Emergence of Cognition from Action

GYORGY BUZSÁKI,[1,2] ADRIEN PEYRACHE,[1] AND JOHN KUBIE[3]

[1]*The Neuroscience Institute, New York University, School of Medicine, New York, New York 10016*
[2]*Center for Neural Science, New York University, School of Medicine, New York, New York 10016*
[3]*Department of Cell Biology, Downstate Medical Center, State University of New York at Brooklyn, New York, New York 11203*
Correspondence: gyorgy.buzsaki@nyumc.org

Theories of brain function have evolved through multiple stages. The first proposition was that brain networks support a set of reflex responses, with current sensory inputs producing immediate motor outputs. The behaviorist paradigm suggested that actions can always be explained as a response to immediate external cues. In response to these views, the cognitive paradigm argued that behavior cannot be understood simply as input–output functions because the hidden layers of brain generate unpredictability. The central processing was termed "cognition." Here we propose a neuroscience-based model of cognition. Our core hypothesis is that cognition depends on internal models of the animal and its world, where internally generated sequences can serve to perform "what if" scenarios and anticipate the possible consequences of alternative actions without actually testing them, and aid in the decisions of overt actions. We support our hypotheses by several examples of recent experimental findings and show how externally guided cell assembly sequences become internalized to support cognitive functions.

Humans, and likely other animals, can imagine far into the future; we spend a great deal of waking time devoted to such covert activities (Killingsworth and Gilbert 2010; Immordino-Yang et al. 2012; Mar et al. 2014). The main thesis of this essay is that such ability of the brain arises from its disengagement from immediate overt actions as a result of experience. We hypothesize that brains, irrespective of size, are predictive devices that exploit regularity and recurrence as a fundamental property of the surrounding world and apply effective heuristics to make such predictions. Prediction is made possible by adaptive mechanisms that are supported by learning rules that either apply across generations (evolutionary adaptation) or within the lifetime of the organism. As a result, organisms can deal with a future occurrence of the same or similar situations more effectively. This is the fundamental organization principle of any adaptive system. Here we describe example organizational brain patterns, with increasing predictive efficiency, and present a hypothesis of how internally generated neuronal network operations, detached from outside inputs, can serve as the necessary substrate for cognition.

One aspect of the argument we will present is that most brain structures can have "dual use": at times, connected to immediate sensory input and output; at other times entrained by internal signals and brain rhythms. For many years, researchers tended to associate preexisting verbal terms, such as memory, planning, envisioning the future, volition, and decision-making, with different and distinct brain structures. Oftentimes, even studying the same structure or system generated seemingly contradictory hypotheses depending on the chosen approach, preconception, or experimental method used. A striking example is the hippocampal system. Lesion data in humans provided evidence that bilateral removal of the hippocampi produced severe and irreversible amnesia (Scoville and Milner 1957). Single unit studies in the hippocampus and entorhinal cortex of animals gave rise to the prominent theory that the fundamental function of the hippocampal system is supporting spatial navigation (O'Keefe and Nadel 1978; Hafting et al. 2005; McNaughton et al. 2006), assisted by inputs from the head-direction system (Ranck 1985; Taube 2007). In contrast, studying the collective behavior of hippocampal neurons by recording the local field potential (LFP) offered the conclusion that hippocampal theta oscillations are an unmistakable reflection of voluntary action (Vanderwolf 1969). These independent ideas of hippocampal functions persisted in parallel for decades without true interactions, apart from occasional polite gestures and references to the competing frameworks (Buzsáki 2005). As discussed below, these barriers are now being dismantled by the recognition that although neurons and networks in the memory/navigational/planning systems are primarily engaged in self-organized activity by theta oscillations, they can robustly respond to environmental or body-derived inputs (Frank et al. 2000; Wood et al. 2000; Fujisawa et al. 2008; Pastalkova et al. 2008).

In "simple" nervous systems, the connection between output and input networks is direct and immediate and the consequences of their action are mediated by the environment, as illustrated in Figure 1A. More complex brains are organized in a "multiple loop" pattern; that is, a series of interacting parallel loops are imposed between input and output, as illustrated in Figure 1B,C. For example, in mammals, new circuits are added onto the monosynaptic
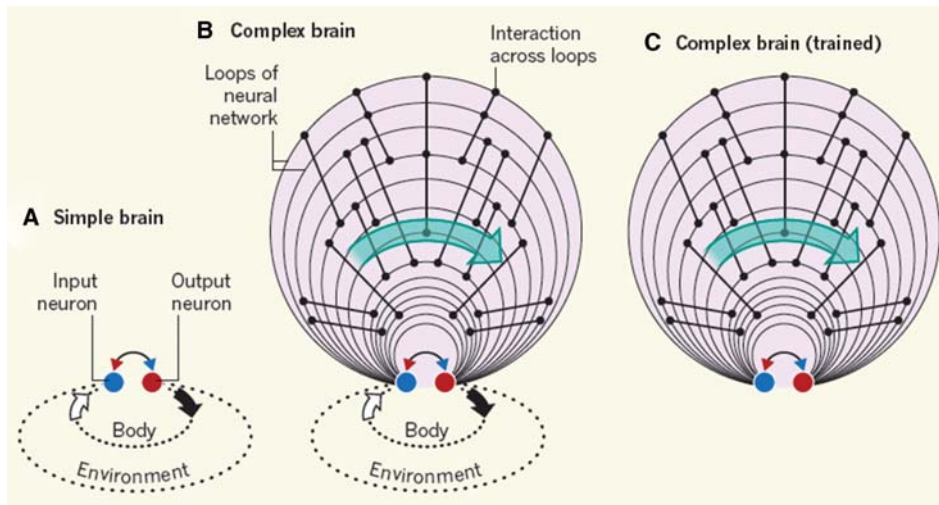
Figure 1. Externally driven and self-organized cell assemblies. (*A*) Evolutionarily simple brains contain simple neural networks. Sensory input from the body and the environment activates input neurons, which interact with output neurons to generate appropriate reflex actions in a short time window. (*B*) In more complex brains, multiple interacting loops of increasing length improve prediction of more elaborate events and can make predictions in more complex environments at longer timescales. (*C*) After extensive training, the loops can sustain self-organized, long-lasting neuronal sequences without reliance on external cues and can, therefore, support cognitive operations such as memory, planning, and imagination. Progression of neuronal operations correlates with elapsed time (green arrows) irrespective of whether the operations are driven externally or internally. (Reprinted, with permission, from Buzsáki 2013.)

spinal cord/brainstem afferent–efferent connections, including the thalamocortical system, cortico-corticortical, corticocerebellar, corticostriatal, and corticolimbic loops. As in the simple system, the parallel loops operate to predict and secure effective outcomes of actions in ever more complex environments and at much longer timescales. These operations require storage of large amounts of past experience and the ability to evaluate and weigh the importance of simultaneously experienced modalities and compare them with similar occurrences in the past and heuristics responses of the organism. Many of these operations require that brain activities continue even after transient or more durable disappearance of sensory inputs and delaying the brain's deliberations in terms of effector outputs. We speculate that such a training process, a result of both phylogenetic and ontogenetic accumulation of experience, makes the longer, more complex neuronal loops "smarter" (i.e., more effective by enabling them to interpolate and extrapolate events surrounding the organism). Our second and critical assumption is that a self-organized computation emerges as a result of training the brain, which can proceed to deliver abstract goals by disengaging the effector networks from overt (muscular or vegetative) actions.

Activity in any system, including the brain, can be maintained by only two mechanisms: externally applied energy or intrinsic resources. By definition, a system that maintains its activity without an external drive must be driven by "self-organized" mechanisms. "Self-organized" is the term most often used in physics and engineering that refers to autonomously generated patterns. In neurophysiology, the term "spontaneous" activity has been in use for decades to refer to the same process. In cognitive science, the terms "internally generated" or "internally organized" refer to the same process. These terms

are synonymous. Therefore, we separate the functioning of complex brains into two modes: First, an "engaged" mode, when the immediate environment controls and feeds back on behavior even when the inputs are quite complex and actions need to be delayed, illustrated in Figure 1B. Second, a "disengaged" or detached mode, when processing continues independent of environmental inputs and behavioral responses of the animal. This input/output disengagement in the presence of continued computation of potential responses is what we refer to as "cognition." This framework may explain why damage to multiple loops of the brain has no or relatively minor impact on operations when input dimensions are limited and input–output gaps are short, but their role becomes evident when the organism needs to evaluate a large and complex parameter space and execute a response most beneficial to the organism's needs.

There are costs and advantages to the disengaged mode. The cost is that it takes greater neural resources to create the "virtual worlds"; in growing brains, these are necessary for processing internally generated patterns without reliance on immediate sensory feedback cues. In the past, Gibson and heirs developed ideas that have come to be known as "embodied cognition" (Gibson 1977; O'Regan and Noë 2001; Thompson and Varela 2001). These authors emphasize that the environment of an animal affords physical and sensory constraints (i.e., reduces the need to compute all possible circumstances) that ease the computational load (Gibson's "affordances"). Without real-world constraints, disengaged processing requires that the brain creates a well-endowed model of the world. With an internalized "virtual world" in the brain, the animal is capable of cognitive vicarious behavior; that is, it can perform "what if" scenarios in its head and anticipate the consequences of

alternative actions without actually testing them. With this mechanism at hand, the brain can create new knowledge not only by interacting with the outside world, but also through "vicarious" (imagined) experience. Disengaged processing, which we call "cognition," provides dramatic advantages in predicting the consequence of actual behavior in complex environments and at long timescales.

There are numerous examples of disengagement of brain activity from afferent inputs. A striking example is the "internalization" of brain rhythms. The first organized neocortical pattern in the developing mammalian brain is the emergence of transient periods of spindle-shaped rhythmic activity at $10-25$ Hz amid seconds- or tens-of-seconds-long silent periods ("tracé alternant"; Dreyfus-Brisac 1962; Anderson et al. 1985; Khazipov et al. 2004). In the newborn rat and premature human babies, all these early spindles are triggered by some kind of a movement, including isolated muscle twitches, limb jerks, and whole body startles (Khazipov et al. 2004). This organization is hypothesized to serve as a mechanism to construct an internal map of the body in the somatosensory system (Khazipov et al. 2004; Buzsáki 2006; Khazipov and Buzsáki 2010). Once the body map is formed ($5-10$ d postnatally in the rat and in the third trimester of pregnancy in humans), the dependence of brain activity on action-triggered reafferentation is no longer needed and spindles become part of sleep, the default organizing mode of the brain. This is a simple example of how brain patterns that are initially correlated with sensorimotor activity can disengage from external inputs, while sustaining their "spontaneous" activity patterns. Below we discuss two more examples to illustrate our disengagement hypothesis of cognition. First, we illustrate how neurons in the head direction system continue their coordinated activity during sleep. Second, we show how internally organized cell assembly sequences can support mental travel of memory and action planning, and how these functions emerge from spatial navigation in the physical environment.

SELF-ORGANIZED ATTRACTORS OF DIRECTION SENSE

The relationship between stimulus-driven and self-organized (or internally generated) activity is a recurring topic in neuroscience. It is often tacitly assumed that the information from the sensory world is "relayed" by the various thalamic nuclei to the neocortex where it is combined with top-down, internally generated cortical activity. How feed-forward, "labeled line" sensory signals in subcortical, thalamic, and cortical networks interact with self-generated ("spontaneous") activity is not well-understood. Part of the difficulty is technical, due largely to the high dimensionality of sensory signal attributes (e.g., intensity, color, orientation in case of vision; Rao and Ballard 1999; Engel et al. 2001; Varela et al. 2001; Buzsáki 2006).

Perhaps the simplest representation animals have is the sense of direction. It is simple in the sense that it has a single dimension, which is the head direction. Head-direction neurons fire robustly when the animal's head points in a specific direction (Ranck 1985; Taube et al 1990a,b; Taube 2007; McNaughton et al. 2006). Thus, the one-dimensional head-direction system offers an opportunity to understand the nature of external and internal interactions experimentally. The head-direction system consists of a largely serially connected brain network, which includes the brainstem, mammillary bodies, anterodorsal thalamic nucleus, postsubiculum, and entorhinal cortex (McNaughton et al. 2006; Sargolini et al. 2006; Taube 2007).

Numerous experiments have showed that head-direction neurons can be controlled by peripheral inputs, mainly the vestibular, visual, and ancillary afferents and the angular velocity of head rotation (Sharp et al. 2001; Taube 2007). How do these neurons behave in the absence of such inputs, when the brain is disengaged from the environment, as is the case during sleep? Computational models have long assumed that head-direction neurons with similar preferred directions fire together, while the other neurons are suppressed. According to these models, the temporally engaged group of head-direction neurons (i.e., a "hill of activity" or "an activity packet") moves on a virtual ring as the animal turns its head (Fig. 2; Skaggs et al. 1995; Redish et al. 1996; Burak and Fiete 2012; Knierim and Zhang 2012). However, experimental demonstration of the existence of internally organized neuronal populations endowed with such properties had to wait until recording technologies matured to be able to record simultaneous population of neurons and compare their dynamic interactions during both waking behavior and sleep. Recent experiments provide such evidence (Peyrache et al. 2015).

Not only do head-direction neurons continue to be active during sleep but their temporal sequential activity patterns in the waking mouse are preserved. At any time, the neuronal population was characterized by the presence of an activity packet that was consistent with its structure during waking: Neurons coding for nearby orientation continued to fire together. Therefore, a "virtual gaze" (i.e., the direction the mouse was "looking") could be reconstructed from firing of neurons in the anterodorsal thalamic nucleus and postsubiculum (Fig. 2). Although both firing rates and sequential firing patterns persisted in all brain states, the temporal dynamics of the drift of the activity packet varied. During REM sleep, when brain electrophysiology is strikingly similar to waking, the packet moved at the same speed as in the waking animal, essentially recapitulating the situation when the mouse moved its head in different direction in the waking animal. During non-REM sleep, the HD signal drifted at a 10 times faster speed, compared with waking and REM sleep, essentially following the faster dynamics of non-REM sleep described in cortical networks (Buzsáki 1989; Wilson and McNaughton 1994). Of course, no real head movement accompanied these changes in the sleeping animal; therefore, the temporal organization had to rely on internally generated mechanisms rather than external stimuli.

The observation that the neuronal code of the head-direction sense is preserved during sleep raises the possi-
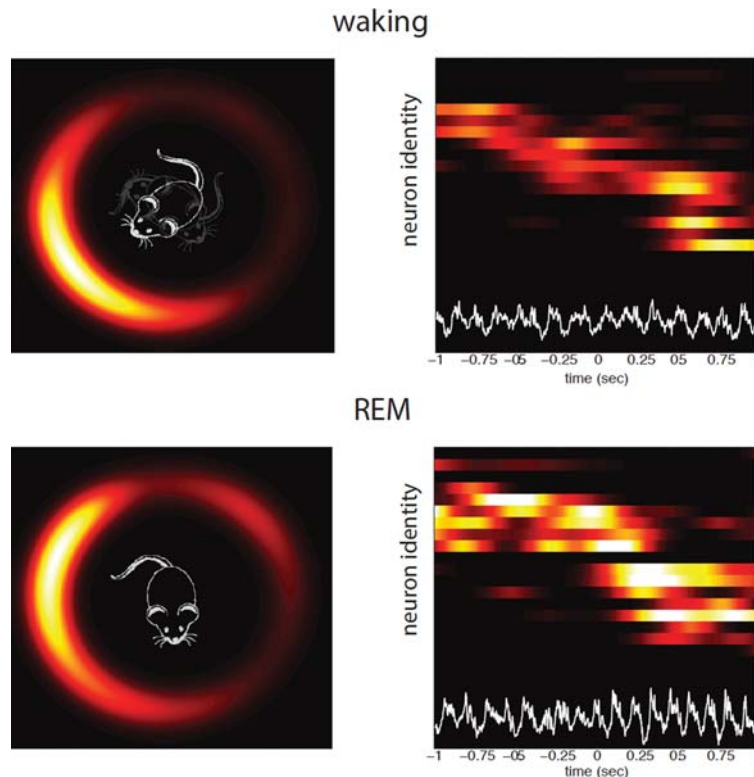
Figure 2. Persistence of head-direction cell assemblies in the anterodorsal nucleus of the thalamus during waking and sleep. Waking. Simultaneously recorded head-direction cells with different direction preferences during exploration (*right*). Each line is a neuron, increasing firing rates are represented by hot colors. Neurons are ordered according to their preferred head-direction. Note sequential activity of head-direction neurons (from *top* to *bottom*). *Bottom* trace (white) is local field potential recording from the hippocampus. (*Left*) Head-direction cell assembly activity displayed on a ring. As the head points to different directions, the hill of spiking activity moves on the ring accordingly. REM, same arrangement as in waking. Note continued sequentially changing activity of the head-direction cells despite the absence of head or body movement. (Reprinted, with permission, from Peyrache et al. 2015.)

bility that self-organized mechanisms continue to be at work also in the waking animal. Such an internal computation can amplify and enhance precision of the head-direction signal by filtering out, in the case of the head-direction, all additional information that is not compatible with a one dimensional angular signal. In the exploring animal, sensory inputs can be combined with the prediction of the internally generated head direction and rapidly adapt to reconfigured environments (Zugaro et al. 2003). In case of ambiguous or conflicting signals, self-organized mechanisms may generate the brain's "best guess" by interpolating across input signals or extrapolating the correct position of the head-direction vector from limited or ambiguous sensory information. Overall, these experimental findings in a simple but fundamentally important sense show that internally organized mechanisms are continuously used to improve the brain's interpretation of the external world (Fig. 1B).

FROM NAVIGATION IN THE PHYSICAL WORLD TO MENTAL TRAVEL

The hippocampal–entorhinal cortex has long been linked to both navigation and memory (Scoville and Mil-

ner 1957; O'Keefe and Nadel 1978). It has been proposed recently that neuronal mechanisms of memory and planning ("mental travel into the past and future"; Suddendorf and Corballis 1997, 2007) have evolved from mechanisms of navigation in the physical world (Buzsáki 2005; Buzsáki and Moser 2013). Below, we discuss that the neuronal algorithms underlying navigation in real and mental space are fundamentally the same. The key distinction is that during real world travel, external cues, such as landmarks, combined with self-motion cues aid in driving cell assembly sequences; when travel is internal, the brain must supply the supporting cues for navigating through a mental space.

To support mental travel effectively, a neural system evolved for navigation must develop the ability to self-generate temporally evolving cell assembly sequences and a capacity to store large quantities of seemingly unrelated episodes. It has been suggested that the evolution of the entorhinal–hippocampal system illustrates the phylogenetic continuity of navigation and memory and has the anatomical and physiological properties that make it especially suitable for meeting the above requirements (Buzsáki 2005; Buzsáki and Moser 2013).

SENSORY-DRIVEN ACTIVITY OF THE HIPPOCAMPAL NAVIGATION SYSTEM

The explicit spatial coordinates of an animal are encoded by a combination of a variety of neuron types, including "head-direction cells" of the thalamus and postsubiculum (Ranck 1985; Taube 2007), "place cells" in the hippocampus (O'Keefe and Nadel 1978), "grid cells" (Hafting et al. 2005), and "border cells" (Solstad et al. 2008) of the entorhinal cortex and surrounding structures. The exact functions of these cell types are yet to be determined, but it has been suggested that the periodically active fields of the grid cells provide a metric for the neural representation of space, in the same way that head-direction cells provide a directional reference frame, whereas border cells can assess allocentric distances in the discoverable environment by triangulation (Moser et al. 2014).

The "temporally evolving sequences" referred to above are, in navigational terms, paths. The paths represent a temporal sequence of neighboring locations on a map. Each map represents a specific environment. "Remapping" refers to the finding that the organization of place cells scramble from one environment to another (Muller and Kubie 1987; Kubie and Muller 1991). Because place cells that have neighboring firing fields will likely not have neighboring firing fields in a second environment, a single sequence of place cell firing not only codes a specific path, but also a specific environment. Recent findings, however, refine this strict "orthogonalized" view of remapping. Although most neurons have single place fields in a given environment, a small minority can have multiple fields and this same minority of neurons continues to fire in multiple environments. In fact, the distribution of the number of place fields per hippocampal neuron follows a lognormal distribution (Buzsáki and Mizuseki 2014). This small subset is not only more active in multiple environments, but their firing rates are higher, emit more spike bursts and their place fields are larger compared with those of the majority neurons. Importantly, the log-based rule organization applies to all cortical regions and multiple neuronal functions from synaptic strength distribution to macroscopic connectivity. The ever-active minority provides the brain's "best guess" and offers "good enough" solutions to get by in any situation. On the other hand, the majority of less active neurons comprises a large reservoir that can be mobilized to precisely distinguish one situation from another and incorporate novel ones as distinct (Buzsáki and Mizuseki 2014).

When the animal explores an environment, head-direction neurons, border cells, grid cells, and place cells continuously and coherently change their firing activity. Traditionally, two sources of information have been considered as driving factors of the temporal evolution of cell assembly sequences (Fig. 2). The first one is the constellation of the landmarks available as the animal navigates (O'Keefe and Nadel 1978). The second may derive from the ideothetic (i.e., body-derived) cues from the vestibular and proprioceptive systems and self-motion-driven optical flow from mostly local cues (McNaughton et al.

1996). Moreover, spatial firing does not require a rich set of landmarks. Landmarks can be impoverished, subsets removed, or the lights turned out, and firing will remain stable and robust (Muller and Kubie 1987; Quirk et al. 1990). These findings indicate that the brain is capable of interpolating between cues and extrapolating (generalizing or "pattern-completing") from the available external information.

However, even small insects, whose nervous systems possess many fewer neurons and are made of simpler circuits, are able to navigate using such strategies (Menzel et al. 1998). In principle, a small-sized network should be sufficient for rodents as well if navigation in a relatively fixed environment and range would be the only goal. Indeed, in laboratory experiments, the explored environment can be mapped at centimeter precision by just a dozen or so grid cells (Fyhn et al. 2007) or place cells (Wilson and McNaughton 1993), What advantage, then, do the complex networks of the entorhinal–hippocampal system of mammals offer? One can speculate that there are two advantages: first, to map specific features of multiple environments; second, that such self-organized networks (Fig. 1B) also have the ability to effectively interpolate and extrapolate from the available cues and hold information over long timescales, as well as use previous experience from similar situations in more complex environments. The ability to rapidly acquire multiple representations, to maintain separation of these representations, to hold the representations over long timescales, and to operate effectively on any, even in degraded conditions, are several of the robust navigational capacities in mammalian brains compared with simpler brains of species where navigation is accomplished by much smaller circuits.

SELF-ORGANIZED ACTIVATION OF THE HIPPOCAMPAL NAVIGATION SYSTEM

A second way of generating sequential firing of cell assemblies is by self-organization (Fig. 3), independent of cues from either the environment or the body. Such internally generated neural assembly sequences have been long assumed to be a necessary substrate of cognition for various mental operations such as memory recall, planning, and imagination (Hebb 1949). Indeed, perpetually changing hippocampal assembly sequences can be observed experimentally. For example, if a rat is trained to run in a wheel with approximately the same speed and facing in the same direction, the environmental and body-derived cues remain constant throughout the run. If such "clamped" behavior is part of a memory task to remember a previously made choice for planning for the appropriate future choice, hippocampal neurons display continuously changing assembly sequences (neuronal "trajectories"; Fig. 4). Importantly, several measures of the firing patterns of the internally organized sequences during the delay period, including the duration of the spiking activity of the neurons and the temporal relationship of their spikes relative to the reference theta oscilla-

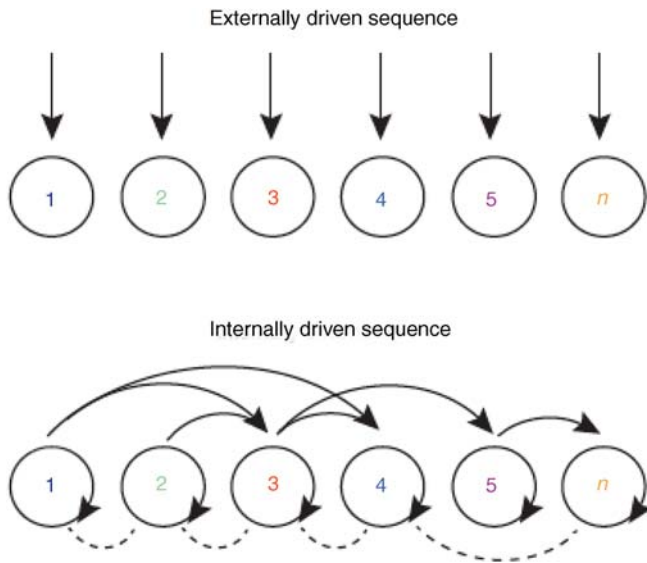**Figure 3.** Cell assembly sequences. Sequential activation of neuronal assemblies neurons ($1-n$) can be controlled by the changing constellation of environmental cues and/or proprioceptive information from the body (*top*). During cognitive activity, sequential activation is supported by self-organized patterning of assembly activity (*bottom*). Not only first-order (neighbor) but also higher-order (nonneighbor) connections can be represented in strongly connected recurrent networks.

tion cycle (O'Keefe and Recce 1993), are virtually identical with those of the place cells during translational behavior (Pastalkova et al. 2008). The implication of these observations is that the physiological mechanisms that govern the progression of cell assembly sequences in the hippocampus during navigation and memory are quite similar. Although theta oscillation is an obligatory mechanism for generating internal cell assembly sequences, environmental cues can maintain sequential place cell

activity even in the absence of theta rhythms (Wang et al. 2015). The cognitive relevance of self-organized sequential activity is emphasized by the observation that identical initial conditions (e.g., a left choice was rewarded) induces a similar assembly sequence each time, whereas different conditions (i.e., different memories) can give rise to uniquely different neuronal trajectories, which accurately predicts upcoming choices in the maze, including erroneous turns (Fig. 4). In accordance
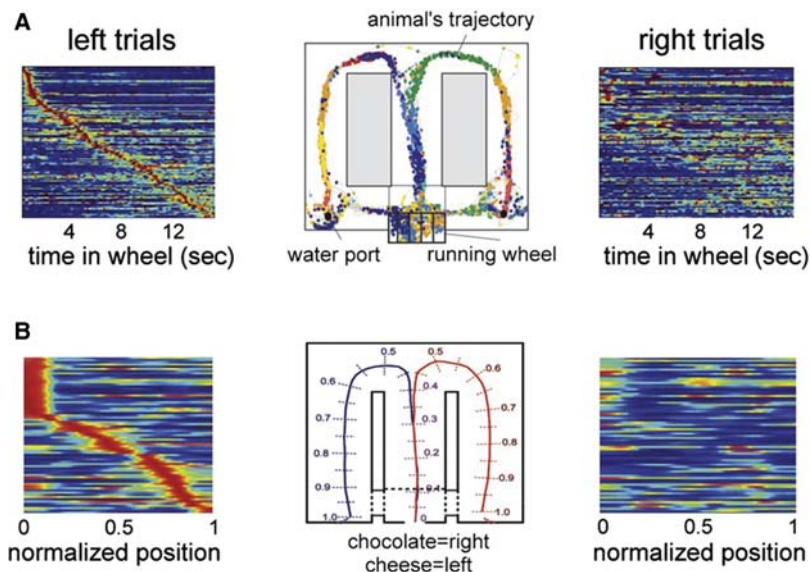

Figure 4. Internally generated assembly sequences during cognitive activity. (*A*) Sequential firing patterns of hippocampal neurons in a memory task. *Center*: Color-coded spikes (dots) of simultaneously recorded hippocampal CA1 pyramidal neurons. The rat was required to run in the wheel facing to the left during the delay between the runs in the maze. (*Left*) Normalized firing rate profiles of neurons during wheel running, ordered by the latency of their peak firing rates during left trials (each line is a single cell). (*Right*) Normalized firing rates of the same neurons during right trials. (*B*) Sequential firing patterns of prefrontal pyramidal cells in a working memory task. (*Middle*) Cheese odor or chocolate odor in the start area signals the availability of cheese or chocolate reward in the left or right goal area (position 1), respectively. Travel trajectories were linearized ($0-1$). (*Left*) Neurons were ordered by the location of their peak firing rates relative to the rat's position in the maze during left trials. Each row represents the position-dependent normalized firing rate of a single neuron. (*Right*) Normalized firing rates of the same neurons during right trials. (*A* and *B* are reprinted, with permission, from Pastalkova et al. 2008 and Fujisawa et al. 2008, respectively.)

with experiments in rodents, single-unit studies in human patients showed that the hippocampus and entorhinal cortex can generate numerous trajectories corresponding to different memory episodes (Gelbard-Sagiv et al. 2008). Subjects first viewed cinematic episodes that depicted short clips of famous people and characters. Several neurons responded selectively to only one or a few of the episodes. In the critical part of the experiment, subjects were asked to recall the characters and scenes without any cues. Remarkably, the neurons that fired slightly before free recall (i.e., a verbal report) corresponded to the same neurons that were activated while watching the cinematic episodes in the learning phase (Gelbard-Sagiv et al. 2008). These results support the view that the neuronal assembly sequences that encode episodes are the same that are active during the mental representation of those episodes (Tulving and Thomson 1973).

Overall, these experiments illustrate that perpetually changing cell assembly sequences can be induced in the hippocampus independent of external cues (Fig. 1C), and such disengagement can support cognitive performance. Generation of neural sentences is not confined to the hippocampal system. In the medial prefrontal cortex of the rat, long neuronal sequences can also reliably differentiate between right and left trajectories in the maze before making a choice, with individual neurons active only for a short duration (Fig. 4; Fujisawa et al. 2008). The possibility exists that choice specific neuronal trajectories in the hippocampus are selected by inputs from the medial prefrontal cortex. Similar neuronal trajectories have been also been described in the parietal cortex (Harvey et al. 2012). In summary, cortical circuits can produce multitudes of unfolding assembly sequences in two different ways: either by responding to environmental/idiothetic stimuli, when such inputs are available, or by generating them internally.

Analogous to spatial navigation, two forms of declarative memory can be distinguished (Squire 1992). Similar to the defining locations of the allocentric map, semantic memory explicitly defines living things, objects, facts, and events of the surrounding world independent from temporal context (Tulving and Thomson 1972; Squire

1992). Episodic memory, in contrast, allows learning and recalling self-referenced (first-person) experiences in the context of both space and subjective time (Tulving and Thomson 1972) for planning actions (Tulving 2002; Buckner 2010), similar to linking together location sequences during exploration. Acquiring a spatial map requires repeated exploration of the landmarks by self-referenced navigation. Similarly, semantic knowledge is generated by repeated episodic encounters until the items lose their spatiotemporal context (Fig. 5; Squire 1992; Eichenbaum et al. 1999; Buzsáki 2005).

On the one hand, declarative memories may be recalled by supplying a variety of real world cues that induce an association. This is, roughly, equivalent to an animal exposed to a few cues from an environment recalling an entire environmental representation. On the other hand, declarative memories may be recalled without real world cue triggers. For example, one may recall an episodic memory or a fact by association with another memory. Such vicarious memory recall is analogous to vicarious navigational recall. We propose they have identical mechanisms.

INTERNALLY GENERATED NEURONAL PATTERNS AND COVERT OUTPUTS IN THE CEREBRAL CORTEX

We hypothesize that internally generated patterns are a common feature of cortical organization which can function as a neuronal substrate to bridge long delays between inputs and action. Connectivity patterns found in neocortex support this notion. Superficial cortical neuronal layers may be viewed as a substrate for inducing such covert outputs. While deep layer cortical neurons project to action systems of the thalamus, colliculi, brainstem, and even the spinal cord, the superficial layer can be viewed as an extra, added loop to the deep layer with subcortical outputs (Shepherd 2013). A novel neuron loop that is added to the sensorimotor short loop is the extensive associational areas, particular the prefrontal cortex. Thus, perhaps it is not surprising that the anatomical
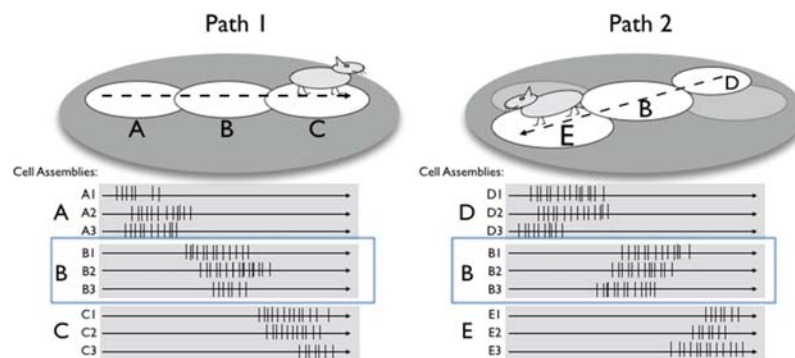

Figure 5. Place cell firing represents both episodic and semantic memory. Paths 1 and 2 are accompanied by cell assembly sequences A−B−C and D−B−E. Internally generated firing of either of these sequences is equivalent to a specific episodic memory: movement along Path 1 or Path 2. After a series of paths that activate cell assembly B, B alone is not associated with a particular path: it explicitly represents a location. In this representation, the activation of cell assembly B, alone, is a semantic fact, location B.

afferent–efferent connectivity and internal organization of the agranular motor cortex and the prefrontal cortex are so remarkably similar (Nimchinsky et al. 1995; Gabbott et al. 2005). The notion is that there are two types of motor output: the familiar "engaged" output, where motor cortical areas activate motor neurons and produce behavior; and vicarious motor output, where an internal analog of motor action is generated, with output remaining entirely within the brain. If the brain contains a sufficient model of the self and world, this type of imagined or vicarious motor action can be examined internally for potential consequences. This vicarious action is what we call a "cognitive process." Such vicarious action can be used for navigation, as in the hippocampal navigational system described earlier, or for other forms of motor behaviors. At the physiological level such cognitive operations require that the self-organized neuronal assembly sequences are propelled forward via an optimal path or by covert trial and error evaluating multiple possible outcomes, solving detours, and finding shortcuts (Gupta et al. 2012), much the same way as we navigate through locomotion in the real world in the presence of environmental feedback.

MEMORY AND THE OPTIMIZATION
OF BEHAVIOR

As discussed earlier, a fundamental goal of many neuronal circuits and the brain as a whole is to predict the future. However, predicting the future is possible only after the organism has already experienced the consequences of its actions and stored the outcomes of those successful and failed actions in memory. Indeed, a barrage of recent papers document the recognition that structures that have been traditionally viewed as memory systems are inseparable parts of planning, imagining, and action systems (Buckner and Carroll 2007; Schacter and Addis 2007; Lisman and Redish 2009; Buckner 2010). What appeared to be memory of the past in the sequential activity of cell assembly sequences, equally well reflected the planned future action of the animal, including commission errors, indicating that the action outcome is the consequence of the brain's *interpretation* of the past experience (Fujisawa et al. 2008; Pastalkova et al. 2008). Numerous sleep "replay" experiments in both rodents and humans show that sleep patterns are far from random but, instead, are reminiscent of waking when the neuronal trajectories evolve under the influence of sensory inputs (Wilson and McNaughton 1994; Nádasdy et al. 1999; cf. Diekelmann and Born 2010). Transfer operations, such as the sharp wave ripples of the hippocampus, which were originally suggested to consolidate and move transient memories from the hippocampus to neocortex (Buzsáki 1989; Wilson and McNaughton 1994), have been shown to generate spike sequences that predict to-be-visited places and delayed actions or a chaining of past and expected events (Diba and Buzsáki 2007; Davidson et al. 2009; Gupta et al. 2010; cf., Carr et al. 2011; Buzsáki and Lopes da Silva

2012). Imaging experiments in humans have repeatedly found that imagining and planning invariably activated structures previously categorized as parts of the memory system (Buckner and Carroll 2007; Schacter and Addis 2007; Buckner 2010). The emerging field of "active sensing" (Katz 1989; Engel et al. 2013; Loeb and Fishel 2014) also emphasizes the primacy of action in acquiring meaning for sensation and perception (Schroeder et al. 2010). These novel (or "revisited") views break away from the tradition of "boxing" man-invented terms into particular brain structures and strategies to identify physiological boundaries for preconceived mental boundaries (Buzsáki 2010). Instead, they emphasize *interactions* of brain systems as the fundamental operations that give rise to behavioral entities that are hard to categorize strictly as memory or planning, overt or covert, conscious or "autopilot." Regarding the self-organized brain activity as the fundamental brain operation and viewing the brain's responses to external perturbations as secondary actions offers a new paradigm, in which "*meaning*" is acquired by matching preexisting neuronal patterns to action-perception (Buzsáki 2006). In support of this conjecture, experiments in developing ferrets show that the similarity between spontaneous and evoked activity increases progressively with age and is specific to responses evoked by natural scenes, suggesting that internal models result from the adaptation of brain states to the statistics of the surrounding world (Fiser et al. 2004). By extension, the construction of an internal representation of the self as a special entity can be accomplished by the reprocessing of internal data simultaneously with information about the environment.

A critical issue that has to be explored in future experiments is the significance of the duration of internally organized sequences for cognition. There are numerous brain operations, especially in the motor system, which require sequential activity of neuronal firing, yet we are not aware of the consequences of those operations and, thus, they have not been considered as cognitive. Neuronal events that are too short (<500 msec; Libet 2004) tend not to evoke awareness. These likely include the short (<100 msec) sharp wave ripple events of the hippocampus (Buzsáki 1989). Yet, these events may be regarded as "subconscious rehearsals" for future actions or priming for memory recall.

Although ample evidence exists to support the view that internally organized cell assembly sequences are the fundamental basis of mental operations, the "clutch" operation is not well understood. However, currently existing and future technologies can be deployed to test the hypothesis that mental travel and planning occur in real time, similar to navigation. By reading out brain signals and connecting them to actuators, such as a navigating robot, thought control of the robot's movements should be possible. Experiments performed within the brain–machine interface paradigm (Donoghue 2002; Fetz 2007; Nicolelis and Lebedev 2009) show the feasibility of such outcomes. Furthermore, neuronal patterns that can be read out during sleep can also be exploited to compare them with waking operations and manipulate

them to establish whether internally generated patterns in the absence of conscious awareness can reflect waking-related neuronal trajectories. This paper highlights recent work in animal physiology that shows the implications of self-organized, internally generated neuronal assembly sequences. These findings support an emerging conception of the widespread importance of internal models as covert, brain-based testing grounds for efficient optimization of behavior.

REFERENCES

Anderson CM, Torres F, Faoro A. 1985. The EEG of the early premature. *Electroencephalogr Clin Neurophysiol* **60:** 95–105.

Buckner RL. 2010. The role of the hippocampus in prediction and imagination. *Annu Rev Psychol* **61:** 27–48.

Buckner RL, Carroll DC. 2007. Self-projection and the brain. *Trends Cogn Sci* **11:** 49–57.

Burak Y, Fiete IR. 2012. Fundamental limits on persistent activity in networks of noisy neurons. *Proc Natl Acad Sci* **109:** 17645–17650.

Buzsáki G. 1989. Two-stage model of memory trace formation: A role for "noisy" brain states. *Neuroscience* **31:** 551–570.

Buzsáki G. 2005. θ rhythm of navigation: Link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* **15:** 827–840.

Buzsáki G. 2006. *Rhythms of the brain*. Oxford University Press, New York.

Buzsáki G. 2010. Neural syntax: Cell assemblies, synapsembles, and readers. *Neuron* **68:** 362–385.

Buzsáki G. 2013. Cognitive neuroscience: Time, space and memory. *Nature* **497:** 568–569.

Buzsáki G, Lopes da Silva FL. 2012. High frequency oscillations in the intact brain. *Prog Neurobiol* **98:** 241–249.

Buzsáki G, Mizuseki K. 2014. The log-dynamic brain: How skewed distributions affect network operations. *Nat Rev Neurosci* **15:** 264–278.

Buzsáki G, Moser EI. 2013. Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat Neurosci* **16:** 130–138.

Carr MF, Jadhav SP, Frank LM. 2011. Hippocampal replay in the awake state: A potential substrate for memory consolidation and retrieval. *Nat Neurosci* **14:** 147–153.

Davidson TJ, Kloosterman F, Wilson MA. 2009. Hippocampal replay of extended experience. *Neuron* **63:** 497–507.

Diba K, Buzsáki G. 2007. Forward and reverse hippocampal place-cell sequences during ripples. *Nat Neurosci* **10:** 1241–1242.

Diekelmann S, Born J. 2010. The memory function of sleep. *Nat Rev Neurosci* **11:** 114–126.

Donoghue JP. 2002. Connecting cortex to machines: Recent advances in brain interfaces. *Nat Neurosci* Suppl: 1085–1088.

Dreyfus-Brisac C. 1962. The electroencephalogram of the premature infant. *World Neurol* **3:** 5–15.

Eichenbaum H, Dudchenko P, Wood E, Shapiro M, Tanila H. 1999. The hippocampus, memory, and place cells: Is it spatial memory or a memory space? *Neuron* **23:** 209–226.

Engel AK, Fries P, Singer W. 2001. Dynamic predictions: Oscillations and synchrony in top-down processing. *Nat Rev Neurosci* **2:** 704–716.

Engel AK, Maye A, Kurthen M, König P. 2013. Where's the action? The pragmatic turn in cognitive science. *Trends Cogn Sci* **17:** 202–209.

Fetz EE. 2007. Volitional control of neural activity: Implications for brain computer interfaces. *J Physiol* **579:** 571–579.

Fiser J, Chiu C, Weliky M. 2004. Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature* **431:** 573–578.

Frank LM, Brown EN, Wilson M. 2000. Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron* **27:** 169–178.

Fujisawa S, Amarasingham A, Harrison MT, Buzsáki G. 2008. Behavior-dependent short-term assembly dynamics in the medial prefrontal cortex. *Nat Neurosci* **11:** 823–833.

Fyhn M, Hafting T, Treves A, Moser MB, Moser EI. 2007. Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* **446:** 190–194.

Gabbott PL, Warner TA, Jays PR, Salway P, Busby SJ. 2005. Prefrontal cortex in the rat: Projections to subcortical autonomic, motor, and limbic centers. *J Comp Neurol* **492:** 145–177.

Gelbard-Sagiv H, Mukamel R, Harel M, Malach R, Fried I. 2008. Internally generated reactivation of single neurons in human hippocampus during free recall. *Science* **322:** 96–101.

Gibson J. 1977. The theory of affordances. In *Perceiving, acting, and knowing: Toward an ecological psychology* (ed. Shaw R, Bransford J), pp. 62–82. Lawrence Erlbaum Associates, Hillsdale, NJ.

Gupta AS, van der Meer MA, Touretzky DS, Redish AD. 2010. Hippocampal replay is not a simple function of experience. *Neuron* **65:** 695–705.

Gupta AS, van der Meer MA, Touretzky DS, Redish AD. 2012. Segmentation of spatial experience by hippocampal theta sequences. *Nat Neurosci* **15:** 1032–1039.

Hafting T, Fyhn M, Molden S, Moser MB, Moser EI. 2005. Microstructure of a spatial map in the entorhinal cortex. *Nature* **436:** 801–806.

Harvey CD, Coen P, Tank DW. 2012. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484:** 62–68.

Hebb DO. 1949. *The organization of behavior*. Wiley, New York.

Immordino-Yang MH, Christodoulou JA, Singh V. 2012. Rest is not idleness: Implications of the brain's default mode for human development and education. *Perspect Psychol Sci* **7:** 352–364.

Katz D. 1989. *The world of touch*. Lawrence Erlbaum Associates, Hillsdale, NJ.

Khazipov R, Buzsáki G. 2010. Early patterns of electrical activity in the developing cortex. In *Oxford handbook of developmental behavioral neuroscience* (ed. Blumberg MS, Freeman JH, Robinson SR), pp. 161–177. Oxford University Press, Oxford.

Khazipov R, Sirota A, Leinekugel X, Holmes GL, Ben Ari Y, Buzsáki G. 2004. Early motor activity drives spindle bursts in the developing somatosensory cortex. *Nature* **432:** 758–761.

Killingsworth MA, Gilbert DT. 2010. A wandering mind is an unhappy mind. *Science* **330:** 932.

Knierim JJ, Zhang K. 2012. Attractor dynamics of spatially correlated neural activity in the limbic system. *Annu Rev Neurosci* **35:** 267–285.

Kubie JL, Muller RU. 1991. Multiple representations in the hippocampus. *Hippocampus* **1:** 240–242.

Libet B. 2004. *Mind time. The temporal factor in consciousness*. Harvard University Press, Cambridge, MA.

Lisman J, Redish AD. 2009. Prediction, sequences and the hippocampus. *Philos Trans R Soc Lond B Biol Sci* **364:** 1193–1201.

Loeb GE, Fishel JA. 2014. Bayesian action&perception: Representing the world in the brain. *Front Neurosci* **8:** 341.

Mar RA, Mason MF, Litvack A. 2014. How daydreaming relates to life satisfaction, loneliness, and social support: The importance of gender and daydream content. *Conscious Cogn* **21:** 401–407.

McNaughton BL, Barnes CA, Gerrard JL, Gothard K, Jung MW, Knierim JJ, Kudrimoti H, Qin Y, Skaggs WE, Suster M, et al. 1996. Deciphering the hippocampal polyglot: The hippocampus as a path integration system. *J Exp Biol* **199:** 173–185.

McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser MB. 2006. Path integration and the neural basis of the 'cognitive map'. *Nat Rev Neurosci* **7:** 663–778.

Menzel R, Geiger K, Joerges J, Muller U, Chittka L. 1998. Bees travel novel homeward routes by integrating separately acquired vector memories. *Anim Behav* **55:** 139–152.

Moser EI, Roudi Y, Witter MP, Kentros C, Bonhoeffer T, Moser MB. 2014. Grid cells and cortical representation. *Nat Rev Neurosci* **15:** 466–481.

Muller RU, Kubie JL. 1987. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J Neurosci* **7:** 1951–1968.

Nádasdy Z, Hirase H, Czurkó A, Csicsvari J, Buzsáki G. 1999. Replay and time compression of recurring spike sequences in the hippocampus. *J Neurosci* **19:** 9497–9507.

Nicolelis MA, Lebedev MA. 2009. Principles of neural ensemble physiology underlying the operation of brain–machine interfaces. *Nat Rev Neurosci* **10:** 530–540.

Nimchinsky EA, Vogt BA, Morrison JH, Hof PR. 1995. Spindle neurons of the human anterior cingulate cortex. *J Comp Neurol* **355:** 27–37.

O'Keefe J, Nadel L. 1978. *The hippocampus as a cognitive map.* Oxford University Press, New York.

O'Keefe J, Recce ML. 1993. Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* **3:** 317–330.

O'Regan JK, Noë A. 2001. A sensorimotor account of vision and visual consciousness. *Beh Brain Sci* **25:** 883–975.

Pastalkova E, Itskov V, Amarasingham A, Buzsáki G. 2008. Internally generated cell assembly sequences in the rat hippocampus. *Science* **321:** 1322–1327.

Peyrache A, Lacroix MM, Petersen PC, Buzsaki G. 2015. Internally organized mechanisms of the head-direction sense. *Nat Neurosci* (in press).

Quirk GJ, Muller RU, Kubie JL. 1990. The firing of hippocampal place cells in the dark depends on the rat's recent experience. *J Neurosci* **10:** 2008–2017.

Ranck J. 1985. In *Electrical activity of archicortex* (ed. Buzsáki G, Vanderwolf CH). Akadémiai Kiadó, Budapest.

Rao RP, Ballard DH. 1999. Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* **2:** 79–87.

Redish AD, Elga AN, Touretzky DS. 1996. A coupled attractor model of the rodent head direction system. *Netw Comput Neural Syst* **7:** 671–685.

Sargolini F, Fyhn M, Hafting T, McNaughton BL, Witter MP, Moser MB, Moser EI. 2006. Conjunctive representation of position, direction, and velocity in entorhinal cortex. *Science* **312:** 758–762.

Schacter DL, Addis DR. 2007. Constructive memory: The ghosts of past and future. *Nature* **445:** 27.

Schroeder CE, Wilson DA, Radman T, Scharfman H, Lakatos P. 2010. Dynamics of active sensing and perceptual selection. *Curr Opin Neurobiol* **20:** 172–176.

Scoville WB, Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* **20:** 11–21.

Sharp PE, Blair HT, Cho J. 2001. The anatomical and computational basis of the rat head-direction cell signal. *Trends Neurosci* **24:** 289–294.

Shepherd GM. 2013. Corticostriatal connectivity and its role in disease. *Nat Rev Neurosci* **14:** 278–291.

Skaggs WE, Knierim JJ, Kudrimoti HS, McNaughton BL. 1995. A model of the neural basis of the rat's sense of direction. *Adv Neural Inf Process Syst* **7:** 173–180.

Solstad T, Boccara CN, Kropff E, Moser MB, Moser EI. 2008. Representation of geometric borders in the entorhinal cortex. *Science* **322:** 1865–1868.

Squire LR. 1992. Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychol Rev* **99:** 195–231.

Suddendorf T, Corballis MC. 1997. Mental time travel and the evolution of the human mind. *Genet Soc Gen Psychol Monogr* **123:** 133–167.

Suddendorf T, Corballis MC. 2007. The evolution of foresight: What is mental time travel, and is it unique to humans? *Behav Brain Sci* **30:** 299–313.

Taube JS. 2007. The head direction signal: Origins and sensory-motor integration. *Annu Rev Neurosci* **30:** 181–207.

Taube JS, Muller RU, Ranck JB Jr. 1990a. Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *J Neurosci* **10:** 420–435.

Taube JS, Muller RU, Ranck JB Jr. 1990b. Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *J Neurosci* **10:** 436–447.

Thompson E, Varela F. 2001. Radical embodiment: Neural dynamics and consciousness. *Trends Cogn Sci* **5:** 418–425.

Tulving E. 2002. Chronesthesia: Conscious awareness of subjective time. In *Principles of frontal lobe function* (ed. Stuss DT, Knight RC), pp. 311–325. Oxford University Press, New York.

Tulving E, Thomson D. 1973. Encoding specificity and retrieval processes in episodic memory. *Psychol Rev* **80:** 352–373.

Tulving E, Donaldson W, Bower GH. 1972. *Organization of memory.* Academic Press, New York.

Vanderwolf CH. 1969. Hippocampal electrical activity and voluntary movement in the rat. *Electroencephalogr Clin Neurophysiol* **26:** 407–418.

Varela F, Lachaux JP, Rodriguez E, Martinerie J. 2001. The brainweb: Phase synchronization and large-scale integration. *Nat Rev Neurosci* **2:** 229–239.

Wang Y, Romani S, Lustig B, Leonardo A, Pastalkova E. 2015. Theta sequences are essential for internally generated hippocampal firing fields. *Nat Neurosci* **18:** 282–288.

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

Wilson MA, McNaughton BL. 1994. Reactivation of hippocampal ensemble memories during sleep. *Science* **265:** 676–679.

Wood ER, Dudchenko PA, Robitsek RJ, Eichenbaum H. 2000. Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron* **27:** 623–633.

Zugaro MB, Arleo A, Berthoz A, Wiener SI. 2003. Rapid spatial reorientation and head direction cells. *J Neurosci* **23:** 3478–3482.