

# Navigating Cognition: Spatial Codes for Human Thinking

Jacob L.S. Bellmund<sup>1,2,3\*</sup>, Peter Gärdenfors<sup>4,5</sup>, Edvard I. Moser<sup>1</sup>, Christian F. Doeller<sup>1,3\*</sup>

1: Kavli Institute for Systems Neuroscience, Centre for Neural Computation, The Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits, NTNU, Norwegian University of Science and Technology, Trondheim, Norway

2: Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands

3: Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

4: Department of Philosophy and Cognitive Science, Lund University, Lund, Sweden

5: Centre for Artificial Intelligence, University of Technology Sydney, Sydney, Australia

\* Corresponding authors. Email: bellmund@cbs.mpg.de; doeller@cbs.mpg.de

The hippocampal formation has long been suggested to underlie both memory formation and spatial navigation. Here, we discuss how neural mechanisms identified in spatial navigation research generalize across information domains to support a wide spectrum of cognitive functions. In our framework, place and grid cell population codes provide a representational format to map variable dimensions of cognitive spaces. This highly dynamic mapping system enables rapid reorganization of codes through remapping between orthogonal representations across behavioral contexts, yielding a multitude of stable cognitive spaces at different resolutions and hierarchical levels. Action sequences result in trajectories through cognitive space, which can be simulated via sequential coding in the hippocampus. Thereby, the spatial representational format of the hippocampal formation has the capacity to support flexible cognition and behavior.

## One sentence summary

We lay out how spatial processing principles of the hippocampal-entorhinal system may provide a geometric code for high-level cognition.

## Introduction

How past experience guides behavior has been a highly contested topic for decades. In 1948, Tolman (1) described evidence for learning beyond stimulus-response driven behavior in rats, foreshadowing the cognitive revolution to follow. He proposed that rats learn comprehensive maps of their environments, which can guide flexible goal-directed behavior such as finding shortcuts. Tolman coined the term *cognitive map* and speculated how these maps might underlie psychological functions. Support for his controversial speculations was provided by the first reports of hippocampal place cells more than 20 years later (2). Place cells are generally only active when an animal occupies the cell's preferred location. These cells were proposed by O'Keefe and Nadel (3) as a potential neural substrate of Tolman's cognitive map. Incorporating the discovery of place cells and findings from lesion studies, the hippocampus was suggested to provide a map-like reference system, signaling both the position of the agent and features of the environment for navigational and mnemonic processing of events (3). The sensitivity of hippocampal coding to items or discrete events in the environment (4–6), differential coding of overlapping path segments based on task state (7) and the widespread deficits in patients with hippocampal lesions (8) have been captured in the development of relational memory theory (8–10). This account builds on the capabilities of the hippocampus for associative binding to link event representations into relational networks and has offered a counterpole to an exclusively spatial processing view of the hippocampus. Episodic memories can be formed by linking successive event representations and episodic trajectories into ensemble patterns stored in hippocampal networks for subsequent retrieval (8, 10–14). Recent advances demonstrate a hippocampal involvement in flexible cognition beyond the domains of navigation and memory (15).

## Space codes as a representational format for cognition

We here describe domain-general core coding principles from spatial navigation research that have the potential to support a wider span of cognitive functions. Specifically, we propose that the hippocampal-entorhinal system represents experience in cognitive spaces (cf. conceptual spaces in (16, 17)). A cognitive space is thought to be spanned by a set of quality dimensions, which can be closely related to sensory inputs but also comprise abstract features (16). A given stimulus can be located in a cognitive space based on its set of feature values along a set of quality dimensions. Each dimension is equipped with an underlying metric and follows geometric constraints satisfying the mathematical notions of betweenness and equidistance (16). Consider the simple example that when planning to buy a new car, you might describe cars along two dimensions, their engine power and their weight (Fig. 1). Depending on the two features, vehicles can be located in this two-dimensional space, for instance sports car might be rather lightweight, but have a strong engine. In this example, we might treat the dimensions as separable. However, stimuli can also vary along integral dimensions, on which a given stimulus cannot be described independently. For example, a car's color can only be fully described when assigning values on all three integral dimensions hue, saturation and brightness, which constitute the color domain (16). Different cognitive spaces might have different underlying metrics, for example the Euclidean or the city-block metric, which allow assigning distances along a dimension reflecting the degree of similarity between locations in cognitive space with similar stimuli located close together (16, 18).

Building on the geometric characterization of quality dimensions, a property is defined as a convex region in some domain with convexity meaning that for all points  $x$  and  $y$  in the region the points between  $x$  and  $y$  also fall into the region (16). In this framework, a property constitutes the simplest form of a concept, based on only one domain, for example describing a car as ‘heavy’. Defining properties as convex regions enables generalization in that a property of two stimuli  $x$  and  $y$  can be inferred to be shared by any stimulus  $z$  falling in between  $x$  and  $y$ . Therefore, the geometric constraints on cognitive spaces allow inference about never-experienced stimuli. Thereby cognitive spaces afford great cognitive flexibility; going beyond associative and transitive inferences also permitted by relational networks (8–10). More complex concepts comprise multiple domains and information about their interrelations. Thus, a concept is defined as a set of convex regions in a number of domains with domains weighted based on salience and additional information about how the regions in different domains are correlated (16). In our example, racing cars would occupy a region characterized by high power and low weight (Fig. 1). Following the spatial definition of concepts, different exemplars correspond to more or less central positions, with prototypical members (19) located centrally in the conceptual region. Using a Voronoi tessellation of a continuous space with a Euclidean metric, where all locations in a given space are discretized as belonging to the nearest prototype, convex regions emerge and allow classification of individual stimuli (16).

We propose that *processing mechanisms in the hippocampal-entorhinal system are well-suited to support cognitive spaces as a domain-general format for flexible, high-level cognition in humans*. Studies of spatial navigation describe how space is mapped by a continuous, multi-scale code of function-specific cells in the hippocampal-entorhinal system. This system enables flexible mapping of different environments and simulations of trajectories through space via temporally compressed sequences. These processing mechanisms might have developed from originally mapping navigable space to also representing cognitive spaces; consistent with an evolutionary preserved circuitry of the hippocampal-entorhinal region across mammals (20), the translation of insights from spatial navigation research in rodents to human navigation and beyond (3, 21, 22) (see Box), and the notion of a core system of geometry (23).

### **From spatial navigation to cognitive spaces**

In the mammalian brain, positional information is conveyed by the spatially constrained firing of place (2) and grid (24) cells during spatial navigation. Place cells in the hippocampus are preferentially active when the animal occupies a certain position within the environment, the cell’s receptive field or place field. The firing fields of the population of place cells are thought to cover the entire environment, thereby providing a map-like representation of the animal’s surroundings (2, 3). While the firing of a place cell is usually restricted to one place in a small environment, grid cells in the medial entorhinal cortex (EC), one synapse from the hippocampus, exhibit multiple firing fields located at the vertices of equilateral triangles tiling the entire environment (24). These regular, six-fold symmetric firing patterns are assumed to support spatial navigation by providing a coordinate system of the environment (24, 25). The entorhinal grid system is assigned a key role in path integration and vector navigation (25–28).

Intracranial recordings in patients navigating virtual reality environments established the existence of place (29) and grid (30) cells in humans. The six-fold symmetry of grid-cell firing has been

translated to non-invasive functional magnetic resonance imaging (fMRI) in healthy volunteers, where grid-like hexadirectional signals have been observed in EC during navigation (22, 31–37) (see Box). The brain's spatial navigation system (25, 38) further includes head direction cells conveying information about the animal's head direction (39), goal and goal direction cells signaling egocentric directions to navigational goals (40, 41) as well as speed cells sensitive to running speed (42) and border (43, 44) or boundary vector cells (45) responding to borders in the environment.

The firing of place and grid cells conveys positional information to navigate Euclidean space. Here, we hypothesize that the spatially constrained firing of place cells and the metric provided by the entorhinal grid system might provide a domain-general mechanism to map dimensions of experience. In this framework, the activity of place cells can be conceived as indexing locations in a cognitive space spanned by the entorhinal grid system. Three further neural coding mechanisms identified in spatial navigation studies illustrate how the hippocampal-entorhinal system may support a core mechanism of mapping cognitive spaces. First, the firing fields of place and grid cells increase in size along the dorsoventral axis of the rodent hippocampus (24, 46–49), in line with mapping of cognitive spaces at different levels of granularity for multi-scale representations of knowledge hierarchies or nested conceptual information. Second, the ability of place cells to undergo global remapping (50–52) allows the flexible formation of a multitude of uncorrelated maps for different cognitive spaces, which can be reinstated via attractor dynamics (11, 13, 53, 54). Third, sequential activity of place (55–57) and grid cells during replay (58–60) and theta oscillations (61) enables the simulation of trajectories (12, 62, 63) through different locations in a cognitive space for adaptive cognition and behavior.

### **A continuous map of experience**

The spatially constrained firing of place and grid cells provides a continuous code for the dimensions of space, in which neighboring positions have similar representations due to partially overlapping firing fields across the population of cells (Fig. 2AB). We build on findings that the continuous code of spatially selective cells maps dimensions of experience beyond Euclidean space, which affords flexible cognition via the formation of cognitive spaces, whose dimensions are geometrically constrained as described above.

Contrary to rodents, the visual system has evolved to be the dominant source of sensory information in primates. During visual exploration of naturalistic images, neurons in primate EC encode gaze position with six-fold symmetric firing patterns that are the hallmark of grid cell firing during navigation (64). Further, some entorhinal neurons preferentially discharged when monkeys explored the edges of the visual stimuli (64), akin to the firing of border cells in rodents (43–45). Others were tuned to saccade direction (65), putatively signaling similar information as head direction cells (39). Together with two recent reports of eye-movement dependent hexadirectional signals in human EC (35, 36) these findings support the notion that representations of visual space follow coding principles identified in spatial navigation research in rodents, suggesting a shared neural basis.

Next to its role in parsing sensory information, this spatial mapping mechanism might also encode a dimension inherent to all experience: time. During space-clamped running throughout a temporal delay, so-called time cells preferentially fire at specific time points (66, 67). The populations of these time cells in the hippocampus and EC overlap substantially with the population of place and grid

cells, respectively, suggesting that cells in these regions might exhibit mixed selectivity for space and time (68, 69). While these hippocampal time cells map repeated intervals in the range of seconds, it remains elusive how time might be discriminated at larger scales. For this, an ongoing temporal signal slowly drifting across hours and days might be transmitted to the hippocampus from the lateral EC, enabling the tagging of time and order to events for mnemonic processing (70).

The mapping of dimensions by the hippocampal-entorhinal system might be a general mechanism to code for task-relevant information. In a sound modulation task, rats manipulated the frequency of an auditory stimulus to 'navigate' to a target frequency (71). Cells in the hippocampus and EC exhibited discrete firing fields with elevated firing constrained to each cell's preferred frequency range, thereby mapping the continuous dimension of sound frequency so that neighboring frequencies exhibited similar firing patterns across cells (Fig. 2C). A subset of the cells carrying frequency-specific information in the sound modulation task could be characterized as place and grid cells during random foraging (71). Hippocampal representations of continuously changing odor concentrations suggest that the directional selectivity of place cells on a linear track might be preserved when mapping a non-spatial dimension (72). This suggests a flexible recruitment of cells to map task-relevant dimensions of experience while maintaining spatial coding properties.

However, we are constantly faced with complex stimuli defined by not one but multiple feature dimensions. Spatial coding has been demonstrated for stimuli varying along two independent dimensions. In one recent study, human participants learned to associate target objects with pictures of stick-figure birds (Fig. 2D) differing in the length of their necks and legs (32). While undergoing fMRI, participants watched trajectories through this two-dimensional feature space and BOLD-signals were analyzed as a function of the angular orientation of these trajectories following the analysis developed to investigate grid coding during virtual navigation (see Box; Doeller et al. (22)). Hexadirectional signals were observed in EC as well as a network of regions (32) involved in human autobiographical memory (63, 73, 74) and exhibiting hexadirectional signals during navigation (22). Within participants, the orientation of the hexadirectional signal was consistent across frontal and medial temporal brain areas and stable over separate sessions more than a week apart (32), suggesting a role of the entorhinal grid system in storage and retrieval of consolidated conceptual knowledge.

In our framework, place and grid cells not only map Euclidean space for navigation, but also map cognitive spaces spanned by relevant feature dimensions (Fig. 2A). Building upon the geometric description of cognitive spaces, this mapping provides an account for how multi-dimensional spaces of experience can be instantiated in the hippocampal formation. If stimuli are located in the space based on their characteristic values on the feature axes, place cell firing might encode stimuli at specific locations in cognitive space based on the respective values along the feature dimensions (16). Similar stimuli are located at nearby positions, whereas dissimilar stimuli might be located further apart in cognitive space. Distances between positions are reflected in population vectors of place cell activity, which will be more similar for nearby positions due to overlapping firing fields of cell ensembles. Therefore, the mapping of cognitive spaces by cell populations in the hippocampal-entorhinal region provides a mechanism to generate similarity between stimuli (16), a concept central to generalization and planning (75). Applying this to the introductory example of a two-dimensional space of engine power and weight, positions encoded by constrained firing of place cells reflect specific combinations of values along the dimensions, for example high engine power

and low weight. This might fall in a region representing the concept of racing cars, which would typically be characterized by this feature combination. The convexity of conceptual regions enables the generalization that a car whose engine power and weight fall between those of two known sports cars also belongs to this category (16). In this framework, the regular firing patterns of grid cells might provide a metric (24, 26–28) for the dimensions of cognitive spaces. This would allow not only for the encoding of positions but also their relations in a similar fashion as proposed for spatial navigation and memory (21, 25). Thereby, the grid system might provide the coordinate system for the dimensions of cognitive spaces, in line with its suggested role in extracting dimensions from task states (76, 77). However, the precise nature of the metric remains elusive as different metrics can underlie cognitive spaces (16) and the impact of deviations from regular hexagonal firing patterns (47, 78, 79) is unclear.

Neural recordings in patients suggest the existence of cells in the hippocampus and EC that selectively respond to a narrow range of stimuli (80). Despite the limitations imposed by the necessity of stimulus selection, the claim that this phenomenon reflects coding of conceptual entities such as famous people or places is supported by the preservation of responses across stimulus modalities. Even though there are no experimentally defined feature axes in these studies, these cells sometimes respond similarly to entities that appear nearby in conceptual space. For example, a cell responding to pictures of the Tower of Pisa also exhibited increased firing in response to the Eiffel Tower (80). This is in line with the assumption that proximal positions in conceptual space are represented by overlapping cell assemblies, comparable to the representations of nearby locations in space by overlapping place cell populations.

The actions and positions of other agents in space are central to interacting with and learning from conspecifics (81). Hippocampal place cells in bats and rodents of an observer animal encode the location of a conspecific or a moving object, indicating the tracking of relevant agents in the environment (82, 83). The human hippocampus also signals the position of others along interpersonal dimensions. In one study, participants were asked to respond to statements of fictitious characters located in a two-dimensional space of power and affiliation (84). Hippocampal activity tracked the position of the counterpart in this social space at times of interaction, exhibiting greater activity during interaction with counterparts with higher power and affiliation (84). Hippocampal encoding of the power dimension has also been demonstrated after learning of social hierarchies (85). These results suggest a hippocampal involvement in the representation of others in both navigable and interpersonal spaces.

It is thus conceivable that coding of fundamental dimensions of experience also underlies the involvement of the hippocampal-entorhinal system in episodic memory, where individual episodes are considered to be embedded in a spatio-temporal context (14). As place and grid cells carry information about both space and time they might thereby signal the context in which events are experienced. Evidence for this stems from a study in which participants encountered objects along a fixed route through a virtual city. Through learning, the similarity structure of multi-voxel activity patterns in the hippocampus changed to reflect remembered spatial and temporal distances between object pairs (86). This is in line with the notion that the hippocampus encodes the dimensions of space and time along which episodic experiences are organized and its role in binding stimuli or events to a context in service of episodic memory, a core function of the hippocampal formation (8, 10, 87).

In conclusion, we suggest that the firing of functionally defined cell types in the hippocampal-entorhinal system prevails across task-relevant dimensions to map dimensions of experience in cognitive spaces. Stimuli are arranged in a spatial format where similarity between positions is reflected in the distance along the dimensions spanning the cognitive space. The representation of cognitive spaces allows not only associative or transitive inference, accounted for via overlapping relational networks in the realm of relational memory theory (8–10), but also generalization and inference to novel stimuli and situations. Below we describe how key coding principles of the hippocampal-entorhinal system make it an ideal candidate for entertaining cognitive spaces.

### **Multiple scales of coding**

One hallmark of abstract knowledge is the representation of information at different hierarchical levels (88, 89). For example, you might identify a particular vehicle characterized by low weight and high engine power as a Porsche (Fig. 3A). On a more general level, you might classify it as a sports car allowing you to infer some properties such as high driving speed, whereas on a more specific level you might wonder about the particular model and its associated characteristics. A likely mechanism for learning and representing information at different scales (88) is described in studies investigating the response properties of place and grid cells along the dorsoventral axis of the rodent hippocampal formation (Fig. 3B) (24, 46–49). On an 18m long linear track, the width of hippocampal place cell firing fields increased from less than 1m in the dorsal to approximately 10m in the ventral hippocampus (48). In human fMRI, this might be reflected in a voxel-similarity gradient along the hippocampal long-axis (90). Similarly, the scale of entorhinal grid cells increases from the dorsal to the ventral parts of the medial EC, reflecting larger firing fields and larger spacing between fields (24, 47, 49, 91). In contrast to place cell firing field width (48), grid scale changes in discrete steps between modules of cells sharing a similar scale and orientation (24, 47, 49, 78).

Different scales of information represented at distinct anatomical locations of the hippocampal formation might serve as a general mechanism across different stimulus domains. Positional decoding in cognitive spaces might benefit from the combination of multiple scales of information analog to navigable space (26, 27). The spacing of firing fields of grid cells responding to locations in visual space increases with anatomical distance to the rhinal sulcus (64), which approximately reflects the anterior-posterior axis of the primate hippocampus corresponding to the dorsal-ventral axes of the rodent hippocampus (92). This is further paralleled by broader tuning of saccade direction cells as anatomical distance to the rhinal sulcus increases (65).

Multiple scales of coding along the dorsoventral axis of the rodent hippocampus are in line with the scale increase of mnemonic networks represented along the anterior-posterior axis of the human hippocampus (Fig. 3C) in participants watching videos of life-like events forming narratives (93, 94). The scale at which these mnemonic networks were represented differed across the hippocampus with the posterior portion representing associations of the most recently linked pair of events, while the mid-portion held information about multiple event pairs and patterns in the anterior hippocampus were indicative of integrated networks of all events in a narrative (93). Multi-scale event representations, interacting with mnemonic processing in the hippocampus, have been dissociated along the cortical hierarchy in humans (95). Representing cognitive spaces at different scales allows for the generalization of specific experience and the formation of contextual codes via more global representations. In rodents performing a context-dependent object discrimination task,

ventral hippocampal neurons exhibited responses generalizing across events within a spatial context while strongly distinguishing between contexts after extended learning, whereas dorsal hippocampal neurons discriminated events within the same context with activity patterns reflecting the hierarchical task structure (96, 97). The representation of integrated codes for overlapping memories might further enable inferential reasoning about related memories. In humans, transitive inference and generalization are supported by anterior portions of the hippocampus (98), whereas more posterior portions are associated with the retention of original memories and element segregation when associations overlap (99, 100). These findings suggest a mnemonic gradient along the hippocampal long-axis in humans paralleling the differences in granularity of spatial representations along the dorsoventral axis of the rodent hippocampus. The most detailed representation might allow for fine-grained discrimination of locations in cognitive space, whereas the representation of larger areas might enable inference and the generalization of behavior to never-experienced stimuli and situations, not limited by the need for associations between nodes in relational mnemonic networks (8–10).

Representing knowledge at different levels of granularity in cognitive spaces requires a cognitive code signaling positions in this space at different resolutions. Generalizing from the above findings to conceptual information, it appears plausible that the gradient of granularity plays an important role in learning and representing hierarchical knowledge structures (88, 89). In these, overarching categories can be conceived as larger areas of cognitive space, putatively represented by place and grid cells with larger firing fields. Subcategories at a finer scale would then correspond to locations nested within these larger areas. Such a nested representation of cognitive spaces could enable inference via the transfer of knowledge from the superordinate category to new exemplars or subcategories (16).

### **Flexible formation of stable cognitive spaces via remapping and attractor dynamics**

The hippocampus has been shown to contribute to a variety of cognitive domains (15). If the hippocampal-entorhinal system maps multitudes of cognitive spaces, this system needs to exert remarkable flexibility not only in terms of the dimensions it can represent, but also to rapidly switch between cognitive spaces (Fig. 4A). This flexibility is demonstrated by the capacity of hippocampal place cells to undergo global remapping (50–52). Different subsets of hippocampal cells will exhibit place fields in two different environments with spatial relationships among cells active in both environments not being maintained, rendering the two maps orthogonal to each other (Fig. 4B) (51). In contrast to place cells, entorhinal grid patterns have varying offsets to boundaries in different environments, but maintain their spatial phase relative to each other, resulting in consistent relations between the firing fields of grid cells (101, 102). These relationships are maintained not only between environmental contexts but also across behavioral states, with essentially identical cross-correlation patterns exhibited by populations of grid cells and other medial entorhinal cells during free foraging and slow-wave or REM sleep (59, 60).

Remapping-like behavior of hippocampal cells has also been observed for time cells encoding temporal intervals during the delay of a memory-based discrimination task (67). When the length of the delay was altered, a subset of time cells remapped. Some cells ceased to be active, became active at different times of the delay or previously silent cells became active, whereas other cells maintained their firing during absolute or relative times of the altered delay (67). The flexible



recruitment of cells to map positions in cognitive spaces is further illustrated by hippocampal representations of conspecifics, where subsets of place cells exclusively encoded the location of the self, the conspecific or an animate object, whereas other cells exhibited firing fields for both the self and other, but at different locations (83). The decorrelation of hippocampal representations has also been observed using multivariate pattern analyses of human episodic memory. Reconfigurations of hippocampal activity patterns reflect associations and narrative insight in increased similarity, but also reduce overlap resulting in decorrelation through experience (103, 104).

If different cognitive spaces are represented by orthogonal subsets of hippocampal cells, how can already-formed representations of spaces be reinstated to provide stable maps over multiple encounters? In rodents, reinstatement of place cell firing patterns has been observed on a trial-by-trial level indicating rapid switching between maps upon re-exposure to a highly familiar circular and square enclosure (Figure 4C), respectively (105). A likely candidate mechanism governing the reactivation of established cognitive spaces may involve attractor networks (11, 13, 53, 54). In intermediate steps of a morph sequence transforming a square to a circular environment, hippocampal patterns resemble the original map until a switch point at which the representation is pulled towards exhibiting the other map (105), though progressive transformations of hippocampal codes have been observed under different experimental conditions (106). Neuroimaging research points towards a role for stored hippocampal representations in perceptual discrimination (107) and attractor dynamics specifically in memory-guided human behavior (Fig. 4D) (108). Participants learned positions of identical sets of objects in two virtual environments distinguished by background cues and were subsequently tested in a series of morph environments following a linear transformation between the original environments. Spatial memory responses in the intermediate environments followed the sinusoidal pattern predicted by the influence of attractor networks and were paralleled by corresponding non-linear changes of hippocampal activity patterns (108).

These findings indicate that response properties of the hippocampal-entorhinal system enable the formation of independent maps for distinct cognitive spaces to map different dimensions. Consider again the car example: In a different context, weight and engine power might not be the most relevant feature dimensions, but rather a new space might be spanned by the dimensions of possible number of passengers and price (Fig. 4A). Via hippocampal remapping, positions in the new space can be mapped and the positions of different cars are re-defined based on their feature values on the new dimensions, resulting in new similarity relations. The more rigid firing properties of the entorhinal grid system might provide a stable metric through its intrinsic coherence across spaces. Once established, rapid switching between the maps of different cognitive spaces through remapping demonstrates the flexibility of the system and permits the representation of relevant information based on behavioral context. Switches between maps might be triggered by internal or external cues, attentional shifts or changes in task demands rendering other dimensions relevant. Attractor dynamics might serve as anchors for stable representations, which enable learning over multiple encounters as well as the generalization of behavior across similar experiences.

### **Simulations and read-out of trajectories for decisions**

Recording studies in rodents have revealed structured activity of hippocampal place cells during rest in which cells fire sequentially, resulting in trajectories reflecting past experience on a maze (55, 56). This so-called 'replay' demonstrates the ability of spatially tuned cells to represent locations beyond

the animals current position (57). During replay sequences, place cells maintain their spatial relationships to each other (55, 109), resulting in trajectories through space occurring both in forward and reverse order (109, 110). Place cell sequences during replay are temporally compressed compared to trajectories during running (56, 57, 111) and might thereby allow fast simulations of trajectories through cognitive space (Fig. 5AB). Sequences replaying trajectories in reverse order might evaluate previous paths and associate the states visited with reward information for learning of adaptive behavior (112–114, 57). Hippocampal sequences may extend to the simulation of potential future paths and their outcomes. In a spatial alternation task, replay of both correct and incorrect future trajectories supports learning and planning (115) and the disruption of sequences during sharp-wave ripples impairs successful performance (116). Further, place cell sequences reflect future paths during goal-directed behavior (117) and trajectories extending into the shock zone of a linear track prior to avoidance behavior (118). These findings dovetail with extensive research on place cell sequences during ongoing theta oscillations, where trajectories towards potential goals are represented within different theta cycles (61), suggesting simulations of possible trajectories through cognitive space.

What role do mental simulations of trajectories play for planning and decision-making in humans? During navigational planning on a circular track, hippocampal activity patterns carry information not only about the start and goal location, but also locations along the optimal rather than suboptimal path, in accordance with sequential simulations of trajectories through space (119). The entorhinal grid system has also been linked to imagining navigation (34) and snapshots from stationary viewpoints (33), implicating it in the planning of trajectories through space in line with replay in the EC (58–60, 120) and the observation of grid cell firing during covert attentional trajectories through visual space (121). Beyond navigable space, the hippocampal-entorhinal system extracts statistical regularities of non-spatial sequences (122, 123) and forms relational maps of the underlying structures the sequences were derived from (104, 124), potentially drawn upon to plan trajectories through task spaces. Indeed, in a non-spatial decision-making task, sequential reactivations of previously visited states reflected reverse trajectories (Fig. 5C) through a space of discrete states represented by objects (125). While the nature of the MEG signal and the analysis approach focusing on visual responses make a direct hippocampal origin of these results unlikely, they might reflect sequential reinstatement of visual representations orchestrated by the hippocampus (125). Initial evidence suggests sequential hippocampal activity can be observed using fMRI (126). Episodic cues can elicit memory-guided simulation of past experience influencing choice behavior (127) and prospective simulations of trajectories through a task's state space have been linked to model-based behavior in a two-stage decision-making task (128). These findings indicate that trajectories through cognitive spaces representing task states can be simulated based on prior experience.

Cognitive spaces enable generalization and finding novel trajectories via the representation of positions along defined dimensions. The rodent hippocampus can recombine separate trajectories across segments of a two-choice T-maze, usually not experienced successively, to infrequently construct never-experienced place-cell sequences (113). Similarly, hippocampal place cells constructed trajectories through previously non-traversed space after rats observed the placement of a reward on one T-maze arm while confined to the stem of the maze (129). In humans, hippocampal simulations have been linked to the construction of imaginary scenarios (12, 62, 63, 73), an ability impaired in patients with hippocampal lesions (130). Mental simulations can be conceived as putative trajectories through the space of episodic experience (12, 63) and have been

shown to influence decision-making in a delay-discounting task (131). Furthermore, in a task in which participants imagined novel compound goods consisting of two familiar foods the hippocampus and mPFC flexibly combine past experiences (132). These findings show that the hippocampus can flexibly draw on past experience to form and simulate novel trajectories through cognitive space allowing adaptive decision-making and behavior.

Simulations of trajectories building on spatially tuned cells have been incorporated in models of episodic memory and consolidation (12, 63, 133). Replay-inspired simulation of experience has also been used to integrate reinforcement learning and deep neural networks in artificial intelligence (134). Here, we suggest that contemplating a number of stimuli can be conceived as a trajectory through cognitive space. In line with its proposed role in future anticipation and prediction (62, 63, 73, 77, 130), the hippocampal-entorhinal system supports these trajectories via sequential activity of spatially tuned cells. The entorhinal grid system might span up a space based on a set of dimensions and thereby provide the framework for flexible simulations of positions and trajectories by the hippocampus. Relevant feature dimensions can range from locations in a maze to abstract state spaces in decision-making problems. Novel trajectories can be generated from past experience and trajectories can be simulated via sequential hippocampal activity to guide future behavior. Drawing upon the geometric characterization of cognitive spaces, experience can be generalized to the outcome of novel trajectories or actions. For example, if of two prior actions one under- and one overshot a goal, then an intermediate trajectory through cognitive space will approximate the goal more closely. While it has been suggested that sequential activity might reflect inherent hippocampal dynamics (135, 136), other accounts highlight hippocampal interactions with state-space representations in frontal regions in sequence generation and action selection, and the role of striatal regions in sequence evaluation (131, 137–140).

### **Open questions and future directions**

We used spatial navigation investigated in rodents and humans (25, 38) as a model system to identify key neural mechanisms and fused them with concepts from cognitive science to describe the central neural coding machinery underlying higher-level cognition in humans. Future research should help to elucidate the generation of cognitive spaces and their governing principles. For example, the conditions of the generation of the continuous code in the hippocampal-entorhinal system are still largely unclear. This concerns both the circuit level interactions of brain regions (141) and their maturation (142), the role of spatial cognition during development (143) and the potential breakdown of spatial codes in aging and disease (31, 37, 144). Of particular interest is how precisely the entorhinal grid code emerges in its hexagonally symmetric form. Advances in techniques such as two-photon calcium imaging (145) will foster the population level understanding of neural codes in rodents. In concert with the rise of high-resolution fMRI and optimized MEG protocols, this should further bridge the gap between systems and cognitive neuroscience to unravel neural mechanisms promoting refinements of processing in cognitive spaces; and spur the mapping of brain structures to specific functions.

Cognitive spaces can be multi-dimensional. Therefore, the question how a continuous code can be extended to map out additional dimensions is fundamental. Research in humans (146), rats (147) and bats has shed light on how spatial coding of place (148) and head direction cells (149) can be extended to the third dimension. However, grid-like coding in 3D as well as evidence for spatial

coding in cognitive spaces of higher dimensionality remains elusive. Related to this is the question how cognitive spaces are spanned by multiple, potentially integral, dimensions. Further, the nature of the metric underlying cognitive spaces can be investigated using measures of neural and behavioral similarity between positional representations. While the assumption of a Euclidean metric (150) might be most intuitive when comparing cognitive to navigable space, there is evidence for topological representations of spaces in both rodents (151) and humans (124, 152, 153).

While grid cells have been implicated in representing dimensions of cognitive spaces, the extent to which they retain their specific firing patterns remains to be explored. For instance, is the modular organization of the EC stable across cognitive spaces? One might expect that grid cells from the same module also show similar orientation and spacing when representing dimensions of a given cognitive space, but differ in their spatial phase from the boundaries of the space. Generalizing from this question, the role of other cell types encoding spatial information in the context of navigation remains to be elucidated. For example, one could envision a role for border cells in signaling event (154–156) or conceptual boundaries when learning to categorize stimuli drawn from a cognitive space as belonging to different concepts. Likewise, head direction (39), goal direction (41) or object vector (157) cells might be involved in representing relationships between stimuli located at different locations in a cognitive space. Similarly, the role networks of brain regions acting in concert with the hippocampus (32, 63, 73, 74) play for cognitive spaces and differences to the encoding of multidimensional stimulus spaces in other brain regions, for example during face processing in monkey inferotemporal cortex (158), should be explored.

An intriguing question concerns the measurement of behavioral benefits of cognitive space formation, for example via the generation of shortcuts through cognitive space or the impaired ability to find shortcuts in lesion patients. Further, how does information encoded in different cognitive spaces interact? Can trajectories encoded in one space be transferred to another and be retrieved there to guide behavior? Contrarily, it might be possible to bring codes from different cognitive spaces into conflict with one another, which might result in interference across spaces, or investigate effects of deformations of firing patterns (47, 78, 79) across spaces to further elucidate how representations of different spaces are entertained by overlapping neural substrates.

## **Conclusion**

In this theoretical article, we propose cognitive spaces as a primary representational format for information processing in the brain. Combining key mechanisms identified in systems neuroscience and concepts from cognitive science and philosophy, we developed a cognitive neuroscience framework for processing and representing information in cognitive spaces in the hippocampal-entorhinal system. Place and grid cells might have evolved to represent not only navigable space, but to also map dimensions of experience spanning cognitive spaces governed by geometric principles. In these cognitive spaces, stimuli can be located based on their values along the feature dimensions mapped by place and grid cells. These spatially specific cells provide a continuous code that allows similar stimuli to occupy neighboring positions in cognitive space, encoded by overlapping population responses. In this framework, concepts are represented by convex regions of similar stimuli. The multi-scale spatial code along the long-axis of the hippocampal formation enables representing stimuli at different granularities for both generalization and maintenance of fine details in hierarchical knowledge structures. Ever-changing demands requiring the flexible

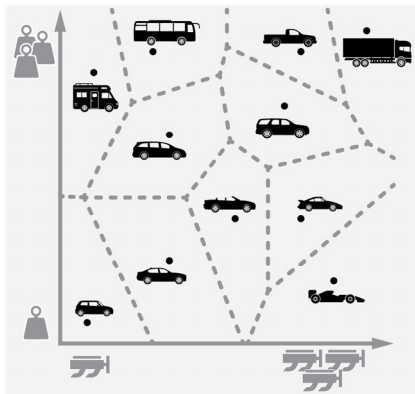
mapping of different dimensions of relevance are met by the capacity of the hippocampus to remap to flexibly form cognitive spaces for which the low-dimensional entorhinal grid code might provide a stable metric. An established mapping of a cognitive space might be reinstated via attractor dynamics and pattern completion to provide stable representations of familiar dimensions. Experiencing a sequence of stimuli results in a trajectory through cognitive space. We propose that sequential hippocampal activity in the form of replay and theta sequences allows simulations of temporally compressed trajectories through cognitive spaces for flexible cognition and adaptive behavior. In sum, we suggest cognitive spaces as a domain-general format for human thinking, thus providing an overarching framework, which can also help to better understand cognitive breakdown in neurodegenerative diseases (31, 144) and to inform novel architectures in artificial intelligence (134).

### **Box: Hexadirectional signals in fMRI**

Grid cells are defined by their six-fold symmetric firing patterns tiling environments in a highly regular fashion (24). Hexadirectional signals serve as a proxy measure for grid-like activity in BOLD-fMRI during trajectories through cognitive spaces to investigate the role of the grid system in higher-level cognition in the healthy human brain (22, 31–37). Three analysis approaches relying on a directional bias of activity have been employed. In the orientation-estimation approach (22) the data are partitioned to estimate the orientation of the hexadirectional signal in one part of the data. The prediction of increased levels of BOLD-activity for trajectories aligned versus misaligned with the estimated orientation is then tested on an independent data partition (see figure). In the second analysis approach, which is based on fMRI repetition suppression, the hexadirectional grid code is reflected in correlations of BOLD-activity during trajectories in a given direction with the time since the last trajectory at an angular offset of 60° (22). The third approach analyzes the similarity of multi-voxel patterns as a function of angular differences between trajectory directions to test the assumption that activity of the grid system should be reflected in a 60°-modulation of entorhinal activity patterns (33). Hexadirectional signals can serve as a showcase example for how insights from rodent electrophysiology might be translated to human navigation in fMRI and explored in human cognition and behavior more broadly (22, 31–37).

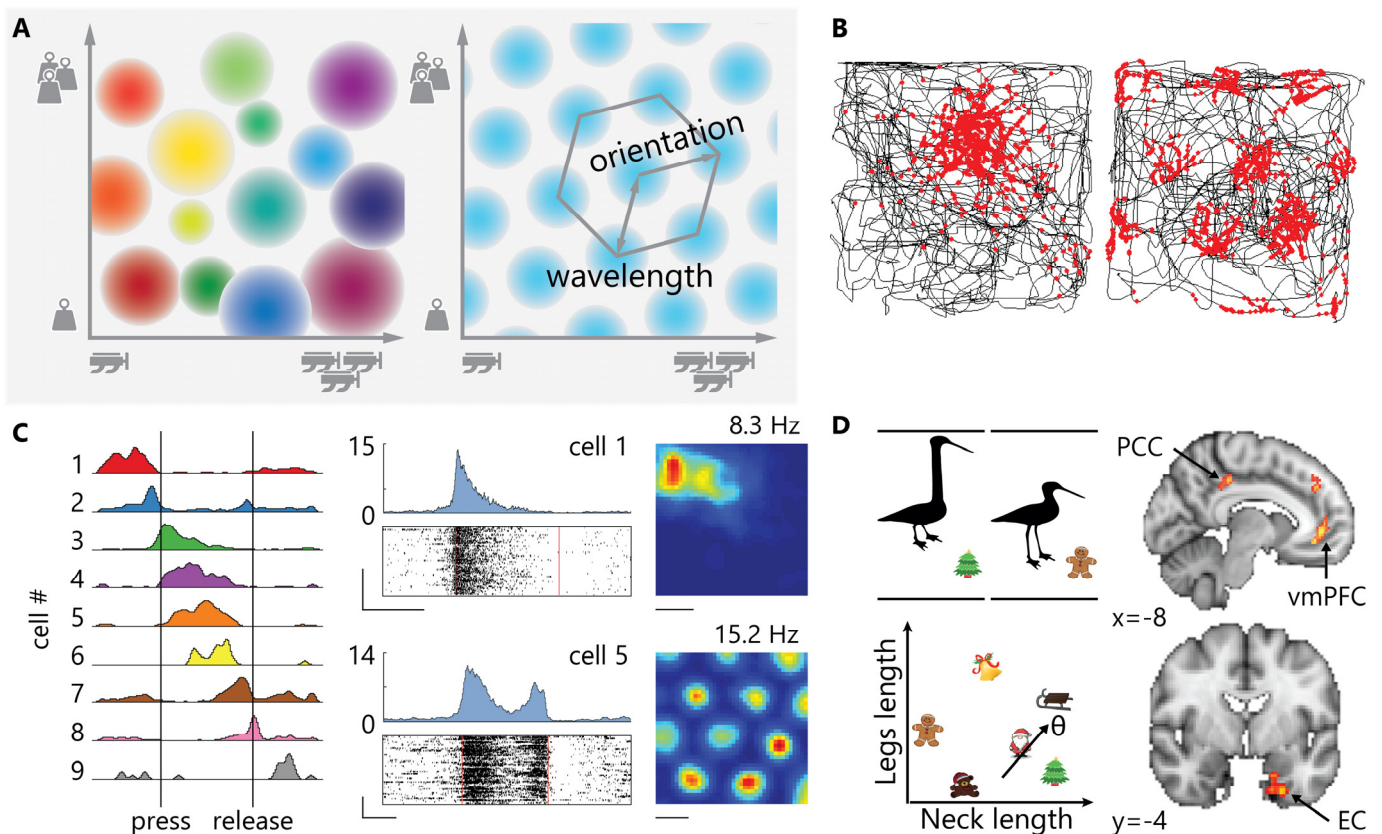
While the BOLD-response measured with fMRI does not reflect single cell activity, it provides a hemodynamic proxy of population activity (159). How can the 60°-symmetry of grid cell firing patterns be translated to a bias in population activity picked up by fMRI during trajectories through cognitive space? While a single grid cell intuitively exhibits hexadirectional activity, a hexadirectional bias on the population level might result if orientations (78, 79) and spatial offsets to boundaries (160) of the firing patterns are clustered across cells; which could also potentially be reflected in directional biases in the local field potential in EC. Additionally, conjunctive grid cells (25) also modulated by heading direction might contribute to the hexadirectional signal if the preferred directions of these cells align with the axes of the cells' grid pattern (22). Technological advances such as two-photon microscopy (145) enabling imaging of cellular responses in larger portions of the rodent brain might allow future research to shed further light on the dynamics underlying hexadirectional signals on the population level.

**Figure 1**



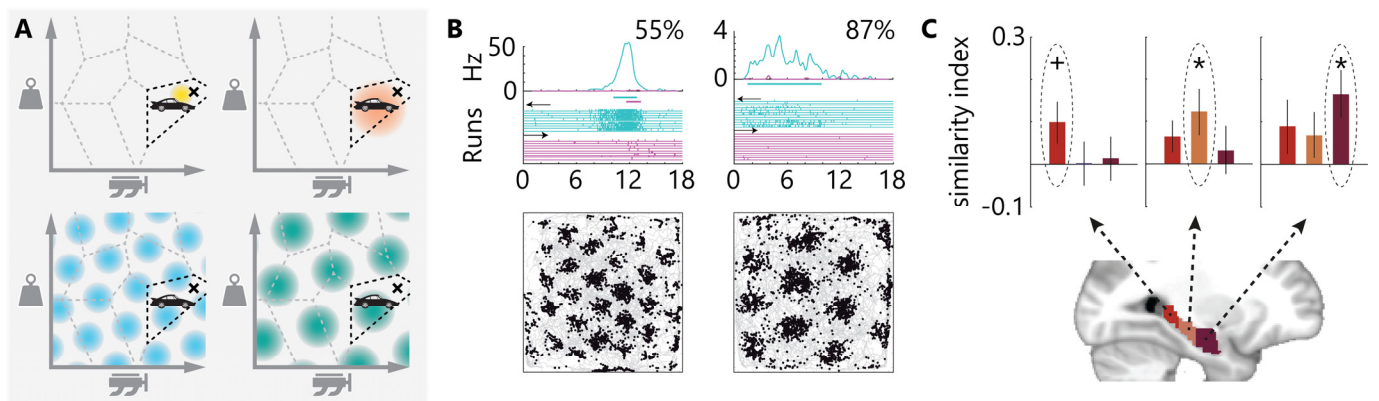
**Fig. 1. Two-dimensional cognitive space.** Schematic of a cognitive space spanned by the dimensions of car weight and engine power. Feature values along the two axes define positions of different cars resulting in stimuli with similar properties being located nearby. Concepts (car icons) are defined as convex regions of the cognitive space and are indicated by dashed lines obtained from a Voronoi tessellation of the space. Under the assumption of a Euclidean metric this discretizes the space into convex regions by assigning each point in space to the region around the closest prototypical exemplar (black dots) with distances based on dissimilarity along the feature dimensions.

**Figure 2**



**Fig. 2. Place and grid cells map dimensions of cognitive spaces.** (A) Colored circles illustrate firing fields of hippocampal place cells in a space spanned by the dimensions of car engine power and weight. Each color represents one individual place cell. Collectively, the firing fields map all locations of the space. The hexagonally symmetric firing pattern of grid cells (right panel shows one cell's pattern) provide a metric for the dimensions of the space. (B) Firing of a place (left) and grid (right) cell recorded from the hippocampus and entorhinal cortex, respectively. Spike locations (red dots) are shown on the animal's path (black line) through a square enclosure (adapted from Moser et al., Annual Review of Neuroscience, 2008). (C) Frequency fields of place and grid cells. While the animal presses a lever to manipulate the frequency of a tone, different place and grid cells fire at different frequencies during the frequency sweep (left). Two example cells active in the sound-modulation task identified as a place and grid cell, respectively, during navigation (right, adapted from Aronov et al., Nature, 2017). (D) Hexadirectional signals in two-dimensional bird space. Participants learned associations of Christmas symbols and stick-figure birds with variable leg and neck length (left column). During trajectories through this space defined by morphing birds, hexadirectional signals (right column) were observed in the entorhinal cortex (ERH) and a network of brain regions implicated in mnemonic processing (adapted from Constantinescu et al., Science, 2016).

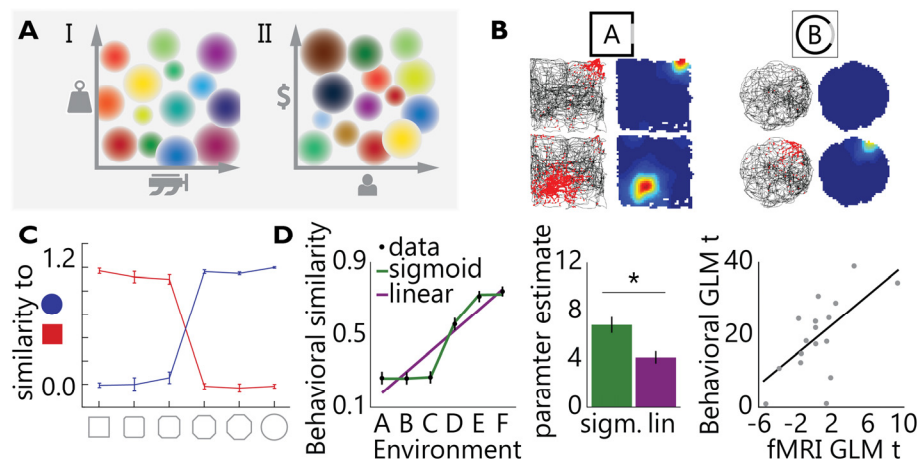
**Figure 3**



**Fig. 3. Multiple levels of representation.** (A) Representations of information at different levels of granularity can be supported by multiple spatial scales in the hippocampal formation. Narrower (left column) or broader (right column) positions in cognitive spaces can be encoded by place (top row) and grid cells (bottom row). Schematics illustrate firing fields of four different cells. Higher-level information (e.g. about the concept of sports cars) can be ascribed to a lower-level stimulus (black cross). For example, when learning that a specific car model is a sports car we can infer that it is likely to have high engine power. (B) The firing field size of place cells (top) and the size and spacing of grid firing fields (bottom) increase from dorsal to ventral recording sites in the rodent hippocampal-entorhinal system (adapted from Kjelstrup et al., *Science*, 2008 and Stensola et al., *Nature*, 2012). (C) The granularity of mnemonic networks scales along the long-axis of the human hippocampus with pairwise associations of elements in posterior and integrated networks in anterior hippocampus (adapted from Collin et al., *Nature Neuroscience*, 2015).

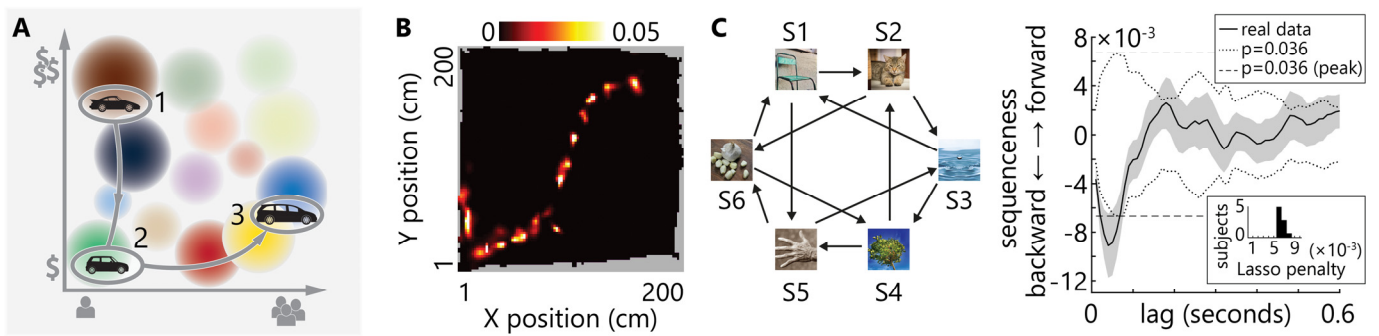


**Figure 4**



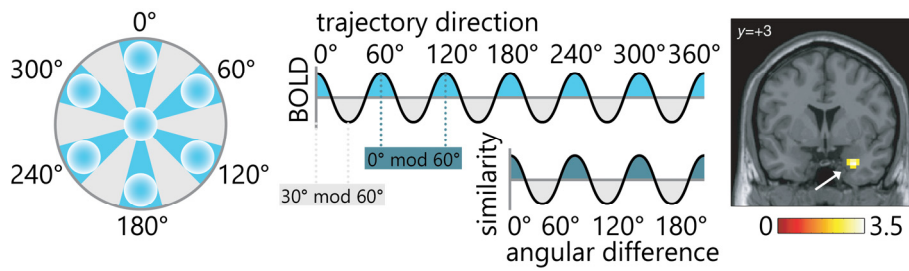
**Fig. 4. Remapping and attractor dynamics for flexible cognitive spaces.** (A) Different task-relevant dimensions (e.g. the maximum occupancy and price instead of car engine power and weight) can be mapped by the recruitment of a different subset of place cells and the rearrangement of firing fields between spaces. (B) In navigating rodents place cells remap to represent different environments. Spike locations (red) overlaid on the animal's path (black) and rate maps (warmer colors indicating increased firing) are shown for two place cells (rows) in two environments (columns). One place cell is active in the square, but not the circular environment, whereas the other exhibits a firing field in both environments, but at unrelated positions (adapted from Fyhn et al., *Nature*, 2007). (C) Attractor dynamics enable rapid switches between established maps. The similarity to established maps of a square (red) and circular (blue) environment is shown for a sequence of intermediate environments (adapted from Wills et al., *Science*, 2005). The data show a sigmoidal rather than linear shift function. (D) Attractor dynamics in human spatial memory. After learning object locations in two base environments A and F, participants were tested in a sequence of intermediate environments. Spatial memory responses (left) exhibited patterns better explained by a sigmoidal than a linear function. Hippocampal multi-voxel pattern similarity was better predicted by a sigmoidal than a linear model of behavior (middle) and the fit of a canonical sigmoidal to spatial memory responses was associated with hippocampal pattern similarity values across participants (right, adapted from Steemers et al., *Current Biology*, 2016).

**Figure 5**



**Fig. 5. Hippocampal simulations via sequential activity.** (A) Sequential activity simulates different positions (1-3) in cognitive space allowing the evaluation of different car types along the trajectory when deciding which car to buy. (B) Simulation of trajectory in a square enclosure. The posterior probabilities of positions from Bayesian decoding applied to different time frames during a sharp wave ripple event result in a trajectory through space (adapted from Pfeiffer & Foster, *Science*, 2015). (C) Reverse trajectories through state space. States were represented by images with arrows indicating possible transitions between states (left). The solid black line shows a ‘sequenceness’ measure indicating the probability of decoding a successive or preceding state (positive and negative values, respectively) at different time points following the decoding of a state during planning (right). Reverse sequences were observed with a lag of around 40ms between state space positions (adapted from Kurth-Nelson et al., *Neuron*, 2016).

### Box Figure



**Hexadirectional signals in fMRI.** The number of grid cell firing fields crossed depends on the direction of a trajectory through cognitive space. More fields are crossed during trajectories aligned with the grid (blue), translating to stronger entorhinal BOLD activity (middle top, see text for details). This effect was first observed during virtual navigation in the entorhinal cortex (right, adapted from Doeller et al., Nature, 2010). The similarity structure of entorhinal multi-voxel patterns exhibits a 60°-modulation when comparing trajectories as a function of their angular difference as the grid is sampled at the same phase every 60° (middle bottom, red), irrespective of its alignment.

## References and Notes

1. E. C. Tolman, *Psychol. Rev.* **55**, 189–208 (1948).
2. J. O’Keefe, J. Dostrovsky, *Brain Res.* **34**, 171–175 (1971).
3. J. O’Keefe, L. Nadel, *The Hippocampus as a Cognitive Map* (Oxford: Clarendon Press, 1978).
4. K. M. Gothard, W. E. Skaggs, K. M. Moore, B. L. McNaughton, *J. Neurosci.* **16**, 823–835 (1996).
5. E. R. Wood, P. A. Dudchenko, H. Eichenbaum, *Nature.* **397**, 613–616 (1999).
6. B. J. Young, G. D. Fox, H. Eichenbaum, *J. Neurosci.* **14**, 6553–6563 (1994).
7. E. R. Wood, P. A. Dudchenko, R. J. Robitsek, H. Eichenbaum, *Neuron.* **27**, 623–633 (2000).
8. N. J. Cohen, H. Eichenbaum, *Memory, amnesia, and the hippocampal system.* (The MIT Press, Cambridge, MA, US, 1993).
9. H. Eichenbaum, N. J. Cohen, *Trends Neurosci.* **11**, 244–248 (1988).
10. H. Eichenbaum, N. J. Cohen, *Neuron.* **83**, 764–770 (2014).
11. D. Marr, *Phil Trans R Soc Lond B.* **262**, 23–81 (1971).
12. M. E. Hasselmo, *How We Remember: Brain Mechanisms of Episodic Memory* (The MIT Press, Cambridge, MA, US, 2011).
13. B. L. McNaughton, R. G. M. Morris, *Trends Neurosci.* **10**, 408–415 (1987).
14. D. Schiller *et al.*, *J. Neurosci.* **35**, 13904–13911 (2015).
15. D. E. Hannula, M. C. Duff, *The Hippocampus from Cells to Systems: Structure, Connectivity, and Functional Contributions to Memory and Flexible Cognition* (Springer International Publishing, Cham, 2017).
16. P. Gärdenfors, *Conceptual spaces: the geometry of thought* (MIT Press, Cambridge, Mass, 2000).
17. C. Balkenius, P. Gärdenfors, *Front. Psychol.* **7** (2016), doi:10.3389/fpsyg.2016.01820.
18. F. Zenker, P. Gärdenfors, Eds., *Applications of conceptual spaces: the case for geometric knowledge representation* (Springer, Cham Heidelberg New York, 2015).
19. E. Rosch, *Cognit. Psychol.* **7**, 532–547 (1975).
20. R. E. Clark, L. R. Squire, *Proc. Natl. Acad. Sci.* **110**, 10365–10370 (2013).
21. G. Buzsáki, E. I. Moser, *Nat. Neurosci.* **16**, 130–138 (2013).
22. C. F. Doeller, C. Barry, N. Burgess, *Nature.* **463**, 657–661 (2010).
23. E. S. Spelke, S. A. Lee, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 2784–2793 (2012).
24. T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, E. I. Moser, *Nature.* **436**, 801–806 (2005).
25. E. I. Moser, M.-B. Moser, B. L. McNaughton, *Nat. Neurosci.* **20**, 1448–1464 (2017).
26. D. Bush, C. Barry, D. Manson, N. Burgess, *Neuron.* **87**, 507–520 (2015).
27. A. V. Herz, A. Mathis, M. Stemmler, *Curr. Opin. Neurobiol.* **46**, 99–108 (2017).
28. B. L. McNaughton, F. P. Battaglia, O. Jensen, E. I. Moser, M.-B. Moser, *Nat. Rev. Neurosci.* **7**, 663–78 (2006).
29. A. D. Ekstrom *et al.*, *Nature.* **425**, 184–188 (2003).
30. J. Jacobs *et al.*, *Nat. Neurosci.* **16**, 1188–1190 (2013).
31. L. Kunz *et al.*, *Science.* **350**, 430–433 (2015).
32. A. O. Constantinescu, J. X. O’Reilly, T. E. J. Behrens, *Science.* **352**, 1464–1468 (2016).
33. J. L. S. Bellmund, L. Deuker, T. Navarro Schröder, C. F. Doeller, *eLife.* **5**, e17089 (2016).
34. A. J. Horner, J. A. Bisby, E. Zotow, D. Bush, N. Burgess, *Curr. Biol.* **26**, 842–847 (2016).

35. M. Nau, T. Navarro Schröder, J. L. S. Bellmund, C. F. Doeller, *Nat. Neurosci.* **21**, 188–190 (2018).
36. J. B. Julian, A. T. Keinath, G. Frazzetta, R. A. Epstein, *Nat. Neurosci.* **21**, 191–194 (2018).
37. M. Stangl *et al.*, *Curr. Biol.* **28**, 1108–1115.e6 (2018).
38. R. A. Epstein, E. Z. Patai, J. B. Julian, H. J. Spiers, *Nat. Neurosci.* **20**, 1504 (2017).
39. K. E. Cullen, J. S. Taube, *Nat. Neurosci.* **20**, 1465 (2017).
40. V. Hok, E. Save, P. P. Lenck-Santini, B. Poucet, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 4602–4607 (2005).
41. A. Sarel, A. Finkelstein, L. Las, N. Ulanovsky, *Science*. **355**, 176–180 (2017).
42. E. Kropff, J. E. Carmichael, M.-B. Moser, E. I. Moser, *Nature*. **523**, 419–424 (2015).
43. F. Savelli, D. Yoganarasimha, J. J. Knierim, *Hippocampus*. **18**, 1270–1282 (2008).
44. T. Solstad, C. N. Boccara, E. Kropff, M.-B. Moser, E. I. Moser, *Science*. **322**, 1865–1868 (2008).
45. C. Lever, S. Burton, A. Jeewajee, J. O’Keefe, N. Burgess, *J. Neurosci.* **29**, 9771–9777 (2009).
46. M. W. Jung, S. I. Wiener, B. L. McNaughton, *J. Neurosci.* **14**, 7347–7356 (1994).
47. C. Barry, R. Hayman, N. Burgess, K. J. Jeffery, *Nat. Neurosci.* **10**, 682–684 (2007).
48. K. B. Kjelstrup *et al.*, *Science*. **321**, 140–143 (2008).
49. H. Stensola *et al.*, *Nature*. **492**, 72–78 (2012).
50. E. Bostock, R. U. Muller, J. L. Kubie, *Hippocampus*. **1**, 193–205 (1991).
51. S. Leutgeb, J. K. Leutgeb, A. Treves, M.-B. Moser, E. I. Moser, *Science*. **305**, 1295–1298 (2004).
52. R. U. Muller, J. L. Kubie, *J. Neurosci.* **7**, 1951–1968 (1987).
53. J. J. Hopfield, *Proc. Natl. Acad. Sci.* **79**, 2554–2558 (1982).
54. J. L. McClelland, B. L. McNaughton, R. C. O’Reilly, *Psychol. Rev.* **102**, 419–457 (1995).
55. M. A. Wilson, B. L. McNaughton, *Science*. **265**, 676–679 (1994).
56. W. E. Skaggs, B. L. McNaughton, *Science*. **271**, 1870–1873 (1996).
57. D. J. Foster, *Annu. Rev. Neurosci.* **40**, 581–602 (2017).
58. H. F. Ólafsdóttir, F. Carpenter, C. Barry, *Nat. Neurosci.* **19**, 792–794 (2016).
59. R. J. Gardner, L. Lu, T. Wernle, M.-B. Moser, E. I. Moser, *bioRxiv*, 198499 (2017).
60. S. G. Trettel, J. B. Trimper, E. Hwaun, I. R. Fiete, L. L. Colgin, *bioRxiv*, 198671 (2017).
61. A. Johnson, A. D. Redish, *J. Neurosci.* **27**, 12176–12189 (2007).
62. R. L. Buckner, *Annu. Rev. Psychol.* **61**, 27–48 (2010).
63. P. Byrne, S. Becker, N. Burgess, *Psychol. Rev.* **114**, 340–75 (2007).
64. N. J. Killian, M. J. Jutras, E. A. Buffalo, *Nature*. **491**, 761–764 (2012).
65. N. J. Killian, S. M. Potter, E. A. Buffalo, *Proc. Natl. Acad. Sci.*, 201417059 (2015).
66. E. Pastalkova, V. Itskov, A. Amarasingham, G. Buzsáki, *Science*. **321**, 1322–1327 (2008).
67. C. J. MacDonald, K. Q. Lepage, U. T. Eden, H. Eichenbaum, *Neuron*. **71**, 737–749 (2011).
68. B. J. Kraus, R. J. Robinson II, J. A. White, H. Eichenbaum, M. E. Hasselmo, *Neuron*. **78**, 1090–1101 (2013).
69. B. J. Kraus *et al.*, *Neuron*. **88**, 578–589 (2015).
70. A. Tsao *et al.*, *Nature* (in press).
71. D. Aronov, R. Nevers, D. W. Tank, *Nature*. **543**, 719–722 (2017).
72. B. A. Radvansky, D. A. Dombeck, *Nat. Commun.* **9**, 839 (2018).
73. D. L. Schacter, D. R. Addis, R. L. Buckner, *Nat. Rev. Neurosci.* **8**, 657–661 (2007).
74. C. Ranganath, M. Ritchey, *Nat. Rev. Neurosci.* **13**, 713–726 (2012).

75. R. N. Shepard, *Science*. **237**, 1317–1323 (1987).
76. Y. Dordek, D. Soudry, R. Meir, D. Derdikman, *eLife*. **5**, e10094 (2016).
77. K. L. Stachenfeld, M. M. Botvinick, S. J. Gershman, *Nat. Neurosci.* **20**, 1643–1653 (2017).
78. J. Krupic, M. Bauza, S. Burton, C. Barry, J. O’Keefe, *Nature*. **518**, 232–235 (2015).
79. T. Stensola, H. Stensola, M.-B. Moser, E. I. Moser, *Nature*. **518**, 207–212 (2015).
80. R. Q. Quiroga, L. Reddy, G. Kreiman, C. Koch, I. Fried, *Nature*. **435**, 1102–1107 (2005).
81. A. Stolk, L. Verhagen, I. Toni, *Trends Cogn. Sci.* **20**, 180–191 (2016).
82. T. Danjo, T. Toyozumi, S. Fujisawa, *Science*. **359**, 213–218 (2018).
83. D. B. Omer, S. R. Maimon, L. Las, N. Ulanovsky, *Science*. **359**, 218–224 (2018).
84. R. M. Tavares *et al.*, *Neuron*. **87**, 231–243 (2015).
85. D. Kumaran, A. Banino, C. Blundell, D. Hassabis, P. Dayan, *Neuron*. **92**, 1135–1147 (2016).
86. L. Deuker, J. L. S. Bellmund, T. Navarro Schröder, C. F. Doeller, *eLife*. **5**, e16534 (2016).
87. A. D. Ekstrom, C. Ranganath, *Hippocampus* (2017), doi:10.1002/hipo.22750.
88. A. M. Collins, M. R. Quillian, *J. Verbal Learn. Verbal Behav.* **8**, 240–247 (1969).
89. J. B. Tenenbaum, C. Kemp, T. L. Griffiths, N. D. Goodman, *Science*. **331**, 1279–1285 (2011).
90. I. K. Brunec *et al.*, *Curr. Biol.* **0** (2018), doi:10.1016/j.cub.2018.05.016.
91. V. H. Brun *et al.*, *Hippocampus*. **18**, 1200–1212 (2008).
92. B. A. Strange, M. P. Witter, E. S. Lein, E. I. Moser, *Nat. Rev. Neurosci.* **15**, 655–669 (2014).
93. S. H. P. Collin, B. Milivojevic, C. F. Doeller, *Nat. Neurosci.* **18**, 1562–1564 (2015).
94. B. Milivojevic, C. F. Doeller, *J. Exp. Psychol. Gen.* (2013), doi:10.1037/a0033746.
95. C. Baldassano *et al.*, *Neuron*. **95**, 709–721.e5 (2017).
96. R. W. Komorowski *et al.*, *J. Neurosci.* **33**, 8079–8087 (2013).
97. S. McKenzie *et al.*, *Neuron*. **83**, 202–215 (2014).
98. A. R. Preston, Y. Shrager, N. M. Dudukovic, J. D. E. Gabrieli, *Hippocampus*. **14**, 148–52 (2004).
99. B. A. Kuhl, A. T. Shah, S. DuBrow, A. D. Wagner, *Nat. Neurosci.* **13**, 501–6 (2010).
100. M. L. Schlichting, A. R. Preston, *Curr. Opin. Behav. Sci.* **1**, 1–8 (2015).
101. M. Fyhn, T. Hafting, A. Treves, M.-B. Moser, E. I. Moser, *Nature*. **446**, 190–194 (2007).
102. K. Yoon *et al.*, *Nat. Neurosci.* **16**, 1077–1084 (2013).
103. B. Milivojevic, A. Vicente-Grabovetsky, C. F. Doeller, *Curr. Biol.* **25**, 821–830 (2015).
104. A. C. Schapiro, L. V. Kustner, N. B. Turk-Browne, *Curr. Biol. CB.* **22**, 1622–7 (2012).
105. T. J. Wills, C. Lever, F. Cacucci, N. Burgess, J. O’Keefe, *Science*. **308**, 873–876 (2005).
106. J. K. Leutgeb *et al.*, *Neuron*. **48**, 345–358 (2005).
107. H. M. Bonnici *et al.*, *Hippocampus*. **22**, 1143–1153 (2012).
108. B. Steemers *et al.*, *Curr. Biol.* **26**, 1750–1757 (2016).
109. K. Diba, G. Buzsáki, *Nat. Neurosci.* **10**, 1241–1242 (2007).
110. T. J. Davidson, F. Kloosterman, M. A. Wilson, *Neuron*. **63**, 497–507 (2009).
111. A. K. Lee, M. A. Wilson, *Neuron*. **36**, 1183–1194 (2002).
112. D. J. Foster, M. A. Wilson, *Nature*. **440**, 680–683 (2006).
113. A. S. Gupta, M. a a van der Meer, D. S. Touretzky, a D. Redish, *Neuron*. **65**, 695–705 (2010).
114. D. Dupret, J. O’Neill, B. Pleydell-Bouverie, J. Csicsvari, *Nat. Neurosci.* **13**, 995–1002 (2010).
115. A. C. Singer, M. F. Carr, M. P. Karlsson, L. M. Frank, *Neuron*. **77**, 1163–1173 (2013).
116. S. P. Jadhav, C. Kemere, P. W. German, L. M. Frank, *Science*. **336**, 1454–1458 (2012).

117. B. E. Pfeiffer, D. J. Foster, *Nature*. **497**, 74–9 (2013).
118. C.-T. Wu, D. Haggerty, C. Kemere, D. Ji, *Nat. Neurosci.* **20**, 571–580 (2017).
119. T. I. Brown *et al.*, *Science*. **352**, 1323–1326 (2016).
120. J. O’Neill, C. N. Boccara, F. Stella, P. Schoenenberger, J. Csicsvari, *Science*. **355**, 184–188 (2017).
121. N. Wilming, P. König, S. König, E. A. Buffalo, *eLife*. **7**, e31745 (2018).
122. D. Kumaran, E. A. Maguire, *PLOS Biol.* **4**, e424 (2006).
123. L.-T. Hsieh, M. J. Gruber, L. J. Jenkins, C. Ranganath, *Neuron*. **81**, 1165–1178 (2014).
124. M. M. Garvert, R. J. Dolan, T. E. Behrens, *eLife*. **6**, e17086 (2017).
125. Z. Kurth-Nelson, M. Economides, R. J. Dolan, P. Dayan, *Neuron*. **91**, 194–204 (2016).
126. N. W. Schuck, Y. Niv, *bioRxiv*, 315978 (2018).
127. A. M. Bornstein, K. A. Norman, *Nat. Neurosci.* **20**, 997–1003 (2017).
128. B. B. Doll, K. D. Duncan, D. A. Simon, D. Shohamy, N. D. Daw, *Nat. Neurosci.* **18**, 767–772 (2015).
129. H. F. Ólafsdóttir, C. Barry, A. B. Saleem, D. Hassabis, H. J. Spiers, *eLife*. **4**, e06063 (2015).
130. D. Hassabis, D. Kumaran, S. D. Vann, E. A. Maguire, *Proc. Natl. Acad. Sci.* **104**, 1726–1731 (2007).
131. J. Peters, C. Büchel, *Neuron*. **66**, 138–148 (2010).
132. H. C. Barron, R. J. Dolan, T. E. J. Behrens, *Nat. Neurosci.* **16**, 1492–1498 (2013).
133. D. Kumaran, D. Hassabis, J. L. McClelland, *Trends Cogn. Sci.* **20**, 512–534 (2016).
134. V. Mnih *et al.*, *Nature*. **518**, 529–533 (2015).
135. G. Dragoi, S. Tonegawa, *Nature*. **469**, 397–401 (2011).
136. A. D. Grosmark, G. Buzsáki, *Science*. **351**, 1440–1443 (2016).
137. M. A. A. van der Meer, A. Johnson, N. C. Schmitzer-Torbert, A. D. Redish, *Neuron*. **67**, 25–32 (2010).
138. N. W. Schuck, M. B. Cai, R. C. Wilson, Y. Niv, *Neuron*. **91**, 1402–1412 (2016).
139. A. M. Wikenheiser, G. Schoenbaum, *Nat. Rev. Neurosci.* **17**, 513–523 (2016).
140. R. Kaplan, N. W. Schuck, C. F. Doeller, *Trends Neurosci.* **40**, 256–259 (2017).
141. E. I. Moser *et al.*, *Nat. Rev. Neurosci.* **15**, 466–481 (2014).
142. F. Donato, R. I. Jacobsen, M.-B. Moser, E. I. Moser, *Science*. **355**, eaai8178 (2017).
143. L. Hermer, E. S. Spelke, *Nature*. **370**, 57 (1994).
144. H. Fu *et al.*, *Neuron*. **93**, 533-541.e5 (2017).
145. D. A. Dombeck, A. N. Khabbaz, F. Collman, T. L. Adelman, D. W. Tank, *Neuron*. **56**, 43–57 (2007).
146. M. Kim, K. J. Jeffery, E. A. Maguire, *J. Neurosci.* **37**, 4270–4279 (2017).
147. R. Hayman, M. A. Verriotis, A. Jovalekic, A. A. Fenton, K. J. Jeffery, *Nat. Neurosci.* **14**, 1182 (2011).
148. M. M. Yartsev, N. Ulanovsky, *Science*. **340**, 367–372 (2013).
149. A. Finkelstein *et al.*, *Nature*. **517**, 159–164 (2015).
150. S. M. Kosslyn, T. M. Ball, B. J. Reiser, *J. Exp. Psychol. Hum. Percept. Perform.* **4**, 47–60 (1978).
151. Y. Dabaghian, V. L. Brandt, L. M. Frank, *eLife*. **3**, e03476 (2014).
152. Z. W. Pylyshyn, *Psychol. Bull.* **80**, 1–24 (1973).
153. W. H. Warren, D. B. Rothman, B. H. Schnapp, J. D. Ericson, *Cognition*. **166**, 152–163 (2017).

154. A. Ben-Yakov, N. Eshel, Y. Dudai, *J. Exp. Psychol. Gen.* **142**, 1255–1263 (2013).
155. Y. Ezzyat, L. Davachi, *Neuron*. **81**, 1179–1189 (2014).
156. I. K. Brunec, M. Moscovitch, M. D. Barense, *Trends Cogn. Sci.* **22**, 637–650 (2018).
157. Ø. A. Høydal, E. R. Skytøen, M.-B. Moser, E. I. Moser, *bioRxiv*, 286286 (2018).
158. L. Chang, D. Y. Tsao, *Cell*. **169**, 1013-1028.e14 (2017).
159. N. K. Logothetis, *Nature*. **453**, 869–78 (2008).
160. J. G. Heys, K. V. Rangarajan, D. A. Dombeck, *Neuron*. **84**, 1079–1090 (2014).



## **Acknowledgments**

**Funding:** CFD's research is supported by the Max Planck Society; the European Research Council (ERC-CoG GEOCOG 724836); the Kavli Foundation, the Centre of Excellence scheme of the Research Council of Norway – Centre for Neural Computation, The Egil and Pauline Braathen and Fred Kavli Centre for Cortical Microcircuits, the National Infrastructure scheme of the Research Council of Norway – NORBRAIN; and the Netherlands Organisation for Scientific Research (NWO-Vidi 452-12-009; NWO-Gravitation 024-001-006; NWO-MaGW 406-14-114; NWO-MaGW 406-15-291). **Author contributions:** J.B. and C.D. conceived the presented ideas. J.B. wrote the manuscript, with input from C.D., P.G. and E.M. All authors contributed to the final manuscript. **Competing interests:** The authors declare no conflict of interest.