Memory Systems and Neural Dynamics

86

Sam McKenzie and György Buzsáki

Contents

Abstract

Philosophers, psychologists, and biologists have long pondered how experiences mark the mind. It is clear that there is no one seat of memory. Within each memory system there are multiple biological mechanisms for information storage. This chapter focuses on how memory could be supported by coincidence of neural activity, changes in synaptic strength, and synchronization of networks. Another major theme in memory research is the localization and specialization of function by different memory systems which can be revealed through appropriate

S. McKenzie
The Neuroscience Institute, School of Medicine and Center for Neural Science, New York University, New York, NY, USA
e-mail: mckens04@nyumc.org

G. Buzsáki (✉)
The Neuroscience Institute, School of Medicine and Center for Neural Science, New York University, New York, NY, USA

Department of Neuroscience and Physiology, School of Medicine, New York University, New York, NY, USA
e-mail: gyorgy.buzsaki@nyumc.org

behavioral tests. Recent technological advances allow for powerful new ways to record and perturb neural circuits to test the veracity of long-held theories about the relationship between biology and cognition.

The fundamental challenge for a memory system is to recognize past experiences that have utility in guiding action. This is difficult because no two situations are identical, and therefore it is necessary for the brain to match which events and stimuli should be grouped together due to common meaning. This is the computation that occurs upon meeting someone new and thinking, "You remind me of an old friend." From this flash of familiarity, recollections of past experiences shape the interaction with the stranger and, in doing so, bias how this new person is seen and which aspects of the experience become committed to memory. This everyday example illustrates the profound influence memory has on learning, behavior, and even perception.

There are several ways to seek an explanation for how the brain uses the past to bias future behavior. One is to design experiments that decrease the probability of an event's occurrence, such as removing a part of the brain and observing memory loss. These loss-of-function experiments elucidate the necessity of a specific phenomenon for the occurrence of another. Other experiments are designed to cause something to happen, such as the delivery of a drug that enhances memory performance. These experiments inform us as to the sufficiency that the presence of one event is enough to observe another. Gain of function and loss of function experiments may perturb the system in unexpected ways, and therefore it is also important to observe the brain with minimal intervention. In such observations, experimenters seek to identify variables that covary, such as the activity of a neuron correlating with some stimulus. A final and important method for understanding the link between biology and cognition is computational and mathematical modeling of neural systems (Hopfield 1995; Marr 1971). These more abstract analyses can offer predictions that, if validated, lead credence to the model from which the prediction was derived, which may provide insight into processes beyond the current limit of experimental validation. Furthermore, there may be many biologic solutions to the same computational problem and modeling efforts help to identify necessary and sufficient overarching principles that unite seemingly disparate observations.

The convergence of these lines of evidence – necessity, sufficiency, correlation, and modeling – has shown that there are multiple memory systems that specialize in storing and processing different kinds of information. In each of these systems, information is thought to be stored in changes in communication efficacy between neurons. The temporal organization of neural activity required for such changes in synaptic strength and for efficient communication is achieved by coordination of brain rhythms across brain regions. This chapter will handle these broad topics with an emphasis on the technological advances in behavioral testing, genetics, imaging, and recording that allowed each new scientific advance. Very few contemporary concepts in memory research are brand new, and therefore it is essential to first discuss the insights of the pioneers that have shaped the current understanding of the neurobiology of memory.

Brief History

Philosophers and scholars have long considered the important role of memory in defining the self, but experimental investigation of memory began relatively recently. Working in the late nineteenth century, Hermann Ebbinghaus tested properties of his own memory on a set of standardized, neutral stimuli, lists of nonsense stimuli. This work introduced the concept of a forgetting curve, that memory fades over time, and savings, that memory gets better with practice. Another early pioneer was Ribot who amassed mostly anecdotal evidence that head trauma following learning can cause forgetting and hypothesized, correctly, the existence of a memory consolidation process that is required for stabilization of memory after learning. Working in the United States, William James was a highly influential psychologist who wrote about the existence of multiple memory systems differentiating short-term and long-term memory and distinguishing those from habits. These concepts, derived in part through introspection and in part through experimentation, still guide much of modern research into the organization of memory.

The early behavioralists working in animals also have enduring influence (Fig. 1). Ivan Pavlov introduced the concept of conditioning in which unconditioned stimuli (US), which are inherently appetitive or aversive and elicit an innate unconditioned response (UR), become associated with neutral, conditioned stimuli (CS) if presented with an appropriately timed presentation. After such classical conditioning, presentation of the CS can elicit a conditioned response (CR), which in Pavlov's famous experiments was the salivation of a dog (CR) in response to the ringing of a bell (CS) that had been paired with food (US). Pavlov also showed that repeated presentation of the CS without the US causes extinction, which is the cessation of the condition response. Burrhus F. Skinner extended this line of work to highlight the important role that behavior has on learning and introduced the concept that behaviors that lead to positive outcome, or those that prevent a negative outcome, become reinforced and repeated; behaviors that do not lead to rewards, or those that lead to punishment, become weakened. This association of action with outcome is known as operant conditioning and much behavioral testing in animals is conducted in operant chambers that allow precisely timed action/outcome contingencies.
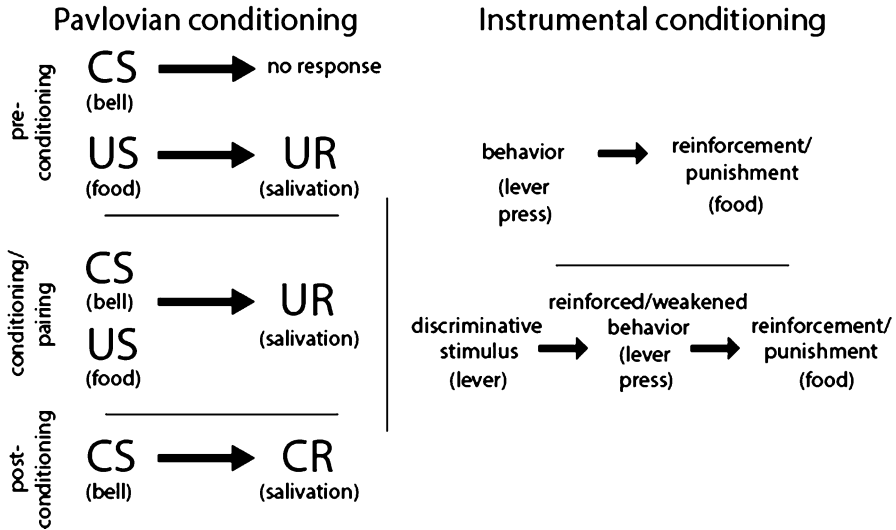
Fig. 1 Pavlovian and instrumental conditioning. In Pavlovian conditioning, the subject initially shows no response to the CS and an innate response (*UR*) to the unconditioned stimulus. After passive pairing, the CR can be elicited by the CS. Instrumental conditioning emphasizes the ability for punishments and reinforces to strengthen and weaken the behaviors that led to the salient outcome. Conditioned behaviors are cued by the predictive, discriminative stimuli

Behavioralism was fundamentally challenged by Tolman's theories of latent learning, the learning associations in the absence of reinforcement. Latent learning can be revealed through its influence on subsequent instrumental or classical conditioning, which he first showed by showing that rats exposed to mazes in the absence of reward can use acquired information about maze layout to guide subsequent behavior. This observation inspired Tolman's theory that the mind is able to create cognitive maps of space, which was ridiculed at the time for its inability to relate to behavior.

Around the turn of the twentieth century, Santiago Ramon y Cajal proposed, through remarkable insight, that neurons are separate entities that communicate with one another with a connection strength that can be modified through experience to form memories. Nearly 50 years later, Donald Hebb built upon this theory by hypothesizing that this anatomical change occurs due to coincident activity among pairs of neurons. He also introduced the concept of a cell assembly, a group of neurons that due to these strong connections become sequentially active in what he called phase sequences. Evidence for Ramon y Cajal and Hebb's postulated memory mechanism would have to wait until the seminal studies of Timothy Bliss and Terje Lømo working in the rabbit hippocampus in the 1970s and Eric Kandel working in the *Aplysia* sea slug around the same time. Together they showed that synaptic strengths can change as a function of neural activity and that such synaptic changes are associated with changes in behavior. This latter point was demonstrated in the case of the *Aplysia* in the study of the gill-withdrawal reflex in which changes in the

synaptic connectivity between sensory and motor neurons predict the classical conditioning properties of the withdrawal of the gill and siphon (CR) due to stimulation of the siphon (CS) after pairing with a tail shock (US). Today, the primary and nearly ubiquitous model for information storage in the brain is activity-induced changes in synaptic communication.

Multiple Memory Systems in the Human

Though psychologists appreciated the existence of multiple types of memory, biologic dissociations of memory systems were more controversial until about 50 years ago. In the 1930s and 1940s, Karl Lashley embarked on a research program to find the physical location of memory, the engram, and systematically removed brain areas and tested maze learning in rodents. He showed that the larger the lesion, the worse the memory deficit and falsely concluded that memory cannot be localized but is distributed across the entire brain and that all brain regions can carry out the function of any other.

Neurological examination of amnesiac patients with brain damage has highlighted instead the functional specialization of different brain structures. The first hints that the hippocampus, an evolutionarily conserved structure located in the medial temporal lobe (Fig. 2), plays a special role in memory emerged in the late nineteenth and early twentieth century, but it was Brenda Milner's seminal work on patient Henry Molaison (H.M.) in the 1950s that ignited decades of research as to how the hippocampus contributes to the conscious recollection of prior experience. As an experimental treatment for intractable epilepsy, H.M. received bilateral resectomy of the hippocampus, amygdala, and surrounding medial temporal lobe cortices including the entorhinal cortex, parahippocampal cortex, and piriform cortex (Scoville and Milner 1957). After the operation, Milner noted that H.M. could "no longer recognize the hospital staff nor find his way to the bathroom, and he seemed to recall nothing of the day-to-day events of his hospital life" (Corkin 2002).

Over 50 years of research on H.M has shown that he was incapable of forming new memories for stimuli of all modalities and a comorbid inability to recall information up to 11 years prior to his surgery but childhood memories remained intact. Subsequent research on other patients with brain damage limited to the hippocampus has confirmed combined memory loss for prior events, known as retrograde memory impairment, alongside an inability to form new memories, known as an anterograde memory deficit. Therefore, the hippocampus is necessary for both storing old episodic memories and making new ones (Tulving 1972). The generality of the deficit across modalities with nearly intact perception are important to keep in mind, as deficits in perception would obviously cause subsequent retrieval deficits.

Clearly, we possess the ability to recall many facts without any knowledge of where or when that information was first learned. This type of memory is known as semantic memory. An impressive dissociation between episodic and semantic memory has been reported by Vargha-Khadem in a 1997 case study of three people with
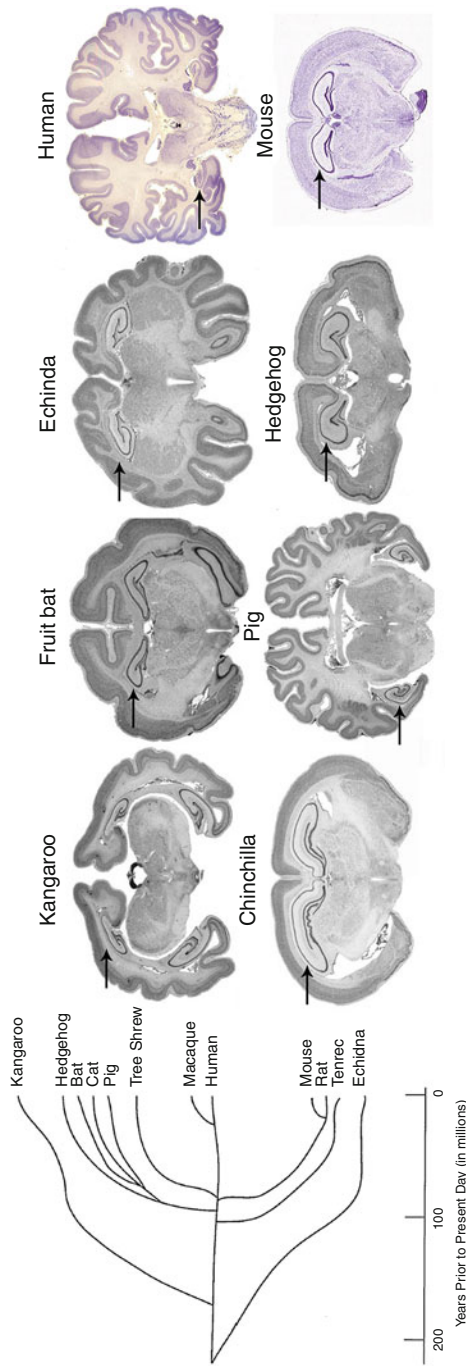
Fig. 2 Coronal slices of mammalian brains demonstrate the phylogenetic preservation of the gross anatomy of the hippocampus (*arrow*) (Figure modified from Manns and Eichenbaum (2006). Mouse brain from http://mouse.brain-map.org/. Other histological images originally from www.brainmuseum.org/)

childhood damage limited to the hippocampus. Remarkably, these patients attended mainstream school, were competent in reading and writing, had normal vocabulary, and showed average performance on a test of verbal IQ that explicitly tests for common-knowledge semantic information. In stark contrast, the same patients showed profound amnesia for everyday events and were unable to navigate familiar environments or orient in time; the hippocampal damage caused a selective episodic memory deficit. Episodic memory is also impaired in subjects with damage to the hippocampus incurred during adulthood, though there is debate as to whether semantic memory is also affected. There are also divergent findings as to whether episodic memories remain dependent on hippocampal function for all times or just within a period of months or years after learning, though there is a consensus that retrieval of recently formed episodic memories depends upon the hippocampus and semantic memories can be formed in the absence of hippocampal function.

It is important to emphasize that not all mnemonic faculties are impaired after hippocampal damage or even medial temporal lobe damage. H.M. had preserved motor learning as demonstrated by an intact ability to trace images while looking in the mirror, a task that requires inversion of the normal rules to eye-hand coordination. H.M. could carry on a conversation which was a clear manifestation of relatively normal working memory. The patients of Vargha-Khadem, with damage limited to the hippocampus, could learn a great deal of new information with a specific deficit for autobiographic information linked to a particular time and place. The near complete inability for H.M. to form new semantic memories, in contrast to the patients of Vargha-Khadem, is likely due to the additional damage to the entorhinal cortex in patient H.M.

Many of the recent advances in memory research have been possible because of animal models for memory and technological advances in imaging and probing neural tools. These topics as they relate to memory will be discussed in the following sections.

Testing Memory in Animal Models

Animal models offer two critical features that are not often available in studies on human patients. First, with the modern pharmacological and genetic techniques it is possible to precisely control the extent and even the duration of a neural intervention such that the issues of "hidden pathologies" that exist in human patients are far less of a concern. Second, in animals it is possible to dictate the timing between learning, neural disruption, and test. However, the major limitation of animal research is that animals cannot verbally report their recollections, and therefore indirect tests of memories are required (Fig. 3).

To justify the use of an animal model, it is first important to show a conserved, cross-species relationship between anatomy and cognition (Squire 1992). After the seminal testing on patient H.M., researchers immediately set to work to find a memory test that could be given in the nonhuman primate that mimicked the pattern of spared and disrupted faculties exhibited after the temporal lobe resection. Delayed
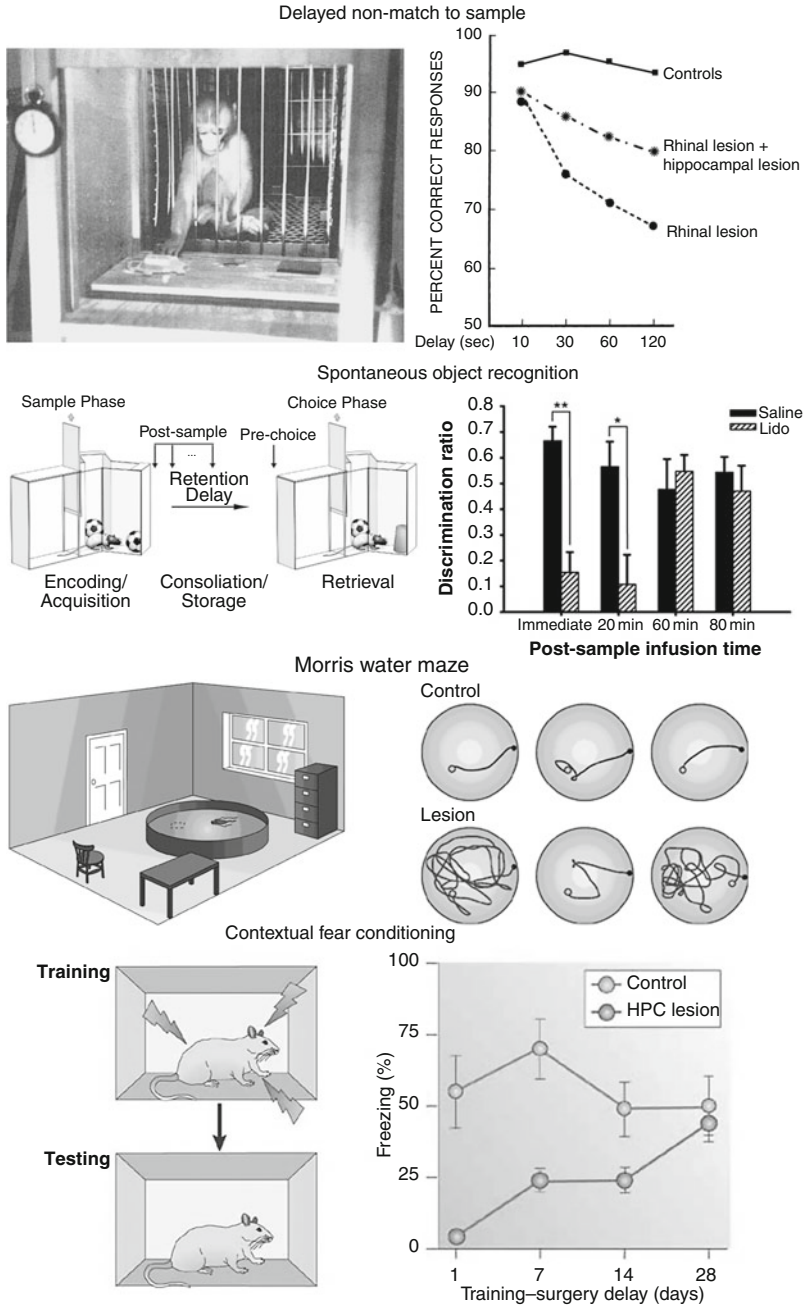
Fig. 3 Common tests for probing memory in animals. Performance on the delayed non-match to sample task depends upon integrity of the rhinal cortices (Based on Meunier et al. 1996). On the spontaneous object recognition task, the discrimination ratio, the ratio of the exploration time for the

match to sample and delayed non-match to sample (DNMS) paradigms resulted from these efforts. In these tests, monkeys are presented with an object and then following a delay are presented with two objects, the one shown just prior and a novel object. In the DNMS monkeys are to choose the novel item, in delayed match to sample the previously presented item is rewarded. A large number of items are needed to observe delay-dependent, medial temporal lobe deficit, highlighting the role that other memory systems can have in guiding behavior. Systematic lesioning studies show the DNMS performance depends most critically on the rhinal cortices, the primary cortical input to the hippocampus, as lesions to the these cortices cause as great a deficit as combined hippocampal and rhinal lesions. Critically, performance without a delay is spared showing that performance deficits were not due to perception or motivation.

Since rats and monkeys, as well as human infants, have a natural tendency to spontaneously attend to novelty, tasks can be designed to compare the time spent exploring novel versus familiar items, with this difference defining the memory strength. In versions of the task that probe the episodic-like recall of what happened where, memory is tested by placing a familiar object in a position that had been previously occupied by a different, familiar object; at test, the space and object are individually familiar, but the combination is novel and detection of this novelty by increased exploration is the sought after memory signal. In monkeys and rodents, hippocampal lesions reliably cause failure to recognize spatial rearrangements of familiar items. In general, hippocampal damage spares recognition of objects but causes deficits in recall of object and place conjunctions, a scenario that models formation of a particular episodic memory of what happened where. On the other hand, lesions to the perirhinal cortex, a part of the medial temporal lobe, cause object recognition deficits.

Another important memory test is contextual fear conditioning. In this paradigm, the subject is placed into a chamber which can be scented or display distinctive cues. After a short duration, the subject receives a low intensity foot shock and is thus conditioned to be afraid of that context. Frightened rodents stop moving or "freeze" and memory can be tested by assessing the level of freezing when the subject is placed back into the conditioned chamber. This memory depends upon integrity of the hippocampus and the amygdala, another subcortical limbic structure with strong bidirectional connections with the hippocampus. Careful manipulations have shown that the hippocampus is responsible for forming a representation of the spatial context (Bouton 1993) that is transmitted to the amygdala to be associated to the

---

Fig. 3 (continued) novel versus familiar item, indicates object recognition. Lidocaine silencing of the perirhinal cortex of the rat impairs performance on this task (Winters et al. 2008). The Morris water maze probes spatial memory for a hidden platform. Hippocampal lesions disrupt memory for the hidden platform as evidenced by the circuitous, indirect trajectories in lesioned animals as compared to the direct routes seen in control subjects (Eichenbaum 2000). In contextual fear conditioning, the subject is shocked in a conditioning chamber and then, after some delay, returned to the testing chamber to assess the level of freezing, the innate fear response. Hippocampal lesions 1 day, and not 28 days, after training decreases freezing on subsequent testing (Frankland and Bontempi 2005)

electrical shock. In a variant of the task in which a tone is given just before the shock, the hippocampus is not required for learning the tone and shock pairing, an association that is disrupted by lesions to the amygdala.

In the contextual fear conditioning task, and others, hippocampal lesions cause a retrograde memory gradient with profound impairments for conditioning that occurred recently and relatively intact memory for learning that took place weeks earlier (Fig. 4). This gradient has been observed in people as well; recall that H.M. had normal childhood memories. This retrograde amnesia gradient has inspired the theory of system's level memory consolidation in which the hippocampus sends some teaching signal to the cortex that allows information to be represented in cortical networks instead of, or in addition to, those in the hippocampus (McGaugh 2000; Debiec et al., 2002).

Rodent research has heavily relied upon tests of spatial memory since the early 1900s. In one such test developed in the 1980s, the Morris water maze, rats are placed in a large tub of cloudy water and learn to swim to a hidden, submerged platform. Memory is then tested by removing the platform and measuring the latency or distance traveled to first reach the platform location and on the amount of time spent swimming in the vicinity of the platform. Rats and mice with hippocampal lesions fail to remember the platform's location and show retrograde memory deficits that do not show a consolidation gradient as described for the contextual fear learning paradigm. These observations, and the physiological ones described in the following section, have led to the hypothesis that the hippocampus stores a spatial map of the environment (Moser et al. 2008), similar to that originally proposed by Tolman in the 1940s.

In another important memory test, the radial arm maze, rodents learn which of several arms contain a reward. In an important experiment that highlighted the dissociation of multiple memory systems, Robert McDonald and Norman White trained rats on three versions of the task that differed by the rule required to find reward (McDonald and White 1993). In the "win-shift" version of the task, all arms contained reward and rats collected food from one arm at a time with a delay between subsequent arm visits, thus requiring animals to keep in mind *where* they are and *where* they had been before to avoid incorrectly entering an arm in which the food had been retrieved. In the "win-stay" version of the task, four arms were lit and baited and rats had to learn to *enter* the lit arms which changed location from trial to trial. In the conditioned cue preference version of the task, rats were first trapped in a lit arm with access to food and a nonlit arm without food and then were subsequently tested for whether they preferred the lit or nonlit arms. Lesioning the hippocampus only caused a deficit on the "win-shift" version of the task consistent with its role place learning. Lesioning the dorsal striatum caused deficits on the "win-stay" version of the task, highlighting its role in learning the type of stimulus-response associations studied by Skinner. Lesioning the amygdala caused deficits on the conditioned cue preference version of the task, which did not involve learning a motor response but rather just the CS-US pairing.

Lesion studies show that a brain region is necessary but provides no information as to how that brain region contributed to the computation required for behavior.
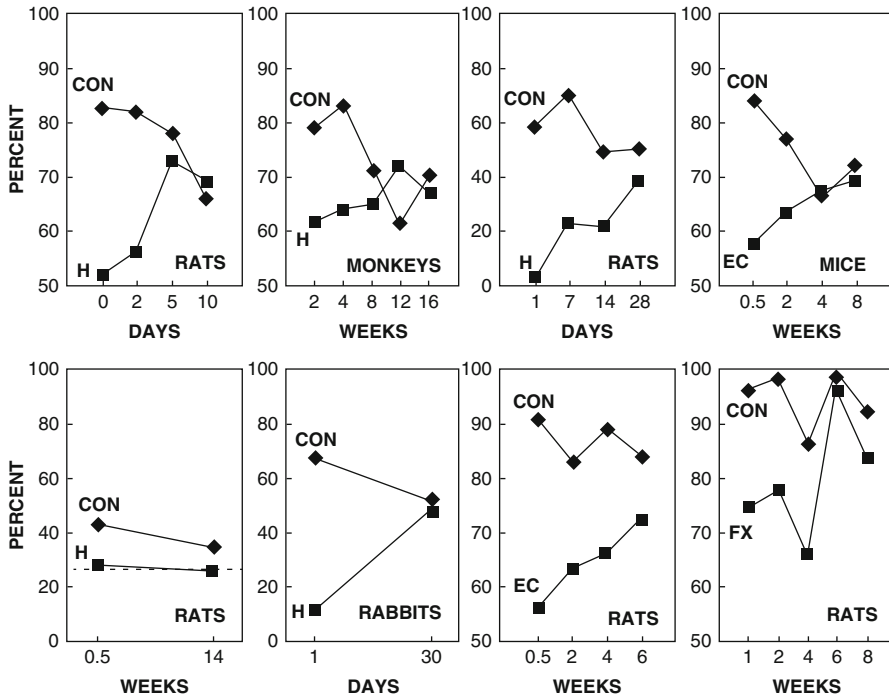
Fig. 4 Summary of findings from studies of retrograde amnesia. Performance of control (*CON*) and operated animals (*H* hippocampus, *EC* entorhinal, *FX* fornix) as a function of the interval between training and surgery. In general, lesions to the hippocampal system long after learning leave memory intact. These observations led to the theory systems consolidation where memory initially depends upon the hippocampus, though through hippocampal-cortical interactions eventually it can be stored in intracortical circuits (Milner et al. 1998)

To study how computation arises from neural tissue, it is necessary to image or record brain activity during such computations. This is the focus of the following section.

Understanding Neural Computation Through Recordings and Targeted Manipulations

Brain activity can be indirectly queried using functional magnetic resonance imaging (fMRI), which measures the amount of oxygenated blood flowing to that region, and the blood oxygen level dependent (BOLD) signal, which correlates roughly with local cell activity. This technology is particularly important in humans as it does not require implanting a recording device, it is noninvasive, and it is possible to assess encoding and retrieval processes while the brain is imaged.

As episodic memories require associating multiple aspects of an experience, experiments have tried to capture this feature by pairing an item (e.g., words,

pictures of objects or faces) with some other information present during the item presentation. During learning, the level of BOLD activity in the hippocampus correlates with subsequent successful retrieval of the paired associate, a phenomenon known as the subsequent memory effect. In addition, greater hippocampal activation is also observed during accurate retrieval of memories made within the lab and during recall of autobiographic episodic memories of everyday life events. These imaging studies support the conclusions reached from neurological examination of patients with brain damage; the hippocampus is involved in both encoding and retrieval.

Interestingly, hippocampal activation also occurs during confident retrieval of the source of a studied word irrespective of accuracy of the memory. Indeed, the hippocampus is active during the imagination of events that have never occurred and together these findings conform to the notion that memory is a constructive process as opposed to a filing cabinet of personal experiences.

The fMRI studies described above bolster the conclusion that hippocampal activity somehow supports memory. However, fMRI has poor spatial resolution ($\sim$1 mm^3) given the density of neurons in the hippocampus ($\sim$10,000 neurons per mm^3) and poor temporal ($\sim$2 s plus the hemodynamic response) resolution given the speed of electrochemical communication ($\sim$2 ms). Therefore, techniques with better spatiotemporal resolution are needed to understand the link between neural computations and cognition. For these experiments, it is generally necessary to turn to an animal model, as increases in spatiotemporal resolution currently require invasive recording technology that is often not available in a patient population and unethical to implement in healthy volunteers. The most popular animal model has been the rodent and the following discussion on the neurophysiology of memory will consequently focus on research on the memory systems of the rat and mouse.

Many biological discoveries have only been possible due to engineering advances, and in the field of electrophysiology this occurred in the 1970s with the miniaturization of voltage amplifiers. This miniaturization is necessary since the potentials recorded in the brain are too small to be carried faithfully across long wires to be amplified elsewhere. With a preamplifier located on the experimental subject, it is possible to record the action potentials from single neurons and also the summed activity from all electrochemical events within the vicinity of the recording lead, the local field potential (LFP), in an awake, behaving animal. Since it is common to record from several neurons on the same electrode, it is often helpful to sample the signal with multiple closely spaced recording sites (e.g., tetrode, silicon probe) that each record slightly different versions of the same signal based off of the distance between the wire and the electrochemical source. In this way, it is possible to assign action potentials to the individual neurons, "single units," that emitted them.

In a seminal report, O'Keefe and Dostrovsky discovered neurons in the hippocampus that fired preferentially when the experimental subject, a rat, was in a particular position and were silent when the rat was outside of the cell's preferred location (Fig. 5) (O'Keefe and Dostrovsky 1971). These cells came to be known as place cells and the location of the environment where they fired, place fields. Recording from a population,
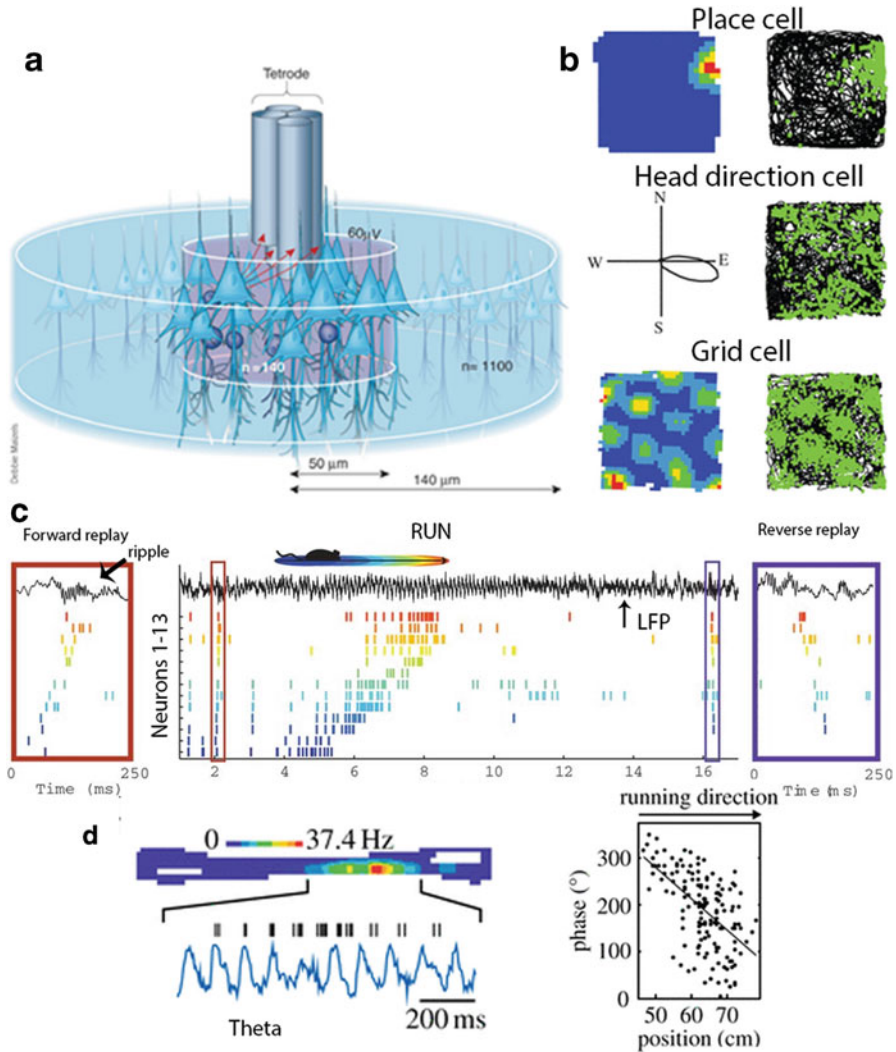
Fig. 5 (**a**) Illustration of a tetrode recording from nearby cells (Buzsaki 2004). (**b**) Three cell types. *Left* The firing rate map of a place cell shows place field of the cell associated with a high-firing rate (*red*). *Right* Top down view of the rat's trajectory (*black*) and the location of the rat when the single place cell fired (*green*). *Left* The polar histogram of the firing rate of a head direction cell as a function of the head direction of the subjects when the cell fired. *Right* The head direction cell fires in many positions but only when the subject faces East. (**c**) Each color coded row is a raster plot from a single neuron. When the rat runs down the track, the neurons fire in the order of their place cells and theta can be observed in the LFP. During ripples, the cells fire ~10–15 times faster in the same order (*forward replay*) or in the reverse order (*reverse replay*) (Diba and Buzsáki 2008). (**d**) Color plot shows the cells place field (*red*). During a single pass through the field, the spikes from the cell can be seen to precess from the late phases of theta to early phases upon place field exit (Hartley et al. 2014)

or ensemble, of such place cells reveals place fields that canvas every portion of the environment such that the rat's position can unambiguously be decoded by assessing the level of activity across the population of neurons. This is an example of a distributed code, since the overall spatial information can only be readout by considering the activity of the group of cells. This is also an example of sparse coding since at any one position only a few cells are active. Other parts of the brain show denser coding in which cells are active in response to a greater diversity of stimuli. Importantly, hippocampal cells have been shown to respond to objects, odors, and rewards suggestive of binding multiple elements of an experience that constitute a memory. Moreover, the location of the place field can change with experience, or remap, which has led to the theory that new hippocampal "maps" are created for each new memory.

Unlike other sensory coding systems, there is no space-sensing organ and localization in the world must be generated by internal processing within the hippocampus or surrounding areas. In the search for how spatial coding arises, several interesting neural correlates have been observed in neighboring and connected brain areas. In the entorhinal cortex, a subset of cells fire in multiple places in an environment and the firing fields are regularly spaced in a tessellating hexagonal grid, for which the "grid cells" get their name (Hafting et al. 2005). To maintain evenly spaced firing fields despite changes in heading and running speed, there must be some way to integrate head direction and speed. In the entorhinal cortex, as well as other connected areas, cells have been found that fire preferentially when the animal is facing in a particular direction, head direction cells. Other cells have been found that fire as a function of running speed and the firing rate of some grid cells is a function of space, head direction, and running speed. Though it was initially believed that grid cells summate to form place cells, it now appears more likely that this is not the case and instead, it is the grid cell code that depends upon the integrity of the hippocampus. Simultaneous recordings from grid cells and place cells and from place cells and head direction cells show that these systems work in concert as remapping in one of these codes is associated with remapping in the other.

The place cell recordings satisfy a necessary covariance requirement of linking hippocampal firing with stimuli that define a memory. Importantly, activation and silencing of hippocampal cells also relates to memory in ways that are expected if their activity is necessary and sufficient to cue recall. In a series of experiments (Liu et al. 2012) in Susumu Tonegawa's lab, mice underwent contextual fear conditioning and the neurons that were active during learning were genetically tagged. The tag was accomplished by introducing DNA that allows expression of the light-activated ion channel, ChR2, an opsin normally found in algae that depolarizes the neuron when exposed to experimentally delivered blue light (Boyden et al. 2005). ChR2 expression was under the temporal control of the experimenter and also transcription factors which are endogenously translated during learning. These transcription factors are known as immediate early genes (IEGs) and are known to bind to the promoter region of DNA to initiate RNA transcription of proteins necessary for synaptic remodeling and memory. This sophisticated genetic system allows cells that were undergoing plasticity at the time of learning to express ChR2 and therefore be under the optogenetic control of the experimenter.

Tagging hippocampal neurons during contextual fear conditioning and then using blue light to reactivate them in a nonconditioned chamber causes the mice to freeze, as though they were recalling the frightening experience. Importantly, light activation of cells tagged in the absence of a shock does not cause freezing; only activating a sparse group of cells that were plastic during experience can drive recall-like behavior. Silencing these cells, using a different hyperpolarizing opsin, causes a memory deficit evidenced by less freezing and also prevents neocortical cells that had been tagged during the original learning experience from becoming reactivated in the conditioned chamber during a recall test. Because these data satisfy the necessary and sufficient benchmarks linking cause and effect, these results are some of the strongest evidence that the hippocampus stores a memory trace, or an "engram," in the identity of which cells originally coded the experience. Moreover, these experiments address the commonly held belief that the hippocampus contributes to retrieval by reinitiating the cortical networks that were active at the time of learning.

Coincident Activity Changes Connection Strength Between Neurons

In the optogenetic experiments described in the previous section, the coincident activity of the previously coactive neurons was artificially recreated by the experimenter; however, in the brain there must be a mechanism to ensure that cells that were coactive before are more likely to become coactive again. This cooperative activity is important since single cells receive inputs from thousands of other neurons and it is rare that any one of the inputs is sufficient to drive the cell to fire. Also, coincidence in the world is usually meaningful in defining parts of a whole and cause and effect. One solution to this problem was originally postulated by the Canadian psychologist Donald Hebb:

> When an axon of cell A is near enough to excite B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased

Recordings conducted in the perforant path inputs from the entorhinal cortex to the dentate gyrus of the hippocampus reveals just such a coincidence dependent change in connection strength. Electrical stimulation of the perforant pathway causes neurons in the dentate to depolarize and fire, resulting in an evoked response in the extracellular local field potential known as a field excitatory postsynaptic potential (fEPSP) and a population spike that reflects the simultaneous spiking of many neurons. The slope of the fEPSP and the timing and size of the population spike reflect the strength of the connection between the stimulated axons and the postsynaptic neurons. Strong, rhythmic stimulation of the perforant pathway causes this connection to increase in a phenomenon known as long-term potentiation (LTP) (Bliss and Lomo 1973). Other stimulation protocols, at slower frequencies, can

cause long-term depression (LTD) which is a weakening of the connection strength. LTP and LTD are dependent upon the IEGs mentioned earlier as disrupting the IEGs disrupts the long-term changes in the connection strength. In general, there is an impressive parallel between the molecules and cellular processes that disrupt LTP and LTD and those that disrupt memory (Fig. 6).

It has been difficult to definitely show that memory is caused by changes in synaptic strengths. Recordings from the auditory cortex to the lateral amygdala show potentiation after classical conditioning of a tone and shock, similar to what can be observed when those connections are artificially strengthened in high-frequency stimulation. Inducing LTP in the hippocampal circuit can cause place cells remap, suggesting that changes in connection strength underlie the remapping of place fields that can be observed during learning. Perhaps the strongest evidence linking LTP and LTD to memory involved pairing a shock with artificially manipulated patterns of brain activation. Control of the cortical neurons was achieved by first transfecting neurons in the auditory cortex and auditory thalamic nucleus with a virus to deliver the DNA for a light-activated ion channel, ChR2. Blue light flashes (CS) delivered to the amygdala, which only affected the terminals of the transfected auditory neurons, were paired with a shock (US) which eventually decreased the rate of lever pressing for reward (CR) in response to the light alone (CS). Subsequently, low frequency light pulses were given to trigger LTD which eliminated the ability for the optical CS to suppress the CR, memory seemed to have blocked or eliminated. Induction of LTP reinstated the ability for the CS to induce the conditioned response. Therefore LTD and LTP could toggle the ability of the CS to elicit the CR and therefore toggle memory expression.

The protocols used to induce LTP and LTD involve stimulation patterns not normally observed in the brain. However, other more realistic protocols have been shown to induce synaptic changes. Paired recordings of pre- and postsynaptic cells in which the precise spike timing of each can be experimentally manipulated by injecting current into the connected neurons has revealed that LTP is observed when the presynaptic cell reliably fires before the postsynaptic, as what would occur if that cells were to causally and synaptically drive the postsynaptic cell to fire (Bi and Poo, 1998). When the cells are stimulated to fire such that the postsynaptic cell fires before the presynaptic, then LTD is observed. The strength of the LTP and LTD is dependent on the temporal offset between the spiking which is a phenomenon known as spike-timing dependent plasticity (STDP). Therefore STDP converts relative spike timing to connection strength.

Brain Rhythms Organize Spike-Timing Relationships Between Neurons

Oscillations are common in dynamical systems that have excitatory and inhibitory feedback. The brain is one such system, and given the multitude of time constants dictating the speed of the feedback, there is also a multitude of frequencies that can be observed even when recording from a single area.
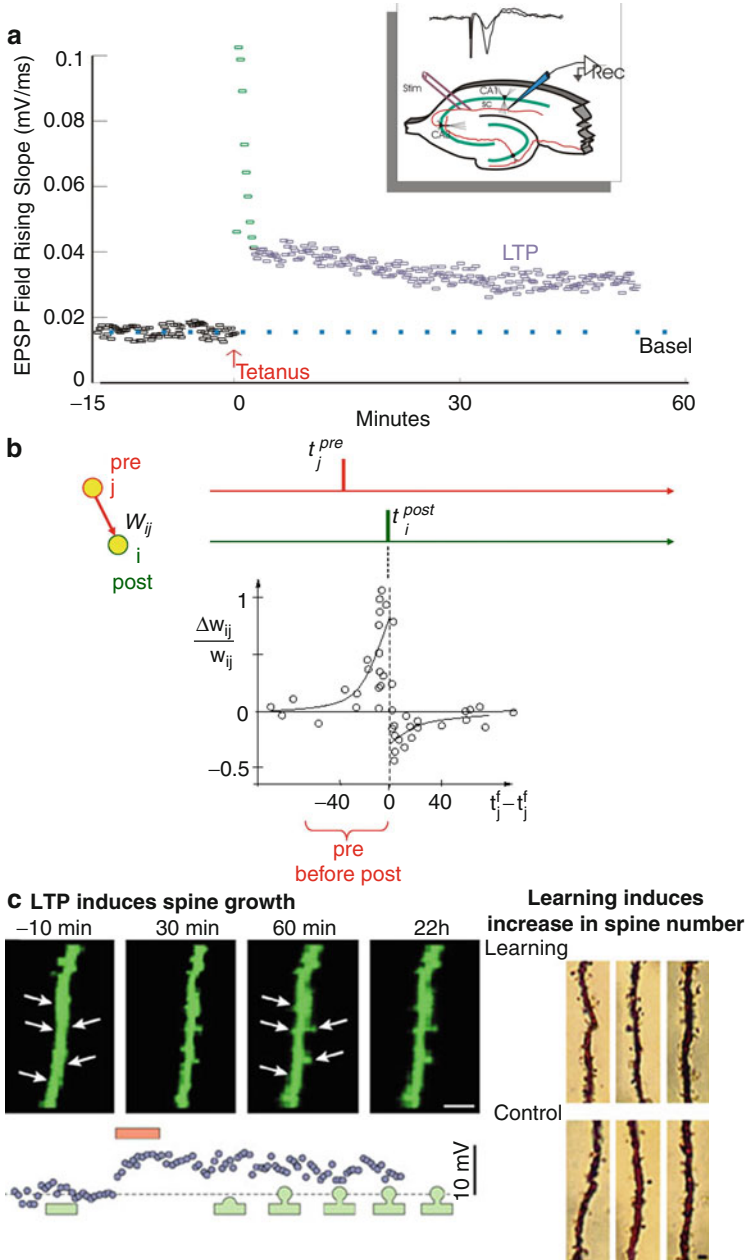
Fig. 6 (**a**) Tetanus, high-frequency stimulation of the Schaffer collateral, connecting area CA3 to area CA1, in the hippocampus increases the slope of the fEPSP in CA1. The increase has been shown to last hours and can depend upon changes in the number of postsynaptic glutamate receptors in CA1. The increase in synaptic strength is known as LTP. (**b**) Spike-timing dependent plasticity depends upon the timing of action potentials in the pre- and postsynaptic cells. When the

In the hippocampus, the prominent rhythm during movement is the 5–12 Hz theta rhythm. This can be observed in the LFP and also in the rhythmic bursting of single cells. There is no single origin of theta. Early models of the origin of theta posited that hippocampal cells oscillated at theta due to an external pacemaker drive from the medial septum since cells in this subcortical region fire at the theta rhythm and septal lesions eliminate hippocampal theta. It is now clear that theta-like activity can be induced in vitro in hippocampal slices, where the medial septum is removed through depolarization of the neurons with cholinergic agonists. Even single cells show resonance at theta frequencies which can be measured by a greater postsynaptic response to stimulation at 8–12 Hz as compared to inputs of different frequencies. The strong rhythmic activity of the theta rhythm provides temporal windows in which presynaptic inputs can be integrated, other windows in which cells fire, and windows of refractoriness in which the network is relatively silent.

The timing of action potentials in the hippocampus is strongly modulated by the theta rhythm. When rat runs through a cell's place field, that cell emits a series of rhythmic bursts where the number of action potentials per burst depends upon the location within the place field, with the fastest firing in the center of the place field. In fact, the rhythmic cell firing is slightly faster than the population average that dictates the frequency of the LFP. The frequency mismatch causes cells to fire at different LFP theta phases depending on where the rat is in the place field (O'Keefe and Recce 1993; Skaggs et al. 1996). Upon place field entry, spikes are emitted late within a cycle and then advance as the rat runs through space such that by the time the rat leaves the place field, the spiking of the place cell is early in the theta period. This systematic advance of spiking phase relative to position is known as theta phase precession (Fig. 5).

By considering a population of cells, each with a place field and each precessing relative to LFP theta, it becomes clear that an important role of theta is to organize cell sequences. Within each theta period, the early spikes reflect cells that code for positions behind the subject, the spikes at the trough reflect where the subject is, and the late spikes are emitted by cells with place fields centered ahead of the rat. Therefore within each theta period, which is about 125 ms, there is a compressed representation of a sequence of positions that will unfold about ten times slower (Dragoi and Buzsáki 2006). The temporal compression is important since it converts spatial distances into short temporal offsets. Based off of the properties of STDP, the temporal offsets can then be converted into synaptic weight differences. These theta sequences have been interpreted as the phase sequences predicted a half century ago by Donald Hebb (Hebb 1949).

Fig. 6 (continued) presynaptic cell fires first, the synapse can be potentiated, though when the postsynaptic cell fires first, the overall connection strength is depressed. (**c**) In many cells, connections are formed on protrusions from the dendrites known as spines. Spines are dynamic and both LTP protocols and memory favor spine growth over spine elimination suggesting this morphological change as a critical component of long-term memory storage (Lamprecht and Ledoux 2004)

Episodic memory is organized in space and time. This theoretical conversion from space, to spike-timing offsets, to synaptic weights provides a physiological mechanism for such an organization. According to computational models, places that are occupied in rapid succession, neighboring places, are encoded by cells that become connected to one another. The asymmetry in the STDP rules ensures the order is captured in the connection strengths. Subsequent reactivation of these sequences allows memory to progress in time as it did during initial learning.

Theta is not the only rhythm in the hippocampus. In fact, theta is predominantly observed when the subject is moving or is otherwise actively engaged. During quiet wakefulness and during slow-wave sleep, ripple oscillations can be seen at 110–200 Hz. Ripples are generated in the CA1 region of the hippocampus due to strong reciprocal connections between excitatory and inhibitory cells. These sharp-wave ripples (SWRs) are the most synchronized event in the mammalian brain. Simultaneous fMRI and electrical recording studies reveal profound brain-wide changes in the BOLD signal during SWRs which suggest that the hippocampus can influence the rest of the brain during the brief (~120 ms) bouts of activity. In addition, blocking SWRs by electrical stimulation of the hippocampus causes spatial memory impairments.

Ripples are not present all of the time but instead come in bursts initiated by strong inputs arriving at CA1 from CA3 via the Schaffer collarerals. This drive can be readily studied because the hippocampus is a laminar, or layered, structure with the cell bodies for the principal, excitatory cells located in a narrow cell layer and axonal input streams from different afferent areas arriving at different parts of the dendrites. This highly organized input anatomy allows conclusions to be reached about the origin of a cellular drive based off of changes in the LFP recorded at different depths. Current source density (CSD) analysis is perfectly suited for such an analysis. CSD requires sampling and comparing the LFP at multiple sites. Such analysis can reveal: current sources which are positive charges in the extracellular sites and current sinks which reflect influx of negative charges. A sink is normally associated with depolarization of nearby cells or cellular processes, since when depolarization occurs when positive ions leave the extracellular space and flow into the cell which causes a hyperpolarization that is recorded as a current sink. Current sources, on the other hand, either reflect the passive return of the ions from inside the cell, or active inhibition, due to negatively charged Cl- ions flowing into the cell, depolarizing the extracellular space. During ripples, CSD analysis reveals current sinks in the stratum radiatum, where the Schaffer collaterals from CA3 terminate onto neurons of CA1. The overall depolarization triggered by CA3 inputs affects the LFP and this influence is called a sharp wave, a depolarization that co-occurs with ripples and neuronal spiking.

The prominence of ripples led researchers to wonder how place cells spike during these events. To address this question, rats are typically trained to run across straight tracks for a food or water reward or to forage in open arenas for randomly placed food rewards. When the animal is permitted to sleep or during bouts of quiet wakefulness, neurons fire in the same order, or in reverse order, as they did during exploration but 10–15 times faster (Fig. 5). These "replay" events can code for

meters of space and, during wakeful planning, tend to end at the locations that a rat will move toward. Ripple events in the hippocampus also occur during reactiviation of cells in the neocortex suggesting cortico-hippocampal coordination. Together, these results support the theory that replays could be an important "teaching signal" necessary for systems consolidation (McClelland et al. 1995).

Outlook

We are now at a point in systems neuroscience where the tools exist to test theories that have existed for decades. Rapid advances in genetics allow targeted delivery of designed DNA to specific cell types at specific times. The advance of optogenetics allows millisecond excitatory and inhibitory control over neural activity, and the rainbow of fluorescent protein families allows imaging of the dynamic subcellular processes, second messenger systems, and even novel features of the gross anatomy. So, what are the next big challenges for the field?

The search for the engram that started 100 years ago is far from complete. The new optogenetics studies reveal that stimulating subsets of cells can induce retrieval-like behavior. Is the engram simply the identity of which neurons were coactive during an experience? This seems unlikely. It is clear that relative firing rate and the temporal ordering of action potentials carries information. How does a rate code and temporal code become embedded within a network during learning? Just as importantly, how are these patterns readout by downstream networks? Coupled recordings of connected brain regions with targeted disruption can begin to answer these tough questions. Also, the activity of the synapse is likely to be more important than the spiking behavior of a neuron in understanding the moment-to-moment state of the network (Buzsaki 2010). Such an understanding can only be achieved by recording the activity across many synapses in vivo, a feat that is not currently possible.

Another fundamental problem that must be addressed pertains to how the brain "chooses" which cells and synapses become modified with experience. This question can also be posed on a gross scale as the competition between memory systems and the allocation of attentional resources. This question is important because it drives at the point made at the opening of the chapter – memory requires the grouping together of stimuli with a common meaning. Modification of a brain region during learning will alter how those networks behave for the memories that previously depended upon those connections; the brain is not a tabula rasa nor filing cabinet. Therefore, how resources are allocated during learning has profound influence on the organization of memory and the stability of old memories during learning. It is now possible, through optical imaging, to record the activity of single cells over months to visualize network modifications that occur during information accumulation. The psychological research program started by Bartlett nearly 100 years ago on how prior knowledge influences learning can finally be subjected to biological investigation.

Human brain research has advanced in large part due to the noninvasive imaging technology afforded by fMRI. Electrophysiological recordings and BOLD imaging

can now be conducted simultaneously. It is important to make the link between electrical activity and the BOLD signal in order to understand what the BOLD signal conveys about electrochemical signal propagation. This advance will allow the community to fully leverage this noninvasive technology beyond brain mapping and hopefully to a computational understanding.

The goal of memory is to drive action. Tolman faced detractors who questioned how memories derived from his cognitive maps could be translated into movement and in fact this remains an open problem today, despite the widespread acceptance that the brain can represent internal models of the world, even those that have never been experienced. Coordination between the medial temporal lobe, (Mishkin et al. 1984) the prefrontal cortex, and striatum is undoubtedly required for action selection, but how the brain systems interact during such decision-making is unknown. Simultaneous, multiregion recordings coupled with behavioral models are necessary to make progress in this field.

Conclusion

In the study of memory, the same phenomena have been rediscovered many times by philosophers, psychologists, behavioralists, and now biologists. For example, the existence of multiple memory systems and learning-induced changes in the communication between brain regions are important themes in guiding memory research. Beyond the intrinsic interest in understanding the origins of thought, solving how information is stored in the brain will have profound ramifications for treating diseases of the mind, such as Alzheimer disease, which are becoming more prevalent as humans live longer. The language the brain uses to communicate with itself is still a mystery, though modern technological advances and theoretical modeling offers promising avenues for answering such hard questions in neuroscience.

References

Bi GQ, Poo MM (1998) Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. J Neurosci 18:10464–10472

Bliss TV, Lomo T (1973) Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol 232:331–356

Bouton ME (1993) Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. Psychol Bull 114:80–99

Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K (2005) Millisecond-timescale, genetically targeted optical control of neural activity. Nat Neurosci 8:1263–1268

Buzsáki G (2010) Neural syntax: cell assemblies, synapsembles, and readers. Neuron 68:362–385

Corkin S (2002) What's new with the amnesic patient H.M.? Nat Rev Neurosci 3:153–160

Debiec J, LeDoux JE, Nader K (2002) Cellular and systems reconsolidation in the hippocampus. Neuron 36:527–538

Diba K, Buzsáki G (2008) Hippocampal network dynamics constrain the time lag between pyramidal cells across modified environments. J Neurosci 28:13448–13456

Dragoi G, Buzsáki G (2006) Temporal encoding of place sequences by hippocampal cell assemblies Nat Rev Neurosci 5:145–157

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. Nat Rev Neurosci 1:41–50

Frankland, P.W., and Bontempi, B. (2005). The organization of recent and remote memories. Nat Rev Neurosci 6:119–130

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

Hartley T, Lever C, Burgess N, O'Keefe J (2014) Space in the brain: how the hippocampal formation supports spatial cognition. Philos Trans R Soc Lond B Biol Sci 369:20120510.

Hebb D (1949) The organization of behavior. Wiley, New York

Hopfield JJ (1995) Pattern recognition computation using action potential timing for stimulus representation. Nature 376:33–36

Lamprecht R, LeDoux J (2004) Structural plasticity and memory. Nat Rev Neurosci 5:45–54

Liu X, Ramirez S, Pang PT, Puryear CB, Govindarajan A, Deisseroth K, Tonegawa S (2012) Optogenetic stimulation of a hippocampal engram activates fear memory recall. Nature 484:381–385

Manns JR, Eichenbaum H (2006) Evolution of declarative memory. Hippocampus 16:795–808

Marr D (1971) Simple memory: A theory for archicortex. Philos Trans R Soc Lond B Biol Sci 262:23–81

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

McDonald RJ, White NM (1993) A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. Behav Neurosci 107:3–22

McGaugh JL (2000) Memory – a century of consolidation. Science 287:248–251

Meunier M, Hadfield W, Bachevalier J, Murray EA (1996) Effects of rhinal cortex lesions combined with hippocampectomy on visual recognition memory in rhesus monkeys. J Neurophysiol 75:1190–1205

Milner B, Squire LR, Kandel ER (1998) Cognitive neuroscience and the study of memory. Neuron 20:445–468

Mishkin M, Malamut B, Bachevalier J (1984) Memories and habits: two neural systems. In: Lynch G, McGaugh J, Weinberger N (eds) Neurobiology of learning and memory. Guilford press, New York, pp 65–77

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

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

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

Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. J Exp Psychol Anim Behav Process 20:11–21

Skaggs WE, McNaughton BL, Wilson MA, Barnes CA (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. Hippocampus 6:149–172

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

Tulving E (1972) Episodic and semantic memory. In: Organization of memory. Academic, New York, pp 381–402

Winters BD, Saksida LM, Bussey TJ (2008) Object recognition memory: neurobiological mechanisms of encoding, consolidation and retrieval. Neurosci Biobehav Rev 32:1055–1070