and hexagonal.

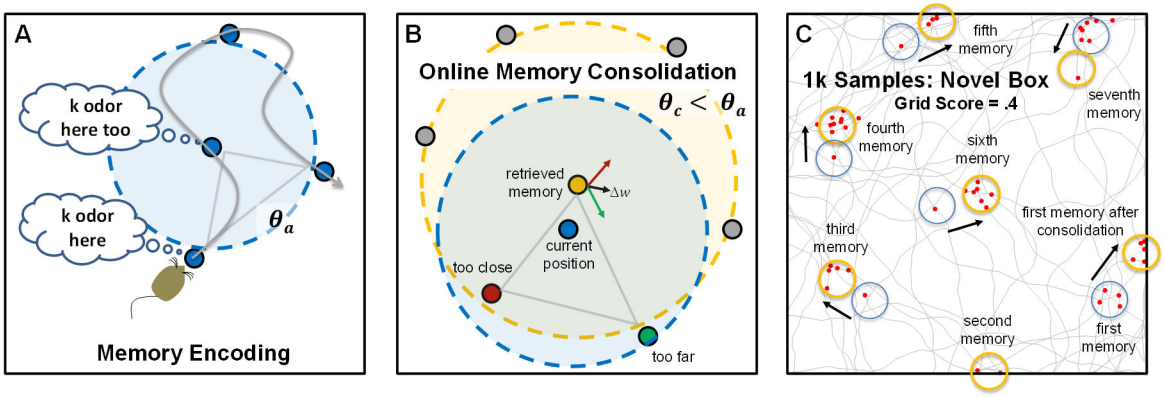

Figure 4. Memory encoding, memory consolidation, and an example sequence. (**A**) When first entering a novel environment, the animal creates a memory of the non-spatial attributes of that environment (e.g., a particular odor) at each location where the attribute is found. The gray curved arrow shows the random path taken by the simulated animal and the blue dots show the positions where memories are created. The activation threshold, θ_a (blue dashed circle), dictates whether previously created memories are retrieved, or if none are retrieved, a new memory is formed. This produces a minimum distance between memories. (**B**) The representations of memories are altered by an online consolidation process that produces unbiased memory representations that tile the environment (i.e., a cognitive map). In this process, the most strongly active memory (the yellow dot retrieved memory) is slightly altered in relation to competing memories that are also activated by the current position (the red and green dots). Other memories (gray dots) remain inactive because they are too dissimilar to the current position (outside the blue dashed circle centered on current position). After initial memory retrieval, the retrieved inputs are used to activate the competing memories and this strength of activation of competitors is compared to a consolidation threshold, θ_c (yellow dashed circle, centered on retrieved memory), which is smaller than the activation threshold, such that consolidation pushes memories to become maximally dissimilar (pattern separation). Competing memories that are more active than the consolidation threshold (red dot) push the weights of the retrieved memory away from the competing memory (red arrow). Competing memories that are less active than the consolidation threshold (green dot) pull the weights of the retrieved memory towards the competing memory (green arrow). For memories arranged in two real-world dimensions, this typically results in activation of three surrounding memories and consolidation makes the triangle formed by these memories an equilateral triangle (gray arrow). (**C**) An example path with 1,000 simulated steps is shown, with the blue circles indicating the initial positions of memories and the yellow circles indicating memory positions after consolidation. The red dots show positions where the simulated grid cell fired. The firing threshold was set such that the cell fires 5% of the time, resulting in 50 positions where the grid cell fired. A movie showing the simulation in **C**, including memories that flash yellow, red, and green as outlined in **B**, can be found at: https://youtu.be/Ts66gBxGdWs.

Equation 2 is the activation, $act(p_i)$, of memory, i, represented by place cell p_i, based on input responses from the 12 entorhinal cells, r_j, that represent the current position ($E/F/G$) and head direction (H), and the previously learned weights, w_{ij}, between each input and the place cell. This equation is divided by 4 because there are 4 dimensions of input ($E/F/G/H$), resulting in a situation where memory activation is 1.0 if all weights exactly equal the input responses (i.e., perfect déjà vu). For simulations that do not include head direction, the summation in the numerator would be over 9 weights (3 dimensions $E/F/G$) and the constant in the denominator would be 3. The activation values for all place cells are then compared to the activation

threshold, θ_a, and if none is above threshold, a new memory is formed by recruiting a new place cell, setting $w_{ij} = r_j$ for all weights connecting inputs to the new place cell.

$$act(p_i) = \frac{\sum_{j=1}^{12} w_{ij} r_j}{4} \quad (2)$$

This mandates a minimum distance between memories at the time of memory formation (***Figure 4A***). In theory, the activation equation should also include the response of the grid cell, k, and the learned weight between grid cell and the place cell. However, because the non-spatial attribute is constant (i.e., the attribute is found everywhere on the surface), its inclusion would only shift all responses by the same constant and can be omitted without changing behavior of the model (i.e., if all activations are increased by a constant, but the activation threshold is also increased by the same constant, then memory formation and retrieval would be unchanged).

Forming a Cognitive Map: Online Memory Consolidation Regularizes the Grid

The encoding assumptions of the model produce a minimum spacing between memories. However, the initial lattice of remembered locations will be somewhat irregular, with some locations more densely surrounded by memories than others as dictated by the path taken by the animal during initial exploration. Similar to the assumption regarding unbiased basis sets of entorhinal inputs, it is advantageous to regularize the remembered locations because doing so provides an unbiased "cognitive map" (O'keefe & Nadel, 1979) that can be used to represent any location with equal precision. This regularization of the memory array could occur offline, as in traditional theories of systems consolidation (O'Reilly & McClelland, 1994). However, the proposed model assumes an online consolidation process that achieves regularization rapidly during initial exploration. Nevertheless, if the proposed consolidation process is a basic mechanism of learning, it will also occur during offline replay (Ólafsdóttir et al., 2018). The goal of the proposed consolidation learning is to ensure that every real-world position is surrounded by an array of memories that are neither too similar nor too dissimilar from each other; if all real-world positions are surrounded by a regular array of remembered locations, this provides a cognitive map by virtue of "triangulation" (i.e., all real-world positions are surrounded by an equilateral triangle of positions with known attributes).

A summary of the consolidation process is provided before considering additional details. First, a set of memories is identified that surround the current real-world situation. Then, the most strongly active of the surrounding memories is adjusted in relation to the other surrounding memories in such a manner as to position the strongly active memory neither too far nor too close to the other surrounding memories. In the case of a real-world space in two-dimensions, the consolidation algorithm can be thought of as ensuring that all positions are surrounded by an equilateral triangle (aka, a "regular simplex"), providing good triangulation to represent knowledge for the attributes of the environment. In the case of three dimensions, which occurs when including head direction, the algorithm ensures that each position is surrounded by a regular tetrahedron of egocentric surrounding memories. More abstractly, this consolidation process can be thought of as providing a "covering map" of surrounding exemplars (Kruschke, 1992). Critically, even positions near the box borders obtain a surrounding array of memories, which occurs naturally during consolidation as some hippocampal place cells are pushed "outside the box". This occurs even though the animal is never given the opportunity to visit

these positions. In other words, false memories for positions outside the box contribute to a cognitive map that supports knowledge for all positions within the box.

Online consolidation quickly repositions memories through nudges to the weights of each previously encoded memory (***Figure 4B***). The consolidation process uses what could be described as top-down delta-rule learning (Rescorla & Wagner, 1972; Widrow & Hoff, 1960): each competing memory that surround a retrieved memory provides a teaching signal to modify the weights of the retrieved memory. Unlike traditional delta-rule learning, which has a fixed learning rate, the learning rate sign and magnitude are modulated according to the BCM rule (Bienenstock et al., 1982), depending on similarity of each competing memory to the retrieved memory. Because the consolidation threshold, θ_c, is smaller than the activation threshold that dictates initial memory formation, memories primarily push away from each other during the earliest stages of exploration, but after further exploration there is an equal mix of pushing and pulling once the enclosure has been fully explored. In other words, during early exploration, there is nothing known about what lies on the other side of each memory, and so there is room to push apart memories. However, once something is known about all positions in the enclosure, the memories are constrained on all sides.

In real systems, the proposed consolidation process might arise from several cycles of alternating bottom-up memory activation and top-down memory recall as coordinated by theta oscillations (see Ólafsdóttir et al., 2016 for evidence that something similar might occur during offline consolidation). First, the entorhinal input corresponding to the real-world situation reminds the animal of the most similar memory (the retrieved memory), as well as a set of partially active competing memories. The retrieved memory re-activates the associated entorhinal inputs owing to top-down memory recall (pattern completion). For instance, a particular X/Y position reminds the animal of some non-spatial attribute that was encountered in a nearby location, and memory retrieval recalls the exact spatial position of that attribute by changing the entorhinal inputs from their real-world values to the recalled values. After this initial top-down memory retrieval, there is a second bottom-up cycle that is now based on the recalled entorhinal inputs. This second bottom-up cycle is used to identify the similarity of the competing memories in relation to the retrieved memory by assessing how strongly the competing memories become active based on the entorhinal inputs of the retrieved memory (critically, the designation of memories as belonging to the set of competing memories is based on the first bottom-up cycle that used real-world inputs such that the consolidation process is in relation to a set of memories that surround the current real-world situation). If the competing memories then re-activate their associated entorhinal inputs (i.e., a second top-down wave, or perhaps through additional cycles if the competing memories are assessed one at a time), this would provide the necessary entorhinal response values for the consolidation learning rule (more specifically, consolidation learning is based on the patterns of entorhinal inputs that correspond to the retrieved memory as well as each of the competing memories).

Equation 3 specifies how the weights, w_{RMj}, connecting the entorhinal input, j, to the Retrieved Memory, RM, are changed by summing over the weight changes in response to each of the, n, Competing Memories, CM. There are typically 3 competing memories for the case of navigating in a space with real-world variation along three-dimensions (i.e., X, Y, and head direction H), corresponding to a surrounding tetrahedron when including the retrieved memory (a surround of 4 memories). ***Equation 4*** specifies the delta-rule for weight changes in response to each competing memory, which multiplies the learning rate for competing memory, α_{cm}, by the difference between the weight connecting the input, j, to the *retrieved* memory and the weight

connecting the input, *j,* to the *competing* memory. Thus, the weight of the retrieved memory becomes more similar to the weight of the competing memory if α_{cm} is negative and more dissimilar if α_{cm} is positive. ***Equation 5*** specifies the sign and magnitude of the BCM learning rate in response to the competing memory. This is calculated by comparing how strongly the inputs corresponding to the retrieved memory (which are the weights of the retrieved memory) activate the competing memory place cell, $act(p_{CM} \mid w_{RM})$, as compared to the consolidation threshold, θ_c. If these retrieved inputs activate the competing memory too strongly ($> \theta_c$), the competing memory is too similar (e.g., too close) and the retrieved memory's weights become dissimilar to the competing memory (pattern separation). If these retrieved inputs activate the competing memory too weakly ($< \theta_c$), the competing memory is too dissimilar (e.g., too far), and the retrieved memory's weights become more similar to the competing memory. If these inputs activate the competing memory to a value equal to the threshold, the competing memory is the desired dissimilarity from retrieved memory and there are no changes to the weights of the retrieved memory.

$$\Delta w_{RMj} = \sum_{CM=1}^{n} \Delta w_{CMj} \tag{3}$$

$$\Delta w_{CMj} = \alpha_{CM}(w_{RMj} - w_{CMj}) \tag{4}$$

$$\alpha_{CM} = act(p_{CM} \mid w_{RM}) - \theta_c \tag{5}$$

Code for model simulations can be found at (https://github.com/dhuber1968/GridCellMemoryModel) and a movie showing the first 1,000 time steps for a simulation that does not include head direction can be found at (https://youtu.be/Ts66gBxGdWs). Pseudocode for each simulated time step is as follows:

1. Move to a new position according to momentum from the prior movement that is partially altered according to a new randomly sampled goal direction
2. Based on the new position and new head direction (***Equation 1***), activate memories (***Equation 2***) and compare to the activation threshold θ_a
3. If no memories are above threshold, create a new memory by recruiting a place cell, setting memory weights equal to the current entorhinal input.
4. If just one memory is above threshold, do nothing (memory retrieval occurred, but no consolidation occurs)
5. If more than one memory is above threshold, note which memory is most active (the *retrieved memory*) and note which other memories are also activated in response to the current inputs (the *competing memories*). Consolidate the retrieved memory in terms of its similarity to the competing memories (***Equations 3-5***), using the consolidation threshold θ_c to determine whether weight changes are towards or away from each competing memory
6. If consolidation occurred, normalize any updated weight values of the retrieved memory to ensure that they correspond to possible real-world inputs

In real systems, the final step of weight normalization might occur through post-learning feedback down to early perceptual regions of the cortex followed by a cycle back up to MTL,

with a comparison between this (feedback) "minus phase" versus (feedforward) "plus phase" (Hinton & Sejnowski, 1986; O'Reilly, 1996). Alternatively, weight normalization might occur through a form of divisive normalization (Carandini & Heeger, 2012). In this case, the normalization ensures that the sum of squared weight values for each basis set is 1.0 and that the sum of weight values for each basis set is the square root of 2, as necessitated by the assumed basis set sine wave tuning functions. Rather than specifying this normalization process, the simulation takes a mathematical shortcut by using the arcsine function to recover the dimensional inputs that are closest to the updated weight values.

Firing Thresholds

The model does not directly implement spiking neurons. Instead, the simulated activation values can be thought of as proportional to the average firing rate of an ensemble of neurons with similar inputs and outputs (O'Reilly & Munakata, 2000). However, in comparing the output of the model to observed spike rate maps, an assumed firing threshold is needed. Consider for instance that there is always top-down activation to the non-spatial attribute cell (*k*) once the enclosure has been fully explored (e.g., all positions trigger memory retrieval indicating that presence of *k*). Nevertheless, a *k*-cell with a sufficiently high firing threshold will become silent when the animal is between preferred memory locations owing to slightly lower memory retrieval strength. The specific firing threshold for each cell is assumed to dynamically adjust to keep each cell at its preferred on-average firing rate over a relatively long timeframe. For the reported simulations of egocentric memories (i.e., ones that include head direction), cells that are active 5% of the time were analyzed. In other words, the simulation unfolds as dictated by the model equations, with continuous activation values for each cell, but when analyzing the results, the simulated steps that produced the top 5% of activation values (***Equation 6***) across the entire recording session were used to specify the spike map for the entorhinal cells. Unlike the entorhinal cells, spike maps for place cells were determined more directly; for place cells, there is a spike every time that place cell activation is greater than the activation threshold, θ_a.

$$act_{mECi} = r_i(d) + R * act(p_{RM}) * w_{RMi} \tag{6}$$

Equation 6 is the real-valued activation of entorhinal cell *i* (e.g., a border cell, head direction cell, or *k* cell) at the time of memory retrieval based on the current bottom-up input response as determined in ***Equation 1***, combined with top-down memory feedback, which is weighted by the retrieval constant, *R*, multiplied by the activation of the place cell for the retrieved memory as determined by ***Equation 2***, and the appropriate bidirectional weight between the entorhinal cell *i* and the retrieved memory. The feedback retrieval constant was set to 3 to ensure that the retrieved memory was able to re-instantiate the entorhinal inputs by overriding the perceptual input to the entorhinal cells (this assumption is not critical and lower values of *R* produce similar results).

Close consideration of the model details reveals an apparent inconsistency; if entorhinal inputs are only active 5% of the time, how can they provide the sine wave neural tuning response functions dictated by ***Equation 1***? This apparent inconsistency is resolved by hypothesizing that different entorhinal cells have different preferred firing rates. For instance, in addition to 5% firing rate cells, other entorhinal cells might fire 10%, 50%, or even 90% of the time. The sine wave tuning function is assumed to reflect the summed activity for a population of entorhinal

cells that have the same weight connections to hippocampus and same weight connections to cortical regions providing excitatory perceptual input to the entorhinal cortex; when adding up the responses of similarly connected entorhinal cells, the weighted sum is a sine wave. Thus, when activating hippocampal memories, it is the sine wave (population code) that dictates which memories are retrieved, but when analyzing the results in terms spike rate maps, specific examples of the population are considered; namely entorhinal cells with 5% firing rates.

The effect of considering cells with different firing rates will be quantitative, rather than qualitative. For instance, border cells with firing rates greater than 5% will fire not just immediately adjacent to the border, but also fire at greater distances from the border. Grid cells with firing rates greater than 5% will exhibit the same grid spacing, but their grid fields will be larger, with smaller gaps between the grid fields.

Results

The Population Response of Border Cells Specifies Locations on a Torus

If memory consolidation of hippocampal place cells makes memories dissimilar from each other, this can explain the finding that grid fields often appear to be centered on positions outside the enclosure, with just the edge of the grid firing field falling inside the enclosure (this behavior can be seen in the raw firing maps of all papers reporting grid cell activity). In the model, this occurs because hippocampal place cell memories created inside the box, but close to a border, can have their place field centers pushed outside the box during memory consolidation. This capacity to represent place cell memories outside the box necessarily occurs with circular tuning functions (e.g., for a circular dimension, approaching the East border from the West is similar to approaching the West border from the East). If the period of the circular border cells was the same as the width of the box, then a memory pushed outside the box on one side would appear on the opposite side of the box, in which case the partial grid field on one side should match up with its remainder on the other side. This would entail complete confusion between opposite sides of the box, and the representation of the box would be a torus (donut-shaped) rather than a flat two-dimensional surface. To reduce confusion between the two sides of the box, the width of the enclosure in which the animal navigated (***Figure 5***) was assumed to be half as wide as the full period of the border cells. In other words, although the representational space is a torus (or more specifically a three-torus), it is assumed that the real-world two-dimensional surface is only a section of the torus (e.g., a square piece of tape stuck onto the surface of a donut). This assumption provides representational space that is outside the box without such locations wrapping around to the opposite side of the box.

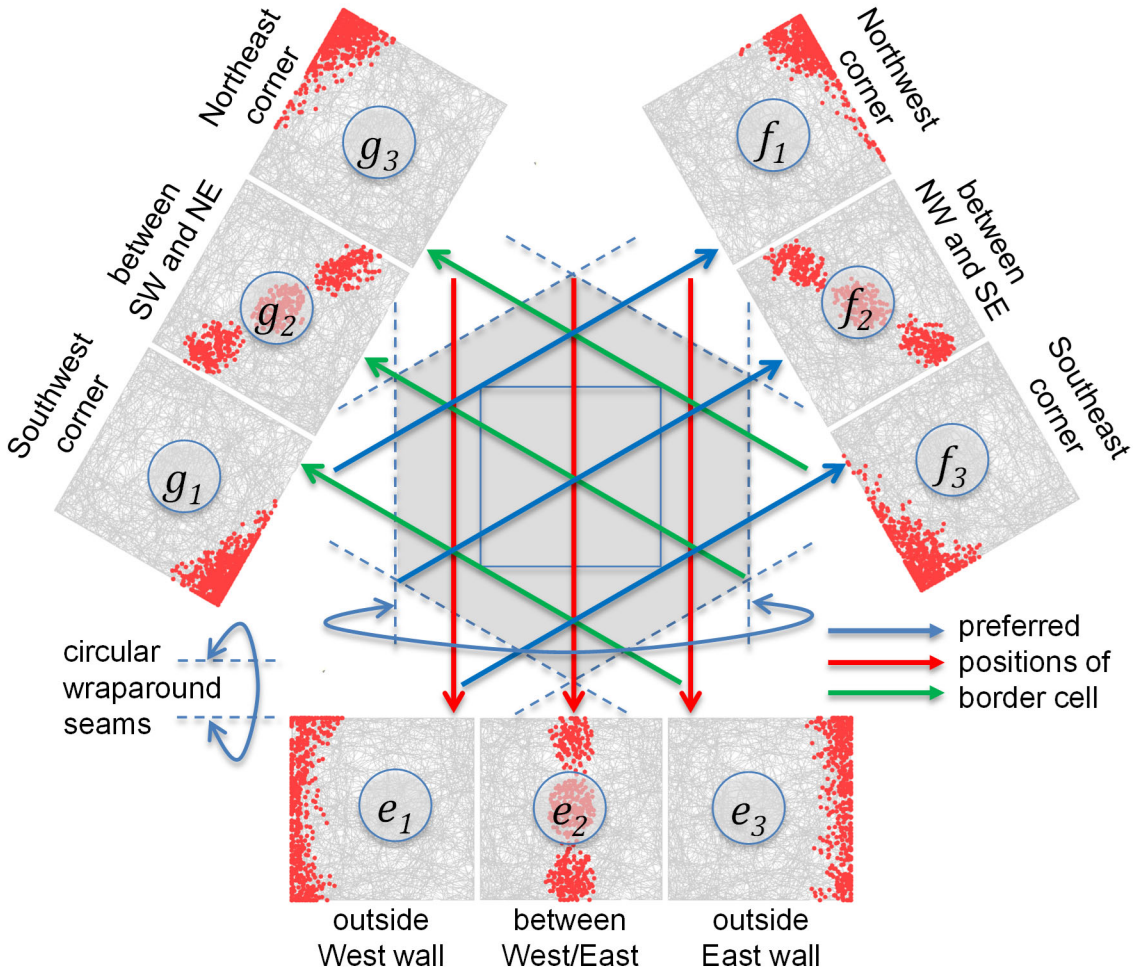

Figure 5. Position of square recording enclosure relative to circular border cell dimensions. The 9 outer graphs show simulated spike rate maps for the border cells under the assumption that each border cell fires 5% of the time based on the summation of its bottom-up and top-down inputs (***Equation 6***). The enclosure was assumed to be half as wide as the full period of the circular border cells, minimizing confusion between opposite borders. The dashed lines indicate wraparound seams such that each dashed line is connected to the opposite dashed line to create a cylinder for that dimension, resulting in a hexagonally connected 3-torus for the entire space across the three non-orthogonal dimensions. The gray curved lines within each of the 9 border cell graphs show the path of a simulated animal across 10,000 steps. The red dots (500 per graph) show positions where the simulated border cell fired. The graphs for the *G* and *F* dimensions are rotated to align the graphs with the corresponding allocentric directions. The preferred positions for each of the 9 cells are indicated by the entire line length of the red, green, or blue arrows that point to the corresponding firing map. The letter labels inside each graph indicate the simulated cell using the same labeling scheme as in ***Figure 2***.

In addition to explaining outside-the-box grid fields, the use of circular border cells explains the finding that the population response of grid cells lies on a torus-like surface (Gardner et al., 2022). More specifically, the model predicts that place cells lie on a torus defined by circular

border cells and because place cells are the cause the grid fields, the population code of grid cells also lies on a torus. This is the first computational model to explain the toroidal nature of grid cells.

Because the three border cells are equally spaced with a period of 2, a box of width 1 results in a situation where the border cells prefer positions immediately outside of the box. For instance, as shown in ***Figure 5***, the West border cell would be most active if the animal were 1/6 of a distance East of the West border (i.e., positions along the red arrow, which is placed just outside the box). Because the animal is never given the opportunity to explore outside the box, this particular cell primarily fires when the animal is immediately adjacent to the West border (the closest allowable position to the cell's preferred position). The other two border cell basis sets are at an angle to the recording box, preferring positions along tipped axes (the blue and green arrows show the preferred positions). These could be described as corner cells, although because they are at a 60-degree tilt rather than 45-degrees, they end up firing primarily along the top and bottom borders.

Border cells are not always uniform in their responsiveness along a particular border (Solstad et al., 2008), although it is not clear whether such inhomogeneities are consistent with the predicted corner cell responses seen in ***Figure 5***. In addition, the model predicts the existence of cells that fire between the borders, but not at the border, although such cells are relatively rare (Solstad et al., 2008). However, these predictions of corner-cells and between-border cells are not crucial to the proposed memory theory of grid cells. For instance, if the model contained 6 different non-orthogonal dimensions, rather than 3, not only would it provide a more accurate approximation to Euclidean distances (see ***Figure 3***), but it would be easier to find border cells that were closely aligned with the orientations of the borders. Similarly, rather than assuming 3 entorhinal cells for each dimension, which requires a 2-to-1 ratio of border cells to between-border cells, it is possible to configure the model with just 2 opposing border cells for each dimension, in which case there are no between-border cells. These assumptions regarding the number of non-orthogonal dimensions and the number of cells per dimension were made primarily to maximize simulation efficiency. The core of the model is its prediction that place cells are junctions of *what* and *where* (i.e., memories) and that grid cells are non-spatial attributes (*what*).

Grid Fields are Immediately Apparent and Align with Enclosure Walls

Behavior of the model was first explored without head direction (i.e., allocentric memories), based on the real-world two-dimensional inputs (*X/Y*) as captured by the three non-orthogonal basis sets (*E/F/G*), in combination with a single non-spatial cell *k*. ***Figure 6*** shows firing maps for the non-spatial cell during initial recording in a novel box (i.e., the first 10,000 simulation steps) as well as the same simulation after 10 prior sessions (familiar box) of experience (e.g., recording after 100,000 prior steps). To explore how the model behaves, this was done with three different consolidation thresholds, which affect grid spacing. For these simulations, the results from cells with a 10% firing rate (rather than 5%) were analyzed to ensure coverage of all grid fields for the case of a closely spaced grid. As discussed in the model methods, the model assumes a population of similarly connected entorhinal cells that collectively produce the sine wave tuning functions (***Equation 1***). As such, the choice to analyze cells with different firing rates amounts to considering grid cells that have larger versus smaller grid fields, but with

equivalent spacing/orientation between grid fields. In this case, 10% firing rate grid cells were selected for their property of exhibiting grid field sizes that allow easy visual assessment of the grid pattern.

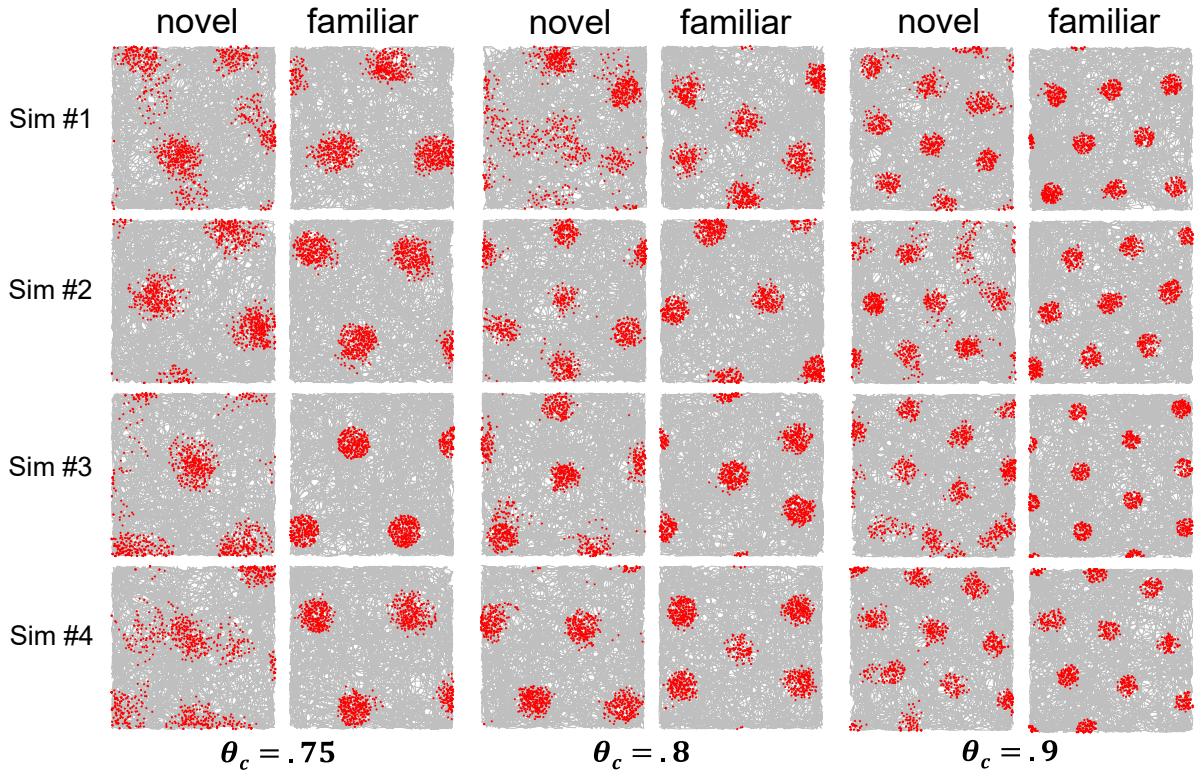

Figure 6. Results for simulated grid cells representing a non-spatial attribute common to a set of place cell memories when not including head direction in the place cell memories (allocentric memories). These simulations are an exploration of how the model behaves with different parameter values. In each case, the first four simulations are shown, regardless of outcome. Results are shown when adopting one of three different consolidation thresholds, θ_c, which produce different spacings between memories. The corresponding activation thresholds, θ_a, were .86, .9, and .92 to make sure that memories were created and activated with a somewhat closer spacing than that dictated by the consolidation threshold. Each pair of novel and familiar firing maps is the same simulation, with the novel firing map showing the first 10,000 simulated steps in a novel environment and the familiar firing map showing 10,000 simulated steps after 100,000 prior steps (e.g., after the equivalent of 10 sessions of prior experience).

As seen in ***Figure 6***, in some cases (e.g., the bottom-right two graphs) the orientation and position of the grid hardly changes with additional experience whereas in other cases (e.g., the bottom-middle two graphs) the position and/or orientation of the grid changed. Because the consolidation threshold remained the same with experience, the grid spacing remained the same. It has been reported that grid spacing tends to shrink with experience (Barry et al., 2012) and this could be captured by gradually increasing the consolidation threshold. In some cases, the novel grid was somewhat smeared out for some of the grid fields, reflecting the ongoing consolidation of memories. This lack of grid regularity for a novel environment has been reported (Barry et al., 2012). In nearly all cases, there were grid fields that appeared to be partially outside the box, as is typically observed in the firing maps of grid cells.

In real data, the grid array tends to align with a straight wall of the enclosure (Krupic et al., 2015) and this is true of many of the simulations in ***Figure 6***. To better quantify this result, an additional 20 simulations were performed at each consolidation threshold after 30 prior sessions experience (i.e., excessive prior experience to ensure stabilization of consolidation). For the .75 consolidation threshold (far spacing), 10 simulations produced a grid that was horizontally oriented and 5 were vertically oriented. The remaining 5 simulations produced a grid with 4 main firing fields that were arrayed in square arrangement rather than a hexagonal arrangement. For the .8 consolidation threshold (intermediate spacing), all grids were hexagonal, with 2 of the simulations producing a grid that was horizontally oriented and the remaining 18 producing a grid that was vertically oriented. For the .9 consolidations threshold (close spacing), all grids were hexagonal, with 6 of the simulations producing a grid that was horizontally oriented, while 6 simulations produced a grid that was vertically oriented. The remaining 8 simulations were tipped at some orientation other than vertical or horizontal. In summary, nearly all grids were hexagonal and the grid usually, but not always, aligned with the vertical or horizontal axis of the square recording enclosure, with this being more likely to occur with far spacings.

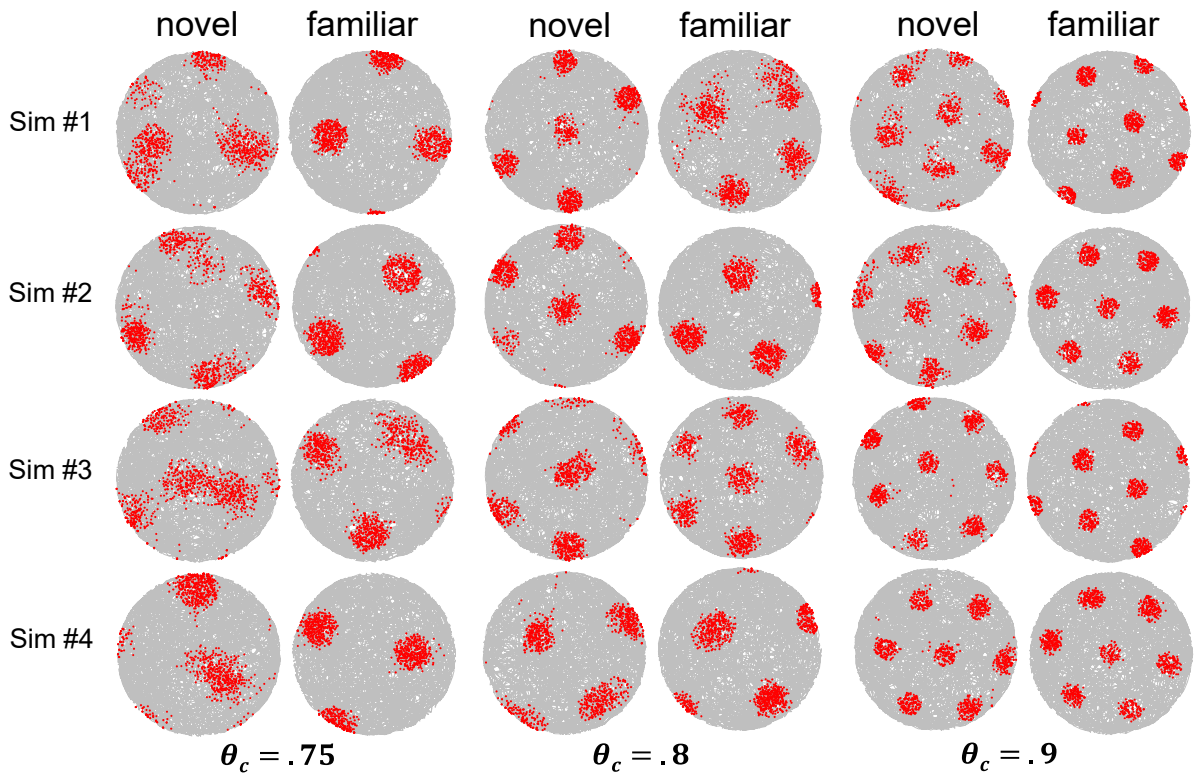

Figure 7. Simulation results when using the same parameters and settings as in ***Figure 6*** for a circular enclosure.

It is not clear from these results whether the tendency of the grid to align with the box reflects the geometry of the box or the assumption that border cell dimension *E* aligns with one of the box borders. To address this question, another set of similar simulations was run with a circular enclosure, but with the dimension *E* still in the horizontal direction, to see if the grid still tended to be horizontally or vertically oriented. As seen in ***Figure 7***, hexagonal grids readily emerge in a circular enclosure (i.e., the grid does not require straight walls). Overall, the

orientation of the grids still seemed to align with the vertical or horizontal directions, where the definition of vertical and horizontal is now relative to the underlying border cell dimension *E*. However, this alignment occurred slightly less often as compared to the square enclosure. To quantify this observation, 20 simulations were run at each of the three grid spacings after 30 sessions of prior experience. For the .75 consolidation threshold (far spacing), 10 simulations produced grids that were horizontally oriented and 2 were vertically oriented. An additional 7 were tipped at random angles and the remaining simulation produced a square grid that was horizontally/vertically aligned. For the .8 consolidation threshold (intermediate spacing), 4 simulations produced horizontally oriented grids and 10 produced vertically oriented grids. An additional 3 were tipped at random angles and the remaining 3 failed to produce a hexagonal grid (the grid was either a square grid or 4 place fields arranged in a Y pattern). For the .9 consolidations threshold (close spacing), one simulation produced a grid that was horizontally aligned and 3 were vertically aligned. For the remaining 16 simulations, it was tipped at some orientation other than vertical or horizontal.

In summary, these simulations show that the grid is immediately established and that the grid has a tendency to align with the underlying border cell directions (e.g., the border cell direction *E* or perpendicular to *E*), particularly with a widely spaced grid in the case of a square enclosure (i.e., the geometry of the enclosure also seemed to play a role in light of the reduced vertical/horizontal alignment within a circular enclosure). Thus, the finding that the grid tends to align with walls of a square enclosure (Krupic et al., 2015) follows from the assumption that one of the underlying border dimensions aligns with the walls of a square enclosure.

Additional analyses revealed that this tendency to align with border cell dimensions is caused by weight normalization (***Step 6*** in the pseudocode). Specifically, connection weights cannot be updated above their maximum nor below their minimum allowed values. This results in a slight tendency for consolidated place cell memories to settle at one of the three peak values or three trough values of the sine wave basis set. This "stickiness" at one of 6 peak or trough values for each basis set is very slight and only occurred after many consolidation steps. In terms of biological systems, there is an obvious lower-bound for excitatory connections (i.e., it is not possible to have an excitatory weight connection that is less than zero), but it is not clear if there is an upper-bound. Nevertheless, it is common practice with deep learning models include an upper-bound for connection weights because this reduces overfitting (Srivastava et al., 2014) and there may be similar pressures for biological systems to avoid excessively strong connections.

Grid Cell Modules Reflect the Dense Packing of Memories in Three Dimensions: An Example with X, Y, and Head Direction

One of the key findings in the grid cell literature is the existence of anatomically arranged modules, with nearby grid cells having the same orientation, spacing, and distortion (i.e., the degree to which the grid is slightly elongated in a particular direction), but different spatial phases (Stensola et al., 2012; Yoon et al., 2013). In other words, the grid of nearby cells is often identical but shifted. Some theories suggest that these grid cell modules provide a two-dimensional Fourier basis set for location coding (Rodríguez-Domínguez & Caplan, 2019; Stachenfeld et al., 2017; Wei et al., 2015). However, the proposed non-spatial memory model of grid cells suggests a different interpretation – rather than tiling a two-dimensional space, grid cell modules might emerge from the consolidation of memories into a three-dimensional volume. For memories that differ from each other in three real-world dimensions, with a fixed minimum

distance between memories, their arrangement is equivalent to the dense packing of equal-sized spheres in a three-dimensional volume (e.g., putting marbles in a glass). In this case, the arrangement of memories should follow the Kepler conjecture (Hales, 2005), which states that there are two ways in which equal-sized spheres can be arranged with the densest possible packing. Both solutions are hexagonal in nature, containing separate layers of spheres that are hexagonally arranged with the same orientation and spacing, but with different phases: face-centered cubic packing (FCP), entails three different hexagonal layers, whereas hexagonal close packing (HCP) entails two different hexagonal layers.

Consider the 3D volume defined by *X*, *Y*, and head direction, *H*. If egocentric memories include head direction (i.e., memories of where a non-spatial attribute can be found and the viewpoint at the time of discovery of the non-spatial attribute), then consolidation may lead to a close-packed FCP or HCP arrangement. In this case, the activity of different head direction cells might capture different hexagonal layers of the dense packed lattice of memories. If so, each head direction cell will exhibit a grid pattern of the same orientation and spacing as other head direction cells, but with the spatial phase of the grid shifted as dictated by FCP or HCP geometry. In the simulation results reported below, the three-dimensional memories naturally settled on the FCP solution (for related theoretical proposals that involve 3D hexagonal tiling, see Mathis et al., 2015; Stella & Treves, 2015), with three layers that were separately captured by the three different head direction cells in the head direction basis set (***Figure 8***).

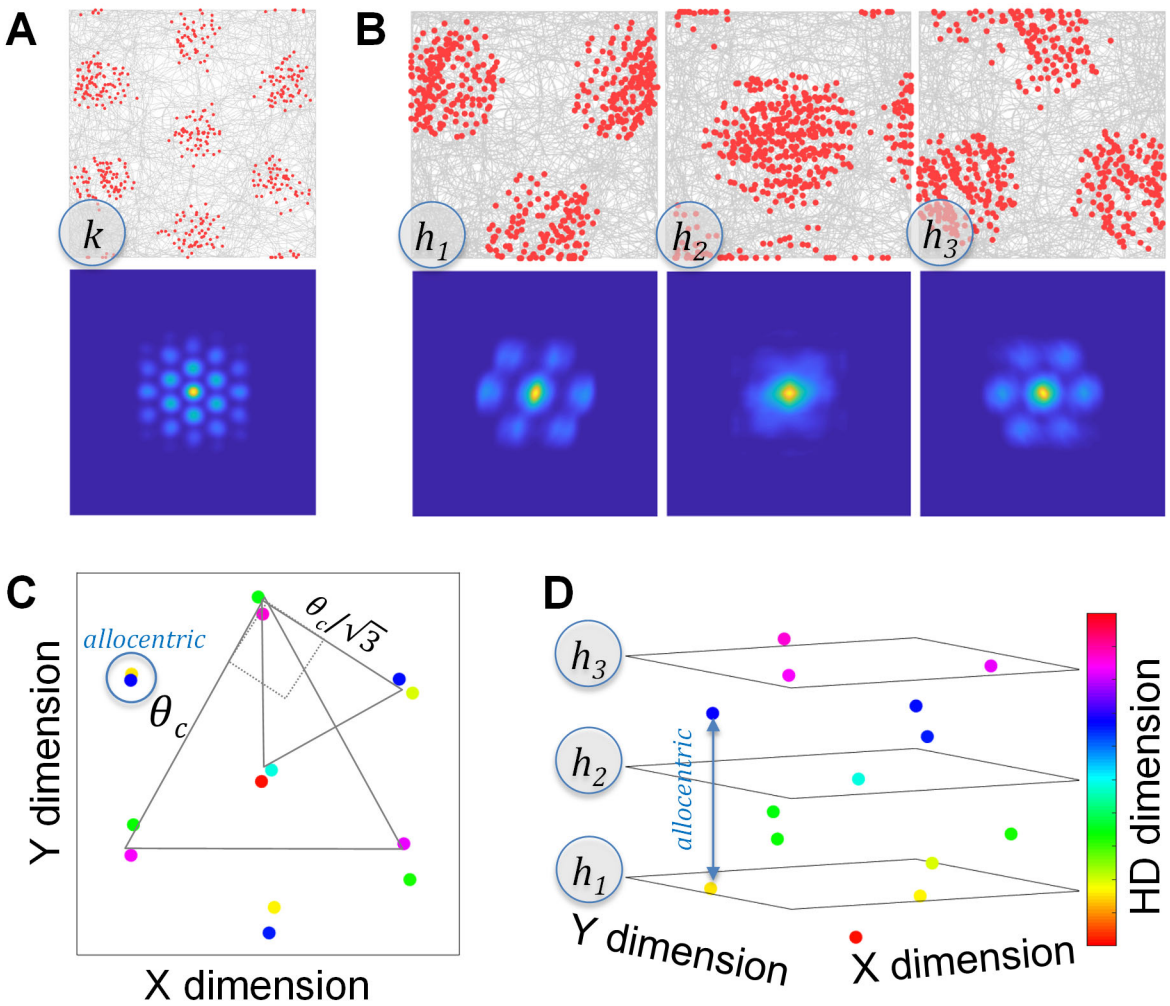

Figure 8. Results from a simulation that included head direction (.8 consolidation threshold). These results are for a familiar box (10 sessions of prior experience). The plotted firing maps (red dots for spikes) and spatial autocorrelation (blue for low correlation and yellow for high correlation) maps are labeled according to the specific cell labels used in ***Figure 2***. The border cell results for this particular simulation are shown in ***Figure 5***. (**A**) The grid fields for the non-spatial attribute, *k*, are vertically aligned. (**B**) In contrast the non-spatial attribute, the grid fields for the head direction cells are horizontally aligned. Cells h_1 and h_3 have the same grid orientation and spacing as each other, as would cell h_2, except that for h_2, its corresponding place cell memories are mostly outside the box. As in real grid cell modules, the grid fields for each of the three head direction cells (h_1 *to* h_3) are shifted relative to each other. (**C**) The head direction grid spacing is the square root of 3 larger than the non-spatial attribute grid spacing, reflecting the superposition of face-centered cubic packing layers. (**D**) For this grid spacing, there were 6 HD layers, such that every third layer produced a nearly identical spatial phase of HD-sensitive place cell memories, except that the preferred HD was 180 degrees opposite. This provides an allocentric arrangement of the egocentric place cell memories whereby the non-spatial attribute at each location was remembered from two 180 degree opposite viewpoints (head directions). The colored dots in the three-dimensional plot are based on the final positions of the 14 memories at the end of the simulation, as calculated from the weight matrix for each place cell, with the color of the dots showing the preferred head direction of the corresponding memory.

Figure 8 shows a typical result for the non-spatial attribute cell (***Figure 8A***) and the three head direction cells (***Figure 8B***) while navigating in a familiar box. The 6 border-cell firing maps for this particular simulation were plotted in ***Figure 5***. The non-spatial attribute cell produced a grid field that was rotated by 90 degrees relative to the grid fields of the head direction cells and a grid spacing that was more closely spaced by a factor equal to the square root of three, reflecting the superposition of FCP layers (***Figure 8C***). This coordination between grid cells of different grid spacings was recently been observed (Waaga et al., 2022), and in the model this coordination arises from memory retrieval for the same set of memories. The place cell memories were arranged into 6 different FCP layers (a repeating sequence of 3 layers), providing an allocentric representation of a familiar box based on combinations of egocentric memories (i.e., based on place cell memories that include head direction). More specifically, each location is remembered from two opposite viewpoints (***Figure 8D***). Two of the three head direction cells exhibited the same grid orientation and spacing but the grid fields were shifted. The third head direction cell would have produced something similar except that its place cell memories were mostly outside-the-box (the spike rate maps show hints of the outside-the-box memories near the borders). In keeping with this result for simulated head direction cell h_2, it has been reported that some mEC cells that initially appear to have a single place cell firing field are in truth grid cells when the box borders are removed – removing the borders reveals that there are additional grid fields outside the box (Savelli et al., 2008).

In summary, this simulation produced an egocentric head direction grid cell module in combination with an allocentric grid array for the non-spatial cell (places where the non-spatial property k can be found), with this second array exhibiting a closer spacing by a factor of the square root of 3 and an orientation that is rotated by 90 degrees.

Stensola et al. (2012) simultaneously recorded as many as 186 mEC grid cells in individual rats, finding approximately 4 different grid modules for each animal. The ratio of grid spacing between modules was found to be approximately the square root of 2 rather than the square root of 3 when comparing modules that had grid spacing values that were close in value. However, an alternative interpretation of the Stensola data would place the modules into interleaved sets of two (each pair consisting of one egocentric and one allocentric module) with the two modules in a pair differing in orientation by 90 degrees and differing in grid spacing by the square root of 3. For instance, perhaps one egocentric/allocentric pair of mEC grid modules is based on head direction (viewpoint) in remembered positions relative to the enclosure borders whereas a different egocentric/allocentric pair is based on head direction in remembered positions relative to landmarks exterior to the enclosure. This might explain why a deformation of the enclosure (moving in one of the walls to form a rectangle rather than a square) caused *some of the grid modules but not others* to undergo a deformation of the grid pattern in response to the deformation of the enclosure wall (see also Barry et al., 2007). More specifically, if there is one set of non-orthogonal dimensions for enclosure borders and the movement of one wall is too modest as to cause avoid global remapping, this would deform the grid modules based the enclosure border cells. At the same time, if other grid modules are based on exterior properties (e.g., perhaps border cells in relation to the experimental room rather than the enclosure), then those grid modules would be unperturbed by moving the enclosure wall.

Similar to head direction, there are other factors that vary during the recording session, which may explain grid modules that don not depend on head direction. For instance, in the current simulation, it was assumed that the animal adopted a new random goal direction at every time step, independent of head direction (e.g., the animal decides to alter its course slightly towards

direction, *g,* for the next step, even though its current head direction is, *h*). To assess whether goal direction might form a grid module, another simulation was run in which the third dimension was based on a basis set of three goal direction cells (g_1 *to* g_3) rather than head direction. The results were essentially identical to that shown in ***Figure 8,*** swapping the labels of h_1 *to* h_3 for g_1 *to* g_3, except that the goal direction cells were insensitive to head direction (close examination of the firing maps in ***Figure 8B*** reveals that they tend to occur on trajectories that align with the preferred head direction of the corresponding head direction cell, but this was not true of the goal direction simulation).

Grid Cells Change their Receptive Fields in the Absence of Hippocampal Feedback

A key prediction of this account is that in the absence of memory retrieval from hippocampal place cells, each mEC grid cell should revert to its underlying bottom-up receptive field. For instance, in the absence of top-down memory retrieval feedback, a head direction conjunctive grid cell should become a head direction cell at all positions rather than just at grid field positions. Such a finding was reported by Bonnevie et al (2013). More generally, this kind of result should occur for all grid cells. For instance, if the bottom-up receptive field for a grid cell is a tuning function for a preferred odor, then that cell should exhibit odor-sensitivity in all positions in absence of hippocampal feedback rather than only responding in remembered positions. In brief, this account assumes that *all grid cells are conjunctive grid cells* except that the nature of the non-spatial attribute that makes them conjunctive is not yet known for many grid cells (see Aronov et al., 2017, for an example of what might be sound conjunctive grid cells).

On this account, grid cells only appear to be non-conjunctive because their preferred non-spatial attribute was constantly present during the recording session (e.g., an attribute that was found everywhere in the enclosure, such as an odor, or an attribute that was a constant property during the recording session, such as an electronic hum). However, if the preferred non-spatial attribute of an entorhinal grid cell was only found in half of the box, that cell would be predicted to exhibit a grid field pattern for that half of the box while remaining silent in the other half. This suggests experiments in which the non-spatial attributes of entorhinal cells are first characterized in the absence of navigation and then, based on these results, the enclosure is given those non-spatial attributes in some regions but not others. In other words, experiments using non-uniform enclosures with salient differences for different surfaces (e.g., an electronic hum that is played when the animal is one half of the box, but not the other half).

Grid Fields Appear Immediately in a Novel Environment whereas Place Cell Positions Take Time to Stabilize

The head direction grid module results reported above were based on a familiar environment, with 10 sessions worth of prior experience. However, grid field patterns are often seen immediately upon entry into a novel environment, although the pattern may be somewhat less regular for a novel environment (Barry et al., 2012). The two-dimensional allocentric simulations (i.e., the simulations that did not include head direction) produced a grid immediately in a novel environment. As shown in ***Figure 9A***, this is also the case for the three-dimensional egocentric memories that included head direction when the results are analyzed in terms of the non-spatial attribute cell. However, for the head direction conjunctive grid cells, the grid pattern emerged

more slowly, owing to ongoing consolidation of the place cell memories (the graph plots the grid score for the head direction cell with the highest grid score of the three head direction cells, h_{max}, considering that some of the head direction cells fail to produce a grid owing to outside-the-box memories, as was the case for cell h_2 in ***Figure 8B***).

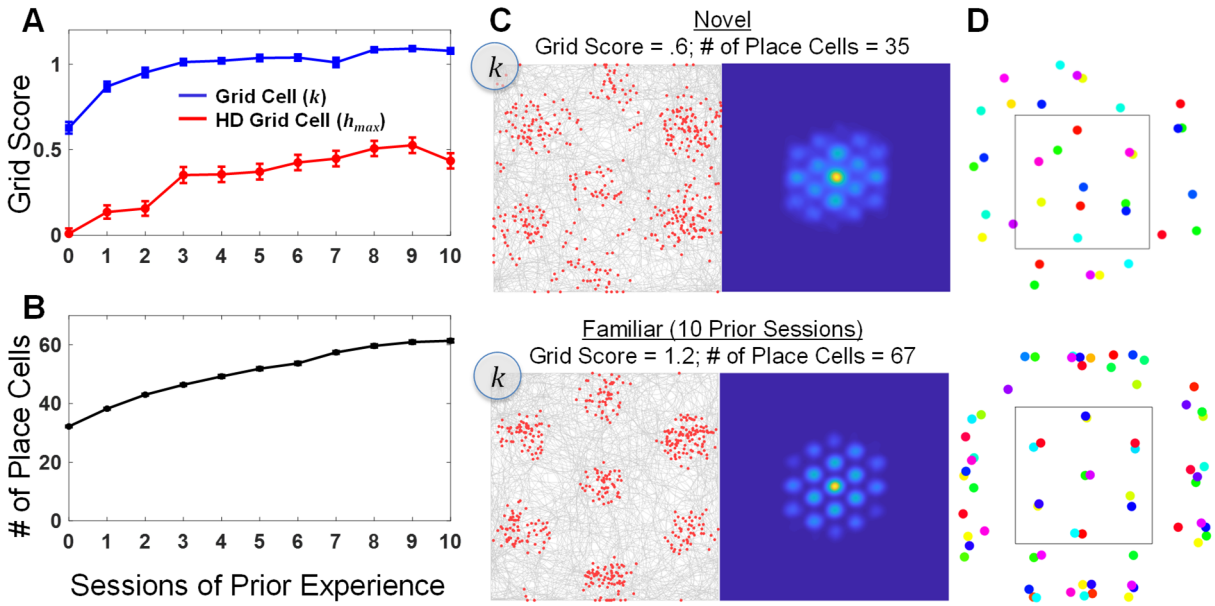

Figure 9. Simulations depicting the learning of non-spatial grid fields (*k*) and head direction conjunctive grid field (*h*), using the same parameter values as in ***Figure 8***. (**A**) Average grid scores across 100 simulations, with prior experience ranging from novel (no prior experience) to 10 sessions of prior experience, show that the grid field of the non-spatial cell, *k*, was immediately apparent whereas the head direction grid cells required prior experience before grid fields stabilized. Because one of the three head direction grid cells tended to have a single central grid field, with potentially outside the box additional grid fields (see cell h_2 in ***Figure 8***), the maximum grid score (h_{max}) from the three head direction cells was used. Error bars are plus and minus one standard error of the mean. (**B**) A plot showing the number of place cell memories created by the end of the simulation reveals that the number of memories grew in a similar manner to the stabilization of the head direction grid fields. (**C**) Representative firing map and spatial autocorrelation results for a novel environment and a familiar one reveal that that the non-spatial cell's grid fields were regular but less precise in a novel environment, reflecting the shifting nature of the place fields. (**D**) Plots showing the final positions of all place cell memories for the representative simulations, reveal that the main effect of prior experience was formation of outer memories that conform to the shape of the box; as the animal learned head direction representations for the borders of the box, the grid cell module based on head direction stabilized. The color of the dots represents the head direction associated with each memory using the same color scale as ***Figure 8***.

Critical to understanding why the non-spatial attribute grid cell reveals a hexagonal grid during the initial recording session in a novel environment is the realization that multiple memories feedback to the non-spatial attribute grid cell and, furthermore, it may be that a particular grid field location reflects multiple memories during the recording session (e.g., a particular grid location reflects retrieval of memory-A at one point in time but later it reflects

retrieval of memory-B, perhaps because consolidation has shifted memory-B to that location). Because the non-spatial cell's grid field reflects on-average memory positions during the recording session (i.e., the locations where the non-spatial attribute is more often remembered, even if the locations of the memories are shifting), the grid fields for the non-spatial are immediately apparent, reflecting the tendency of place cells to linger in some locations as compared to other locations during consolidation. More specifically, the place cells tend to linger at the peaks and troughs of the border cell tuning functions (see the explanation above regarding the tendency of the grid to align with border cell dimensions). By analogy, imagine a time-lapsed birds-eye view of cars traversing the city-block structure of a densely populated city; this on-average view would show a higher density of cars at the cross-street junctions owing to their tendency to become temporarily stuck at stoplights. However, with additional learning and consolidation, the place cells stabilize their positions (e.g., the cars stop traveling), producing a consistent grid field for the head direction conjunctive grid cells. This slow stabilization of place fields is a known property (Bostock et al., 1991; Frank et al., 2004).

In summary, these model simulation results indicate that some grid patterns are learned more slowly than others for a completely novel environment (i.e., an environment that is completely foreign to the animal). More specifically, egocentric grid cells (e.g., head direction conjunctive grid cells) require stabilization of the place cell memories in the face of ongoing consolidation whereas allocentric grid cells reflect on-average place field positions. However, it should be noted that these simulations were for a "blank slate" animal that had no prior experience with any similar enclosures. This is unlikely to be the situation in the real world, and it may be that the empirical finding that head direction conjunctive grid cells can be found immediately reflects memories of similar experiences in similar enclosures rather than very recent memories from the last few minutes or seconds (the role of memories prior to the recording session are considered in the remapping simulations below, which produced immediately apparent head direction conjunctive grid cell responses).

Consolidation of Interior Place Cells Depends on Exterior Place Cells: The Cognitive Map Emerges after Learning Specific Views of the Boundaries

Figure 9B shows that the main effect of prior experience is the addition of new place cell memories, which produce more precise grid fields for the non-spatial attribute cells (***Figure 9C***); these new place cells result in stability (i.e., consolidation no longer moves the positions of the place cells). The place field locations for these additional memories were primarily outside of the box (***Figure 9D***). Considering that place fields are the cause of individual grid fields for the model, this may relate to the finding that encounters with boundaries can serve to stabilize grid fields (Hardcastle et al., 2015). More specifically, memories initially formed inside the box in positions adjacent to the borders push against each other during consolidation owing to pattern separation (Chanales et al., 2017), causing some memories to move outside the box. This doesn't necessarily mean that the animal remembers positions outside the box. Instead, these memories could be thought of as characterizations of the borders; by placing their preferred positions outside the box, these memories become highly selective for positions immediately inside the borders of the box (i.e., the inside border locations are the only allowable real-world circumstances that come closest to their preferred inputs). By analogy, the faces of famous individuals are more easily recognized when drawn with impossibly exaggerated facial features (Mauro & Kubovy, 1992), suggesting that the memory representation is a pattern-separated

impossible exaggeration that uniquely identifies that individual. In addition to being highly selective to the borders of the box, these exaggerated border memories constrain the interior memories, which stabilizes interior place cells, revealing the grid fields of the head direction cells.

The effect of outside-the-box place cells is further explored by considering in detail the firing maps of the place cells (***Figure 10***) for a familiar box (this particular simulation is different than the one shown in ***Figure 9***, but used the same parameters as ***Figure 9***). Similar to the "familiar" results shown in ***Figure 9***, ***Figure 10A*** shows the final positions for the 48 place cells that were learned after 10 sessions of prior experience, as well as the firing maps for the three head direction cells (cells h_1, h_2, and h_3 using the cell labeling from ***Figure 2***). Although similar results were show in ***Figure 9***, the results for this particular simulation are shown such that the preferred location of each place cell (***Figure 10A***) can be directly compared to the observed place field firing maps (***Figure 10B***) of each place cell. This comparison highlights how the outside-the-box place cells constrain the nature of the cognitive map.

The place cells were grouped into two types according to observed firing rate: ones that fired less than 2.5% of the time versus ones that fired more than 2.5% of the time (for place cells, firing rates reflect the portion of stimulated steps for which the cell was above the activation threshold). The cutoff of 2.5% was chosen because this cleanly divided the place cells into those with preferred positions outside the box versus inside the box. As shown in left-side of ***Figure 10B***, the less active place cells only fired when the animal was near a border, as expected considering that these place cells had their preferred locations outside-the-box. Therefore, these low activity place cells are termed "exterior" place cells, which capture memories for specific locations near a border, as seen from a particular viewpoint. This viewpoint dependency is specified by the color coding of the spikes shown ***in Figure 10B***, which shows that the exterior place cells only fire when the animal is not only close to the border, but also when the animal has its head pointed in a very specific direction (i.e., each cluster of spikes, which correspond to a particular exterior place cell, has a very specific color).

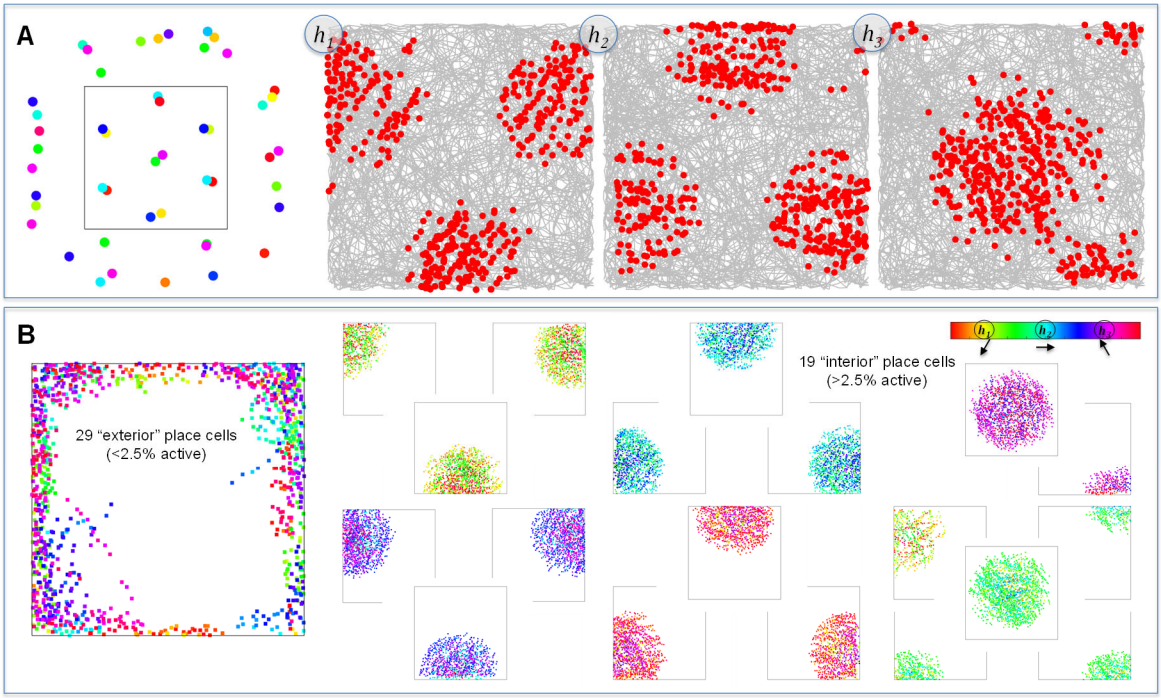

Figure 10. Using the same parameters as ***Figures 8-10***, a comparison of place cell firing maps for interior versus exterior place cells demonstrates how head direction sensitivity depends on location and how interior place cells give rise to head direction conjunctive grid cells. (**A**) The final positions (preferred positions and preferred head directions) of the 48 place cell memories and head direction cell firing maps. (**B**) Firing maps of the place cells. The place cells were divided into 29 "exterior" place cells that were active less than 2.5% of the time (these are accumulated onto the same graph on the left) versus 19 "interior" place cells (one graph per cell on the right). The 8 interior place cells in the top two rows of were the cause of the head direction conjunctive grid cell firing map shown immediately above in panel. The color bar that relates head directions to colors includes the labels for the head direction cells (h_1, h_2, and h_3) and the adjacent arrows depict the three head directions. The 11 interior place cells in the bottom two rows selected for head directions that were between the preferred head directions of the head direction cells (i.e., the other 3 hexagonal layers from the 6-layer hexagonal close packing). The color of each spike for the place cells shows the head direction of the simulated animal at the time of the spike. For the place field center of the interior place cells, head direction sensitivity was weak (i.e., a greater range of colors in the center of each place field), whereas the firing maps for the exterior place cells were highly view dependent (consistent color for each cluster of spikes).

In a few instances, the exterior place cells produced a couple of more interior spikes. This reflects situations in which an interior memory was created during the recording, producing a few more central spikes (e.g., the animal experienced a novel view of a particular border), before the memory was consolidated to an exterior position.

The right-hand side of ***Figure 10B*** shows the "interior" place cells. These 19 cells are arranged in a precise hexagonal pattern, and they are the source of hexagonal grid fields exhibited by the head direction cells. The firing maps for the 8 hippocampal place cells in the top two rows plot the results for head direction-sensitive place cells that have the same preferred head direction as the head direction cell plotted immediately above each graph. In contrast, the

11 place cells in the bottom two rows are sensitive to head directions that were 180 degrees opposite (i.e., head directions that were between the preferred head directions of the three head direction cells). Unlike the exterior place cell firing maps, which were highly selective for head direction, these interior place cells were less selective, firing to a range of head direction values, particularly when the animal was in the center of the place field (i.e., at the center of each cloud of spikes, the head direction colors take on hues other than the preferred head direction of the corresponding head direction cells).

Head Direction Sensitivity of Place Cells Depends on Position Relative to Borders and Place Field Center

For the parameter values used in ***Figure 8-10***, the place cells were fairly sensitive to head direction. Although real hippocampal place cells often show this kind of head direction sensitivity (Rubin et al., 2014), other place cells appear to be insensitive to head direction, particularly in an open field enclosure (Muller et al., 1994). One possibility is that some place cell memories do not receive head direction input, as was the case for the simulations reported in ***Figures 6/7*** – for those simulations, the place cells were entirely insensitive to head direction owing to a lack of input from head direction cells (i.e., each hippocampal place cell memory was necessarily allocentric). However, it has been found that the removal of head direction input to hippocampus affects place cell responses (Calton et al., 2003) and grid cell responses (Winter et al., 2015), suggesting that head direction is a key component of the circuit. Furthermore, if place cells represent episodic memories, it seems natural that they should include head direction (an egocentric viewpoint). When including head direction in the place cell memories, it may be that different parameter values produce place cells that are less sensitive to head direction. If the activation/consolidation threshold parameters are low enough, a hippocampal place cell can become active solely based on location, regardless of head direction. To examine situations with relatively low head direction sensitivity, a new simulation was run by setting the consolidation threshold and activation thresholds to the widely spaced grid values from ***Figures 6/7***, but with the inclusion of head direction in the hippocampal place cell memories.

As seen in ***Figure 11A***, which shows results from a familiar environment, the firing maps for the non-spatial attribute cell (a.k.a. grid cell) and the three head direction cells were very similar to those seen in ***Figures 8-10*** despite the larger grid spacing that arises from a lower consolidation threshold. The pressure to place memories farther apart was satisfied by arranging the place cell memories into 3 face-centered cubic packing hexagonal layers rather than 6 layers. In other words, this change of parameter values resulted in a wider spacing between the memory layers along the head direction dimension. This is shown in ***Figure 11B***, where the plot of the final positions of the memories are non-overlapping rather than producing two oppositely oriented memories in an allocentric pair (the simulation resulted in 45 memories, although the exterior place cell memories are clipped off in the graph to make it easier to see the face-centered cubic packing of the interior 17 memories).

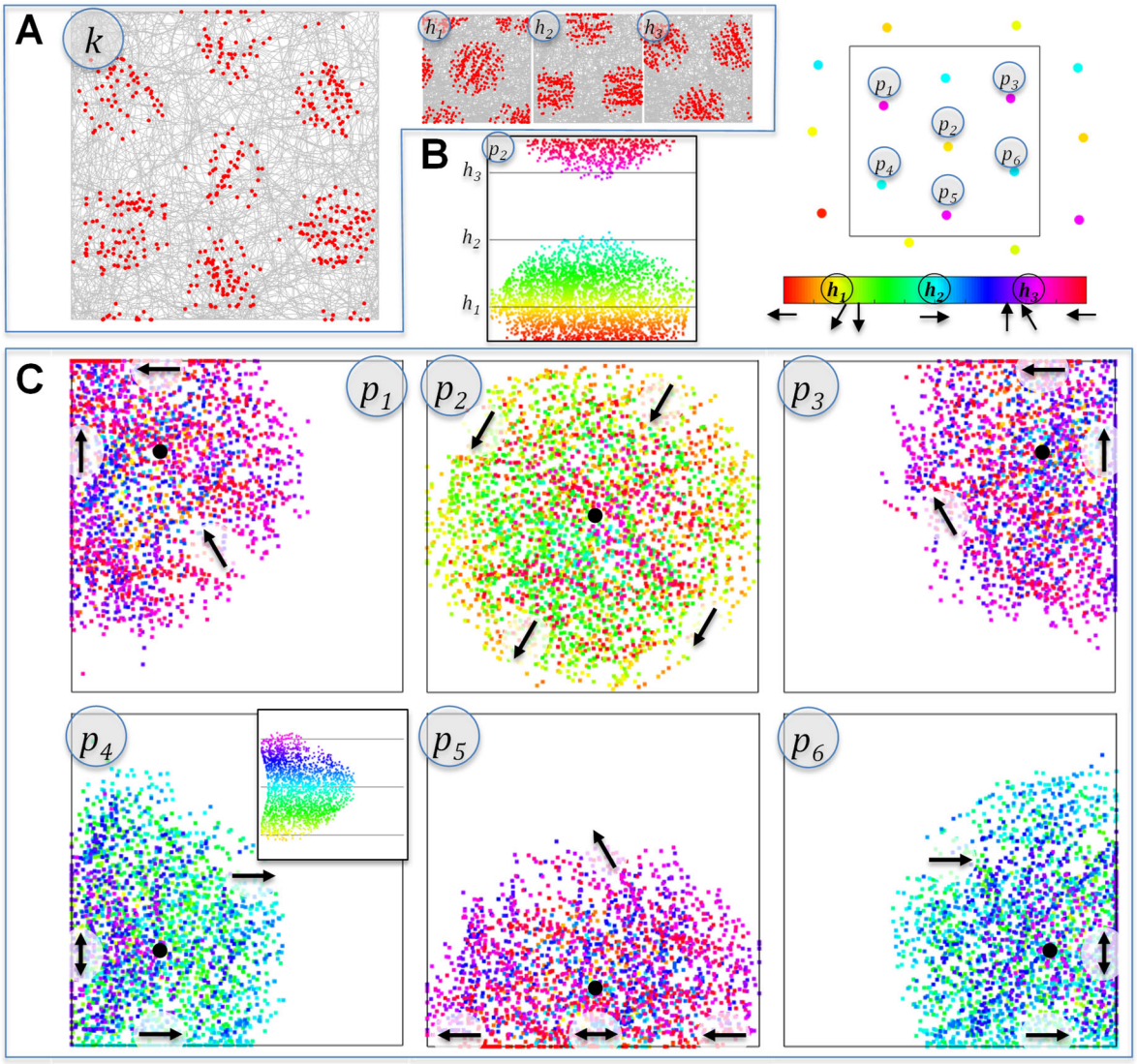

Figure 11. Simulation results using the wide-spacing activation/consolidation threshold parameter values from ***Figure 6/7*** but when including head direction in the hippocampal place cell memories. (**A**) The non-spatial attribute cell (k) and head direction cells (h_1 – h_3) produced similar grid field patterns to those seen in ***Figures 9/10***. (**B**) The right-hand graph shows the final preferred positions of hippocampal place cells, with color indicating preferred head direction. The wider spacing was satisfied through 3 rather than 6 layers along the head direction dimension (only 3 different colors). Six interior place cells (p_1 – p_6) are selected for additional analyses. The lower-left graph plots x-position by head direction for the firing map of cell p_2, collapsing over y-position: the cell was active for a full 2/3 of head directions at its place field center. The color of each spike indicates the head direction of the spike according to the color map. (**C**) Head direction sensitivity of the place cells (shown in six firing map plots for each of six place cells) for the six interior place cells. The black dots show approximate place field centers, and the black arrows show approximate head direction sensitivity in different locations. For cell p_2 head direction sensitive increases with distance from the center. For the other 5 cells, a different pattern emerges along the borders of the box. Along the borders, some head directions are not empirically observed, as shown by the inset graph for cell p_4, which plots x-position by head direction. The absence of data for a subset of head directions gives the spurious appearance of head direction sensitivity that is in line with the borders in one direction or the other, or possibly both directions.

These results demonstrate that the head direction grid module can emerge with other parameter settings. Furthermore, these results indicate that there is more than one way in which an allocentric cognitive map can emerge – even though this simulation did not produce allocentric pairings of oppositely oriented place cells in the same position, the interior place cells were largely insensitive to head direction when the animal was in the exact center of the place field. This is shown in the graph labeled p_2 in ***Figure 11B***, which plots the head direction of each spike for place cell p_2 as a function of x-position, collapsing over y-position. This particular cell preferred the center of the enclosure and responded to a full 2/3 of the possible head directions at the place field center. This occurred even though the head direction of the egocentric memory was in direction h_1 (the yellow color on the head direction color map).

Figure 11C plots the firing maps for 6 of the interior place cells, using the labels from ***Figure 11B***, with head direction coded by color. The black dots show the approximate place field centers for each place cell, and it is at these positions that the cell is relatively insensitive to head direction (i.e., responding to a full 2/3 of possible head directions). But for positions farther from the black dots, the place cell becomes more selective in its head direction, provided that the position is not along a border (discussed below). In other words, these place cells are active if the animal is in the preferred location regardless of head direction, or active for non-preferred positions provided that the head direction matches the preferred direction.

The situation near the borders is more complicated owing to missing data (Muller et al., 1994; Peyrache et al., 2017). Specifically, there is often a failure to observe head directions pointing towards or away from border positions if the animal tends to run alongside the borders (this happens in the simulation owing to momentum). The problem of missing data is highlighted with the black arrows in ***Figure 11C***, which show head direction sensitivity in different locations of the place field. For instance, place cell p_4 appears to prefer both up and down head directions along the West border even though the preferred head direction of this place cell memory was to the East. This occurred because head direction towards the West was never observed at the West border. The inset graph shows the same firing map, but with axes of x-position and head direction, collapsing over y-position, highlighting the absence of East-facing head direction samples at the lowest (i.e., Westernmost) x-positions.

The complicated nature of hippocampal place cell head direction sensitivity seen in the simulation results in ***Figure 11*** may provide an alternative explanation of the results reported by Jercog et al. (2019), which were taken to indicate that head direction sensitivity of hippocampal place cells is relative to a reference point. The reference point model in that study described a number of different behaviors of place cells, including: 1) place cells that seemed to prefer the animal heading towards or away from the place field center; 2) place cells that preferred the animal heading in a circular direction clockwise or counterclockwise around the pace field; and 3) place cells that preferred the animal heading in a particular direction that was far outside the enclosure. Some of these behaviors are approximated by the simulation results in ***Figure 11***, and future work could directly compare this account to the reference point model.

Head Direction Sensitivity of Place Cells Depends on Enclosure Geometry

As reported in ***Figure 10B***, the outside-the-box exterior place cells are more selective for head direction considering that the animal is not allowed to visit these cells' preferred locations: that is, the animal is only ever located in the periphery of these cells' true place fields, where head direction sensitivity is greater. This is equally true for the simulation in ***Figure 11***. This

aspect of model may help explain the empirical result that place cells seem to exhibit greater head direction sensitivity when the animal navigates a familiar narrow passage, such as occurs in a radial arm maze or an elevated track (McNaughton et al., 1983; Muller et al., 1994). Using the same parameters as ***Figure 11***, ***Figure 12*** reports the results of an animal navigating a familiar but extremely narrow enclosure (only 5% as tall as it is wide). In this case, nearly all place cells (14 of the 15 memories) were exterior place cells and, as a result, nearly all place cells exhibited strong head direction sensitivity, preferring travel in one direction or the other along the narrow passage (blue-East or red-West), but not in both directions. In brief, place fields for narrow passages have greater head direction sensitivity because the "true" place field center has been consolidated to a position outside the passage.

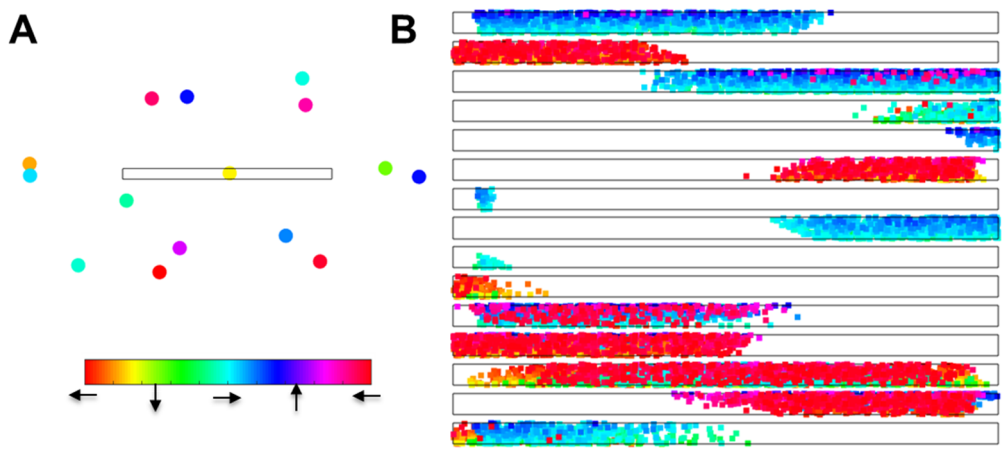

Figure 12. Place cell responses using same parameters as ***Figure 11*** but with an enclosure that is 5% as tall as it is wide, simulating behavior in a highly familiar narrow passage, such as an arm of a radial arm maze or an elevated track. (**A**) The positions of all 15 consolidated memories are shown, with the color indicating preferred head direction according to the color map. (**B**) The firing maps show the place field for each place cell with spike color indicating head direction. Even though these same parameter values produced place cells that were relatively insensitive to head direction in the open field enclosure of ***Figure 11***, in this case each place cell appeared to only prefer one direction or the other (blue or red) along the narrow passage.

As seen in ***Figure 12***, because all but one of the place cells was exterior when the simulated animal was constrained to a narrow passage, the hippocampal place cell memories were no longer arranged in a hexagonal grid. This disruption of the grid array for narrow passages might explain the finding that the grid pattern (of grid cells) is disrupted in the thin corner of a trapezoid (Krupic et al., 2015) and disrupted when a previously open enclosure is converted to a hairpin maze by insertion of additional walls within the enclosure (Derdikman et al., 2009).

Remapping of Place Cells with Changes in Enclosure Geometry

It is beyond the scope of this study to provide a full model of the situations that cause hippocampal place cells to remap their place field positions (Geva-Sagiv et al., 2016; O'keefe & Nadel, 1979) or keep their place field positions but change their firing rate (Leutgeb et al., 2005). However, because this is a memory model, it holds the potential to explain such effects (see Sanders et al., 2020 for a recent review of the remapping literature and application of a belief state model). According to the proposed memory account of place cells, a similar enclosure or a

similar context (i.e., non-spatial attributes) as compared to prior experiences can trigger memory retrieval (see Maurer & Nadel, 2021 for a similar proposal). For instance, similar locations might trigger memory retrieval, but perhaps to a lesser degree (i.e., rate remapping) or perhaps the set of retrieved memories might entail a completely different arrangement if those memories occurred in a similar context but with a different enclosure (i.e., global remapping). Furthermore, as the new enclosure/context becomes familiar, new memories (place fields) may be created and the positions of old memories may be changed owing to consolidation and pattern separation.

As currently implemented, the only non-spatial contextual input to the model is the non-spatial grid cell (the *k* cell). Future work could expand this non-spatial input by including basis sets of different contextual inputs (e.g., a basis set for odor, k_1-k_3, a basis set for surface texture, l_1-l_3, etc.), in which case the model could apply to situations where the geometry and location of the enclosure are kept the same, but the non-spatial characteristics are changed. Nonetheless, the current model can be applied in its current form to changes in enclosure geometry, which has been found to produce both rate remapping (Leutgeb et al., 2005) as well global remapping, depending on how drastically the geometry is changed (Wills et al., 2005).

Using the same parameter values as for ***Figures 8-10***, ***Figure 13A*** shows the grid cell firing map (*k*), hippocampal place cell memory positions (color coded for head direction), head direction grid module firing map (h_1-h_3), and border cell firing maps (e_1-e_3, f_1-f_3, and g_1-g_4) for a recording after 10 sessions of prior experience with a circular enclosure. Critically, the same non-spatial attribute is found for the familiar circular enclosure and the novel square enclosure and both enclosures have same width/height. This situation might correspond to several days of testing with the circular enclosure followed by testing in a square enclosure that is placed in the same broader context (i.e., the same experimental testing room) as prior testing. Prior to the change, it can be seen that the familiar circular enclosure produced a 6-layer face-centered head direction grid module, similar to the results reported in ***Figures 8-10***. The border cell firing maps are shown to highlight that as currently configured, the model tends to produce somewhat patchy border cell responses (i.e., multiple separated hotspots along a line), particularly for the interior border cells (e_2, f_2, and g_2). This occurs because the border cells receive stronger memory retrieval feedback in the locations of hippocampal place cell memories. This behavior is parameter dependent; for instance, border cells with a higher firing rate will be more “filled in” between the hotspots where memory feedback occurs.

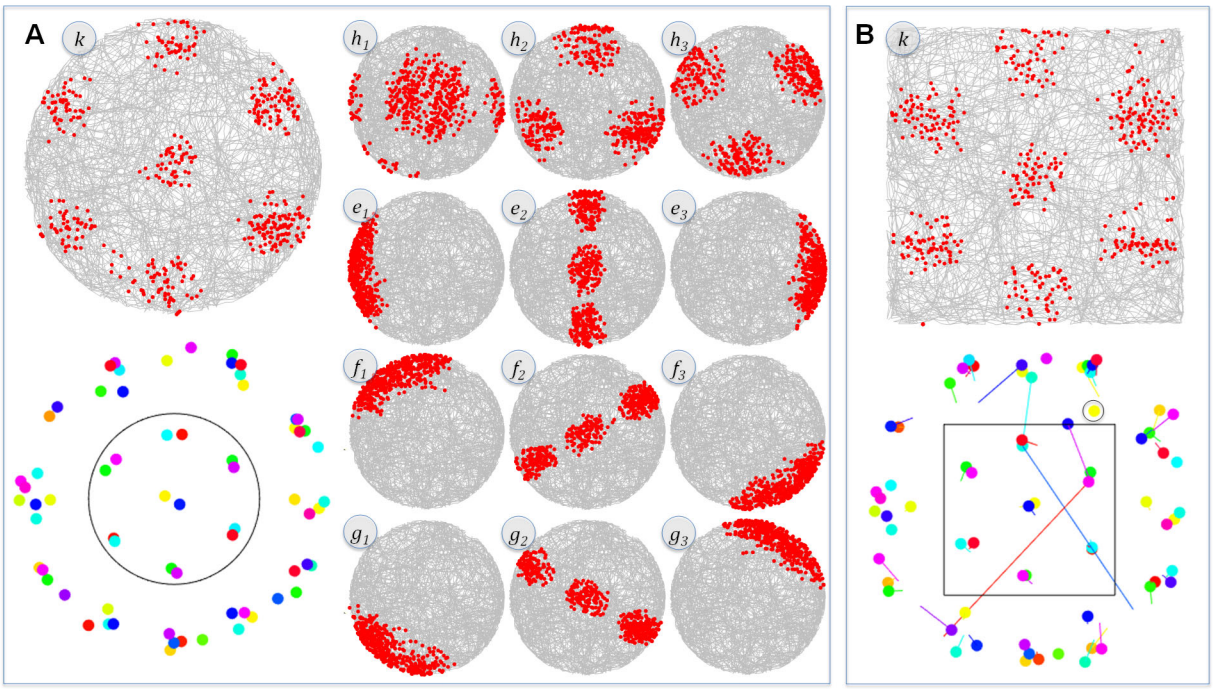

Figure 13. Simulation results for a familiar circular enclosure followed by a change to a square enclosure placed in the same context, containing the same non-spatial attribute. (**A**) The memory locations (dots colored by head direction) and firing maps for the familiar circular enclosure show a 6-layer face-centered cubic packing for the head direction grid module (h_1-h_3). The border cells (e_1-e_3, f_1-f_3, and g_1-g_4) captured one side or the other except for the between-border cells, which show a patchy response owing to memory feedback. (**B**) Upon entry into a novel square enclosure that contained the same non-spatial attribute as the circular enclosure and was the same width/height as the circular enclosure, the head direction grid module was immediately apparent (in contrast to the results reported in ***Figure 9***). This occurred because the memories created in the circular environment were recalled in the square environment and the geometry was similar enough that very few of the retrieved memories changed their locations. The "tails" emanating from each place cell memory location show the shift from the prior location in the circular enclosure to the new location in the square enclosure. The color of the tail indicates the prior head direction of the memory if head direction of the memory changed with consolidation (e.g., the blue memory on the upper border has a pink tail). As shown, there was a subtle change in the positions of the exterior place cells to accommodate the novel square shape. Two exterior place cells migrated to interior positions (long red and blue tails) and, correspondingly, two interior place cells migrated to exterior positions (cyan and pink tails). The length of the tails is sometimes misleading owing to the circular nature of the border cells (in truth, the two very long red and blue tails migrated in a wraparound fashion, which was a much smaller change). Finally, one new memory was formed (yellow memory encircled in black).

Figure 13B is a recording of the same simulated animal after switching to a novel square enclosure that contains the same non-spatial attribute and has the same width/height as the familiar circular enclosure. Of note, experiments that switched animals between square and circular enclosures have produced global remapping and, correspondingly, rotation of the grid cell responses (Fyhn et al., 2007). However, to keep things simple in this simulation, it was assumed that the border cells lined up in the same manner for both enclosures (i.e., the *E* dimension was East-West for both enclosures), such as might occur if the animal was primarily

attuned to a salient characteristic of the testing room rather than the enclosure walls. But if the border cells had changed their alignment with the new enclosure (e.g., if the *E* border dimension aligned with the North-South borders), then the place cells would have appeared to undergo global remapping as their positions rotated by 90 degrees and the grid pattern would have also rotated.

As seen in the non-spatial grid cell firing map and in the memory locations at the end of the recording session, the head direction grid module was immediately apparent in this novel environment (unlike the results plotted in ***Figure 9***, for a "blank slate" simulated animal with no prior experience with any similar enclosures). Thus, the immediate appearance of head direction grid modules may reflect memory retrieval of similar enclosures. In brief, because the shape of the novel square enclosure was similar to the circular enclosure (i.e., of similar height and width), and because both enclosures were in the same experimental room and in the same position, and because both enclosure contained the same non-spatial attribute, the exterior place cells that were learned for the circular enclosure were readily adapted to make sense of the square enclosure, and this gave stability to the hexagonal 3D arrangement of the hippocampal place cells. In other words, this situation is analogous to rate remapping results in which previously learned memories from similar enclosures placed in the same global context produce the same place fields, but with different firing rates (Leutgeb et al., 2005).

The current simulation did not actually produce a substantial change in firing rates of the place cells considering that the non-spatial context of the situation was assumed to be identical between the circular enclosure and the change to the square enclosure. In this regard, this is more accurately described as partial remapping (nearly all place fields were unaffected) rather than rate remapping. Despite the adaptation of memories from the circular enclosure to the novel square enclosure, there were some changes. More specifically, 12 of the 14 interior place cells remained in their same positions. However, two of the previously interior place cells moved to the exterior (the cyan and pink tails) and, correspondingly, two of the previously exterior place cells moved to the interior, filling the now vacant memory locations (red and blue tails, which actually took the shorter path wrap-around direction, rather than the long paths shown in the figure). In addition, one new memory was formed (the encircled yellow dot). Another important aspect of this partial remapping is the set of subtle changes in the exterior place cells to accommodate the square shape.

Using the same parameters and same situation of recording in a familiar enclosure before a change to a novel enclosure placed in the same global context and containing the same non-spatial attribute, ***Figure 14*** shows a simulation that is more accurately be described as global remapping (Geva-Sagiv et al., 2016; O'keefe & Nadel, 1979). In this case, the familiar enclosure was a rectangle that was only 25% as tall as the square. As seen in ***Figure 14A***, the exterior place cells conformed to the shape of the rectangle. When moved to the square enclosure (**Figure 14B**), this gave the animal access to previously unexplored regions within the global context (the regions above the North wall of the rectangular enclosure and below the South wall of the rectangular enclosure). As a result of learning and consolidation for these previously unexplored regions, the exterior place cells above and below the remembered enclosure were pushed a substantial distance to accommodate the larger square shape. Furthermore, 4 new interior memories were created (as well as one new exterior memory) and some of the interior memories changed from one interior position to a different interior position. By the end of the recording in the novel square enclosure, the memories formed into a 6-layer face-centered head direction grid module, demonstrating that even with global remapping, memories from other enclosures can be

adapted to provide a rapid understanding of a novel enclosure. However, in this case, this occurred in combination with a greater degree of new memory formation and more radical changes in the positions of place cells than the remapping that occurred with the change from a circular enclosure to a square enclosure.

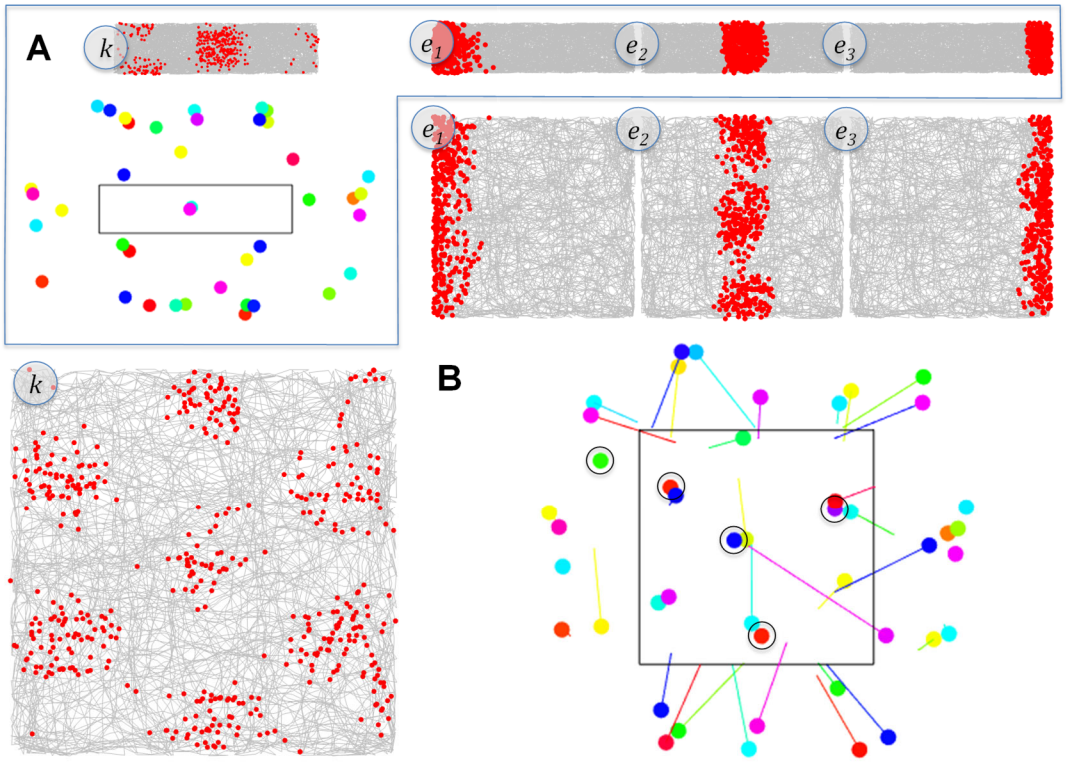

Figure 14. Simulation results for a familiar rectangular enclosure followed by a change to a square enclosure placed in the same context, containing the same non-spatial attribute. (**A**) As with the narrow passage in ***Figure 12***, the familiar rectangular enclosure was too narrow to produce a highly regular hexagonal grid. (**B**) In switching to the much larger square enclosure, most place cells changed their locations (global remapping) and 5 new memories were formed. Despite the considerable disruption in the place fields, the memories from the rectangular enclosure allowed for the rapid establishment of the hexagonal grid, including head direction conjunctive grid cells (see ***Figure 13*** caption for additional information).

Finally, one notable behavior of the border cells in ***Figure 14*** is that East/West border cells appeared to elongate their responses when the East and West borders were elongated to create the novel square enclosure. Exactly this result has been reported in the literature (Solstad et al., 2008), and in the model it occurs because the border cells are more accurately described as an allocentric playing field in which different enclosures can be represented, with a tendency for the playing field to align with borders. In other words, the border cells define a toroidal space (a donut) on which knowledge for the enclosure can be created (e.g., placing memory "sprinkles" on the donut in the shape of the enclosure).

Discussion

These simulation results demonstrate that a memory model can explain many of the results in the rodent electrophysiology literature that were previously assumed to reflect the operations of an entorhinal/hippocampal navigation system. The ability of the model to explain these results raises the possibility that place cells are in fact memory cells, representing conjunctions of *what* and *where* and that grid cells are in fact non-spatial *what* cells, with select locations triggering retrieval for the non-spatial attributes (i.e., remembering the non-spatial attributes that exist at a particular location). Because there are many different non-spatial attributes that contribute to hippocampal memories, only a small proportion of the cells in entorhinal cortex are truly spatial on this account. Specifically, the true spatial cells are the head direction and border cells. The grid cells only appear to be spatial because their activity was recorded while the animal navigated an open field enclosure in which the non-spatial attributes were the same at all locations (i.e., navigation in an enclosure devoid of salient landmarks). This uniformity of the non-spatial attributes leads to a situation in which memories become arrayed in a hexagonal grid owing to pattern separation and rapid memory consolidation.

Key Predictions Made by this Memory Model

This account predicts that all grid cells are conjunctive grid cells and the attributes that are conjoined with location are awaiting discovery (see predictions listed in ***Box 1***). Providing examples of this, head direction conjunctive grid cells become head direction cells in the absence of hippocampal memory feedback (Bonnevie et al., 2013) and some garden-variety grid cells become sound-frequency cells when the animal performs a sound-frequency task (Aronov et al., 2017). Another unique prediction of this memory model is that place cell fields should be arrayed in a hexagonal grid. This does not mean that *all* place cell fields should be part of the *same* grid. Instead, the key prediction is that *sets* of place cells should be arrayed in a grid *if the attribute that defines the set of place cell memories is found uniformly throughout a previously explored two-dimensional surface*. A set of place cells is defined as place cells that represent the same attribute in conjunction with location information (e.g., all places with a particular odor). The attribute that links a set of place cells could be spatial, such as with head direction (i.e., a remembered viewpoint at all locations), and one way to test this idea would be to find sets of place cells that are sensitive to the same head direction, with a test of whether their corresponding place fields are arrayed in a grid. But more generally it will be difficult to identify sets of place cells linked by a non-spatial attribute considering that non-spatial attributes are not typically manipulated during a navigation experiment (more to the point, the set of place cells is predicted to be arrayed in a grid precisely because the non-spatial attribute is *not* manipulated). One approach for testing this prediction would be to use a particular grid cell to reference the non-spatial attribute (i.e., identify place cells that are also active when a particular grid cell is active). Once identified, the corresponding set of place cells should have their place centers arrayed in a hexagonal grid. In other words, there should be a one-to-one correspondence between each grid field of a particular grid cell and a corresponding place field.

The one-to-one correspondence between each grid field and a corresponding place field arises in the model from hippocampal feedback. In other words, the place fields are the cause of the grid fields. However, the model is bidirectional and recurrent. For instance, memory consolidation of the place cells arises from cycles of bottom-up and top-down activation between

entorhinal cells and hippocampal place cells. As such, changes in the properties of the grid cells may entail changes in the properties of place cells. This recurrence might explain the finding that manipulations of grid spacing are associated with an increase in the size of place fields (Mallory et al., 2018). For this memory model, place cell field size directly relates to the stability of place cells. If the place cells are actively moving their place centers with ongoing consolidation (e.g., moving to acquire a larger spacing between memories), they will appear to have larger place fields because the recording session entails a time period in which the place cells are actively changing their preferred locations. In the simulation results, this occurred with a novel enclosure (see for instance the larger grid field sizes, which arise from unstable place cells, for the novel network results in ***Figures 6/7***), and to a lesser extent with remapping (***Figure 13/14***).

This memory model predicts that grid-like responses do not require spatial navigation for their existence. Instead, grid-like responses might occur in any situation involving a highly familiar two-dimensional space in which nothing of interest varies aside from the two dimensions. In such cases, the memories that represent points in the two-dimensional space arrange themselves into an equally spaced hexagonal pattern owing to consolidation and pattern separation. If this pattern separation consolidation process is a general property of cortical neurons, grid-like responses may be found throughout the cortex. This might explain why well-learned non-spatial two-dimensional spaces such as cartoon bird neck/leg length (Constantinescu et al., 2016) and pine/banana odors (Bao et al., 2019) produce grid-like responses in areas other than the MTL. It might also explain why visual search in the two dimensions of the visual field can produce grid-like responses in humans (Julian et al., 2018) and primates (Killian et al., 2012).

This non-spatial, memory account of grid cells places important limitations on when grid-like responses should appear during navigation. If the two-dimensional space is cluttered (inhomogeneous placement of non-spatial attributes) or narrow, there will still be memories and memory retrieval for what exists in different locations (i.e., hippocampal place fields), but the memories will fail to arrange themselves in a precise hexagonal pattern. Furthermore, even when the space is uncluttered and the enclosure is sufficiently large, there needs to be some salient/memorable non-spatial attributes that are found throughout the enclosure.

The dependence of grid cell responses on memory may help explain why grid cells have been found for bats crawling on a two-dimensional surface (Yartsev et al., 2011), but three-dimensional grid cells have never been observed for flying bats. More specifically, two-dimensional surfaces have spatially fixed attributes (e.g., a location of a food source) and it is useful to retrieve these non-spatial attributes when revisiting previously explored positions. However, a location in the three dimensions of air is less likely to contain spatially fixed non-spatial attributes and the spatial cueing of memories may be less useful. Thus, the memory system of the bat may be better configured to cue memory retrieval in relation to surface positions (i.e., the ground) rather than positions in air. Although three-dimensional place fields are found in flying bats (Yartsev & Ulanovsky, 2013), these 3D place cells might reflect memory for the geometry of the environment (i.e., a true place cell, rather than a what/where memory conjunction). This highlights the key hypothesis of this memory theory of grid cells. If grid cells are non-spatial and reflect memory rather than navigation, and if non-spatial attributes are more typically affixed to surfaces rather than arbitrary locations in air, then bats should exhibit two-dimensional grid cells responses but not three-dimensional grid cell responses even though they navigate in three dimensions.

Other Grid Cell Models Assume that Grid Cells are Spatial

The regularity of grid cell responses is fascinating and unparalleled in neuroscience – the existence of the grid pattern assuredly tells us something important about the operation of the nervous system and, correspondingly, there have been dozens of models aimed at explaining how the regular grid pattern emerges. These models are briefly considered to highlight the most important differences between them and this memory model of grid cells. The first wave of grid cell models built the grid cell responses in a purely bottom-up fashion, proposing that place cells learn their positions as defined by a population code of grid cells (Hasselmo, 2009; Mhatre et al., 2012; Solstad et al., 2006). Subsequent work called into question these models (although see Lian & Burkitt, 2022), based on the finding that inactivation of the hippocampus eliminates grid responses (Bonnevie et al., 2013) and the finding that place cells exist before grid cells during development (Bjerknes et al., 2014). To accommodate such results, a second wave of grid cell models assumed that place cells serve as a guide for the development/emergence of grid cell responses (Castro & Aguiar, 2014; Stepanyuk, 2015; Widloski & Fiete, 2014). However, it is less clear why the grid cells are needed in these models, considering that place cells could directly provide spatial knowledge, although some have proposed that grid cells are useful for charting novel navigational paths (Bellmund et al., 2016; Bush et al., 2015; Sorscher et al., 2023). Recently, there are several models that take a broader information-theoretic view of the MTL by asking what kinds of representations are useful for representing different kinds of spaces in a parsimonious manner (Mok & Love, 2019; Rodríguez-Domínguez & Caplan, 2019; Wei et al., 2015), or useful for predictive coding (Stachenfeld et al., 2017). These models can capture some of the grid cell results presented in the current simulations, including extension to non-spatial grid-like responses. However, unlike the present model, these models still assume that entorhinal grid cells represent space rather than a non-spatial attribute.

Conclusions

Returning to the original question posed in this study – how can the function of the MTL be the creation and retrieval of episodic memories if most of the cells in MTL are spatial? The proposed theory and model present a potential answer. The appearance of rodent MTL as primarily concerned with navigation rather than memory may be spurious if grid cells have been misclassified as spatial. Instead, place cells may be “memory cells” that combine spatial location with many other dimensions into complex conjunctions representing episodic memories, and the precise hexagonal firing pattern of grid cells might reflect memory encoding for non-spatial attributes (“what”) that are found throughout the two-dimensional surface.

Acknowledgments

I thank Trygve Solstad and Rosie Cowell for many discussions during the development of this model, Nina Kazanina for providing feedback on a draft of this report, Josh Jacobs for creating an initial version of Figure 1A, and Tim Xia for his work reanalyzing the data of Jercog et al. (2019).

References

Aronov, D., Nevers, R., & Tank, D. W. (2017). Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature*, *543*(7647), 719–722.

Bao, X., Gjorgieva, E., Shanahan, L. K., Howard, J. D., Kahnt, T., & Gottfried, J. A. (2019). Grid-like neural representations support olfactory navigation of a two-dimensional odor space. *Neuron*, *102*(5), 1066–1075.

Barry, C., Ginzberg, L. L., O'Keefe, J., & Burgess, N. (2012). Grid cell firing patterns signal environmental novelty by expansion. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(43), 17687–17692. https://doi.org/DOI 10.1073/pnas.1209918109

Barry, C., Hayman, R., Burgess, N., & Jeffery, K. J. (2007). Experience-dependent rescaling of entorhinal grids. *Nature Neuroscience*, *10*(6), 682–684.

Barry, C., Lever, C., Hayman, R., Hartley, T., Burton, S., O'Keefe, J., Jeffery, K., & Burgess, N. (2006). The boundary vector cell model of place cell firing and spatial memory. *Reviews in the Neurosciences*, *17*(1–2), 71–98.

Bellmund, J. L., Deuker, L., Navarro Schröder, T., & Doeller, C. F. (2016). Grid-cell representations in mental simulation. *Elife*, *5*, e17089.

Bhatia, A., Moza, S., & Bhalla, U. S. (2019). Precise excitation-inhibition balance controls gain and timing in the hippocampus. *Elife*, *8*, e43415.

Bienenstock, E. L., Cooper, L. N., & Munro, P. W. (1982). Theory for the Development of Neuron Selectivity—Orientation Specificity and Binocular Interaction in Visual-Cortex. *Journal of Neuroscience*, *2*(1), 32–48.

Bjerknes, T. L., Moser, E. I., & Moser, M.-B. (2014). Representation of geometric borders in the developing rat. *Neuron*, *82*(1), 71–78.

Boccara, C. N., Sargolini, F., Thoresen, V. H., Solstad, T., Witter, M. P., Moser, E. I., & Moser, M.-B. (2010). Grid cells in pre-and parasubiculum. *Nature Neuroscience*, *13*(8), 987–994.

Bonnevie, T., Dunn, B., Fyhn, M., Hafting, T., Derdikman, D., Kubie, J. L., Roudi, Y., Moser, E. I., & Moser, M. B. (2013). Grid cells require excitatory drive from the hippocampus. *Nat Neurosci*, *16*(3), 309–317. https://doi.org/Doi 10.1038/Nn.3311

Bostock, E., Muller, R. U., & Kubie, J. L. (1991). Experience-dependent modifications of hippocampal place cell firing. *Hippocampus*, *1*(2), 193–205.

Bush, D., Barry, C., Manson, D., & Burgess, N. (2015). Using grid cells for navigation. *Neuron*, *87*(3), 507–520.

Buzsáki, G., & Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature Neuroscience*, *16*(2), 130–138.

Calton, J. L., Stackman, R. W., Goodridge, J. P., Archey, W. B., Dudchenko, P. A., & Taube, J. S. (2003). Hippocampal place cell instability after lesions of the head direction cell network. *Journal of Neuroscience*, *23*(30), 9719–9731.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62.

Castro, L., & Aguiar, P. (2014). A feedforward model for the formation of a grid field where spatial information is provided solely from place cells. *Biological Cybernetics*, *108*, 133–143.

Chanales, A. J., Oza, A., Favila, S. E., & Kuhl, B. A. (2017). Overlap among spatial memories triggers repulsion of hippocampal representations. *Current Biology*, *27*(15), 2307–2317.

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468.

Corkin, S. (1984). *Lasting consequences of bilateral medial temporal lobectomy: Clinical course and experimental findings in HM*. *4*(02), 249–259.

Derdikman, D., Whitlock, J. R., Tsao, A., Fyhn, M., Hafting, T., Moser, M. B., & Moser, E. I. (2009). Fragmentation of grid cell maps in a multicompartment environment. *Nat Neurosci*, *12*(10), 1325-U155. https://doi.org/Doi 10.1038/Nn.2396

Doeller, C. F., Barry, C., & Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature*, *463*(7281), 657–661.

Frank, L. M., Stanley, G. B., & Brown, E. N. (2004). Hippocampal plasticity across multiple days of exposure to novel environments. *Journal of Neuroscience*, *24*(35), 7681–7689.

Fyhn, M., Hafting, T., Treves, A., Moser, M. B., & Moser, E. I. (2007). Hippocampal remapping and grid realignment in entorhinal cortex. *Nature*, *446*(7132), 190–194. https://doi.org/Doi 10.1038/Nature05601

Gardner, R. J., Hermansen, E., Pachitariu, M., Burak, Y., Baas, N. A., Dunn, B. A., Moser, M.-B., & Moser, E. I. (2022). Toroidal topology of population activity in grid cells. *Nature*, *602*(7895), 123–128.

Geva-Sagiv, M., Romani, S., Las, L., & Ulanovsky, N. (2016). Hippocampal global remapping for different sensory modalities in flying bats. *Nature Neuroscience*, *19*(7), 952–958.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*(3), 325–331.

Grieves, R. M., & Jeffery, K. J. (2017). The representation of space in the brain. *Behavioural Processes*, *135*, 113–131.

Grossberg, S., & Grossberg, S. (1982). How does a brain build a cognitive code? *Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition, and Motor Control*, 1–52.

Hafting, T., Fyhn, M., Molden, S., Moser, M. B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, *436*(7052), 801–806. https://doi.org/Doi 10.1038/Nature03721

Hales, T. C. (2005). A proof of the Kepler conjecture. *Annals of Mathematics*, 1065–1185.

Hardcastle, K., Ganguli, S., & Giocomo, L. M. (2015). Environmental boundaries as an error correction mechanism for grid cells. *Neuron*, *86*(3), 827–839.

Hasselmo, M. E. (2009). A model of episodic memory: Mental time travel along encoded trajectories using grid cells. *Neurobiology of Learning and Memory*, *92*(4), 559–573.

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory*. Psychology press.

Hinton, G. E., & Sejnowski, T. J. (1986). Learning and relearning in Boltzmann machines. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, *1*(282–317), 2.

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, *112*(1), 75.

Høydal, Ø. A., Skytøen, E. R., Andersson, S. O., Moser, M.-B., & Moser, E. I. (2019). Object-vector coding in the medial entorhinal cortex. *Nature*, *568*(7752), 400–404.

Jacobs, J., Weidemann, C. T., Miller, J. F., Solway, A., Burke, J. F., Wei, X.-X., Suthana, N., Sperling, M. R., Sharan, A. D., & Fried, I. (2013). Direct recordings of grid-like neuronal activity in human spatial navigation. *Nature Neuroscience*, *16*(9), 1188–1190.

Jercog, P. E., Ahmadian, Y., Woodruff, C., Deb-Sen, R., Abbott, L. F., & Kandel, E. R. (2019). Heading direction with respect to a reference point modulates place-cell activity. *Nature Communications*, *10*(1), 2333.

Julian, J. B., Keinath, A. T., Frazzetta, G., & Epstein, R. A. (2018). Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nature Neuroscience*, *21*(2), 191–194.

Killian, N. J., Jutras, M. J., & Buffalo, E. A. (2012). A map of visual space in the primate entorhinal cortex. *Nature*, *491*(7426), 761–764. https://doi.org/Doi 10.1038/Nature11587

Krupic, J., Bauza, M., Burton, S., Barry, C., & O'Keefe, J. (2015). Grid cell symmetry is shaped by environmental geometry. *Nature*, *518*(7538), 232–235.

Kruschke, J. K. (1992). ALCOVE: an exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22.

Leutgeb, S., Leutgeb, J. K., Barnes, C. A., Moser, E. I., McNaughton, B. L., & Moser, M.-B. (2005). Independent codes for spatial and episodic memory in hippocampal neuronal ensembles. *Science*, *309*(5734), 619–623.

Lever, C., Burton, S., Jeewajee, A., O'Keefe, J., & Burgess, N. (2009). Boundary vector cells in the subiculum of the hippocampal formation. *Journal of Neuroscience*, *29*(31), 9771–9777.

Lian, Y., & Burkitt, A. N. (2022). Learning Spatiotemporal Properties of Hippocampal Place Cells. *Eneuro*, *9*(4).

Mallory, C. S., Hardcastle, K., Bant, J. S., & Giocomo, L. M. (2018). Grid scale drives the scale and long-term stability of place maps. *Nature Neuroscience*, *21*(2), 270–282.

Marr, D., Willshaw, D., & McNaughton, B. (1991). *Simple memory: A theory for archicortex*. Springer.

Mathis, A., Stemmler, M. B., & Herz, A. V. (2015). Probable nature of higher-dimensional symmetries underlying mammalian grid-cell activity patterns. *Elife*, *4*, e05979.

Maurer, A. P., & Nadel, L. (2021). The continuity of context: A role for the hippocampus. *Trends in Cognitive Sciences*, *25*(3), 187–199.

Mauro, R., & Kubovy, M. (1992). Caricature and face recognition. *Memory & Cognition*, *20*, 433–440.

McNaughton, B. L., Barnes, C. A., & O'Keefe, J. (1983). The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. *Experimental Brain Research*, *52*, 41–49.

Mhatre, H., Gorchetchnikov, A., & Grossberg, S. (2012). Grid cell hexagonal patterns formed by fast self-organized learning within entorhinal cortex. *Hippocampus*, *22*(2), 320–334.

Milivojevic, B., & Doeller, C. F. (2013). Mnemonic networks in the hippocampal formation: From spatial maps to temporal and conceptual codes. *Journal of Experimental Psychology: General*, *142*(4), 1231.

Mok, R. M., & Love, B. C. (2019). A non-spatial account of place and grid cells based on clustering models of concept learning. *Nature Communications*, *10*(1), 5685.

Moser, E. I., Kropff, E., & Moser, M.-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, *31*, 69–89.

Muller, R. U., Bostock, E., Taube, J. S., & Kubie, J. L. (1994). On the directional firing properties of hippocampal place cells. *Journal of Neuroscience*, *14*(12), 7235–7251.
O'Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Experimental Neurology*, *51*(1), 78–109.
O'Keefe, J., & Krupic, J. (2021). Do hippocampal pyramidal cells respond to nonspatial stimuli? *Physiological Reviews*, *101*(3), 1427–1456.
O'keefe, J., & Nadel, L. (1979). Précis of O'Keefe & Nadel's The hippocampus as a cognitive map. *Behavioral and Brain Sciences*, *2*(4), 487–494.
Ólafsdóttir, H. F., Bush, D., & Barry, C. (2018). The role of hippocampal replay in memory and planning. *Current Biology*, *28*(1), R37–R50.
Ólafsdóttir, H. F., Carpenter, F., & Barry, C. (2016). Coordinated grid and place cell replay during rest. *Nature Neuroscience*, *19*(6), 792–794.
Olsen, S. R., Bhandawat, V., & Wilson, R. I. (2010). Divisive normalization in olfactory population codes. *Neuron*, *66*(2), 287–299.
O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8*(5), 895–938.
O'Reilly, R. C., & McClelland, J. L. (1994). Hippocampal conjunctive encoding, storage, and recall: Avoiding a trade-off. *Hippocampus*, *4*(6), 661–682.
O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. MIT Press.
Peyrache, A., Schieferstein, N., & Buzsáki, G. (2017). *Transformation of the head-direction signal into a spatial code. Nat Commun 8, 1752*.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. *Classical Conditioning, Current Research and Theory*, *2*, 64–69.
Rodríguez-Domínguez, U., & Caplan, J. B. (2019). A hexagonal Fourier model of grid cells. *Hippocampus*, *29*(1), 37–45.
Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(5), 558.
Rubin, A., Yartsev, M. M., & Ulanovsky, N. (2014). Encoding of head direction by hippocampal place cells in bats. *Journal of Neuroscience*, *34*(3), 1067–1080.
Sanders, H., Wilson, M. A., & Gershman, S. J. (2020). Hippocampal remapping as hidden state inference. *Elife*, *9*, e51140.
Sargolini, F., Fyhn, M., Hafting, T., McNaughton, B. L., Witter, M. P., Moser, M. B., & Moser, E. I. (2006). Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science*, *312*(5774), 758–762. https://doi.org/DOI 10.1126/science.1125572
Savelli, F., Luck, J., & Knierim, J. J. (2017). Framing of grid cells within and beyond navigation boundaries. *Elife*, *6*, e21354.
Savelli, F., Yoganarasimha, D., & Knierim, J. J. (2008). Influence of boundary removal on the spatial representations of the medial entorhinal cortex. *Hippocampus*, *18*(12), 1270–1282.
Solomon, S. G., & Lennie, P. (2007). The machinery of colour vision. *Nature Reviews Neuroscience*, *8*(4), 276–286.
Solstad, T., Boccara, C. N., Kropff, E., Moser, M. B., & Moser, E. I. (2008). Representation of Geometric Borders in the Entorhinal Cortex. *Science*, *322*(5909), 1865–1868. https://doi.org/DOI 10.1126/science.1166466

Solstad, T., Moser, E. I., & Einevoll, G. T. (2006). From grid cells to place cells: A mathematical model. *Hippocampus*, *16*(12), 1026–1031. https://doi.org/Doi 10.1002/Hipo.20244

Sorscher, B., Mel, G. C., Ocko, S. A., Giocomo, L. M., & Ganguli, S. (2023). A unified theory for the computational and mechanistic origins of grid cells. *Neuron*, *111*(1), 121–137.

Squire, L. R., & Zola-Morgan, S. (1991). The medial temporal lobe memory system. *Science*, *253*(5026), 1380–1386.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, *15*(1), 1929–1958.

Stachenfeld, K. L., Botvinick, M. M., & Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, *20*(11), 1643–1653.

Stella, F., & Treves, A. (2015). The self-organization of grid cells in 3D. *Elife*, *4*, e05913.

Stensola, H., Stensola, T., Solstad, T., Froland, K., Moser, M. B., & Moser, E. I. (2012). The entorhinal grid map is discretized. *Nature*, *492*(7427), 72–78. https://doi.org/Doi 10.1038/Nature11649

Stepanyuk, A. (2015). Self-organization of grid fields under supervision of place cells in a neuron model with associative plasticity. *Biologically Inspired Cognitive Architectures*, *13*, 48–62.

Sutherland, R. J., & Rudy, J. W. (1989). Configural association theory: The role of the hippocampal formation in learning, memory, and amnesia. *Psychobiology*, *17*(2), 129–144.

Taube, J. S., Muller, R. U., & Ranck, J. B. (1990). Head-Direction Cells Recorded from the Postsubiculum in Freely Moving Rats .1. Description and Quantitative-Analysis. *Journal of Neuroscience*, *10*(2), 420–435.

Teyler, T. J., & Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: Updating the index. *Hippocampus*, *17*(12), 1158–1169.

Tulving, E., Markowitsch, H. J., Craik, F. I., Habib, R., & Houle, S. (1996). Novelty and familiarity activations in PET studies of memory encoding and retrieval. *Cerebral Cortex*, *6*(1), 71–79.

Waaga, T., Agmon, H., Normand, V. A., Nagelhus, A., Gardner, R. J., Moser, M.-B., Moser, E. I., & Burak, Y. (2022). Grid-cell modules remain coordinated when neural activity is dissociated from external sensory cues. *Neuron*, *110*(11), 1843–1856.

Wei, X.-X., Prentice, J., & Balasubramanian, V. (2015). A principle of economy predicts the functional architecture of grid cells. *Elife*, *4*, e08362.

Widloski, J., & Fiete, I. R. (2014). A model of grid cell development through spatial exploration and spike time-dependent plasticity. *Neuron*, *83*(2), 481–495.

Widrow, B., & Hoff, M. E. (1960). *Adaptive switching circuits*. Stanford Univ Ca Stanford Electronics Labs.

Wills, T. J., Lever, C., Cacucci, F., Burgess, N., & O'Keefe, J. (2005). Attractor dynamics in the hippocampal representation of the local environment. *Science*, *308*(5723), 873–876.

Winter, S. S., Clark, B. J., & Taube, J. S. (2015). Disruption of the head direction cell network impairs the parahippocampal grid cell signal. *Science*, *347*(6224), 870–874.

Yartsev, M. M., & Ulanovsky, N. (2013). Representation of three-dimensional space in the hippocampus of flying bats. *Science*, *340*(6130), 367–372.

Yartsev, M. M., Witter, M. P., & Ulanovsky, N. (2011). Grid cells without theta oscillations in the entorhinal cortex of bats. *Nature*, *479*(7371), 103–107. https://doi.org/Doi 10.1038/Nature10583

Yoon, K., Buice, M. A., Barry, C., Hayman, R., Burgess, N., & Fiete, I. R. (2013). Specific evidence of low-dimensional continuous attractor dynamics in grid cells. *Nature Neuroscience*, *16*(8), 1077–1084.