

Computational models of the hippocampal region: linking incremental learning and episodic memory

Mark A. Gluck¹, Martijn Meeter¹ and Catherine E. Myers²

¹Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, 197 University Avenue, Newark, New Jersey 07102, USA

²Psychology Department, Rutgers University, Newark, NJ, USA

The hippocampal region, a group of brain structures important for learning and memory, has been the focus of a large number of computational models. These tend to fall into two groups: (1) models of the role of the hippocampal region in incremental learning, which focus on the development of new representations that are sensitive to stimulus regularities and environmental context; (2) models that focus on the role of the hippocampal region in the rapid storage and retrieval of episodic memories. Rather than being in conflict, it is becoming apparent that both approaches are partially correct and might reflect the different functions of substructures of the hippocampal region. Future computational models will help to elaborate how these different substructures interact.

And so these men of Indostan
Disputed loud and long,
Each in his own opinion
Exceeding stiff and strong,
Though each was partly in the right,
And all were in the wrong!
–The Blind Men and the Elephant, John Godfrey Saxe
(1816–1887)

A review of theories and computational models of the function of the hippocampal region in learning and memory produces a startlingly diverse and conflicting set of views about the functional role of this brain region. In particular, those who study the function of the hippocampal region in incremental learning paradigms, such as classical conditioning and human associative learning, have argued that this region is crucially involved in the development of novel and flexible representations, and that it is sensitive to stimulus–stimulus regularities and learning context [1–6]. By contrast, those who study episodic memory tasks, such as recall and recognition, have emphasized the role of the hippocampus in the rapid

encoding of episodic memories [7–14]. Although both camps emphasize a key role for the hippocampal region in encoding and learning, it is sometimes hard to accept that they are talking about the same region of the brain.

The diversity of hippocampal-region models seems as irreconcilable as the claims of the blind men in the Hindu fable who argued about the nature of an elephant. One man touched the trunk and exclaimed, ‘The elephant is like a snake!’ Another man touched the side and proclaimed that the elephant was like a wall, while a third touched the leg and concluded that the elephant was like a tree. In the end, of course, each was partly right and all were totally wrong. Might the same be true of theories of hippocampal-region function?

Early approaches to understanding the hippocampal region through animal models were hampered by imprecise lesion techniques that incompletely destroyed key areas while extending into nearby extrahippocampal areas. Similar problems occur in studying humans, because damage is seldom limited to a single brain structure. As such, the data available until recently allowed only the broadest of generalizations about the aggregate function of the whole hippocampal region. However, recent advances in anatomical, surgical and physiological techniques have made it clear that many functions previously ascribed to the hippocampus proper might, in fact, be subserved by other fields in and around the hippocampus, including the entorhinal cortex, the dentate gyrus and the subiculum (Fig. 1). We will refer to these regions and the hippocampus proper together as the ‘hippocampal region’, although the nomenclature varies in the literature.

Computational models play an important role in extending our understanding of the neural bases of learning and memory. The most valuable of these models are born amidst a wealth of experimental studies and justify their existence by inspiring further empirical research. Most models of the function of the hippocampal region have their roots in psychological theories and research on learning and memory; these have led to a wealth of insights into how diverse behaviors in animals

Corresponding author: Mark Gluck (gluck@pavlov.rutgers.edu).

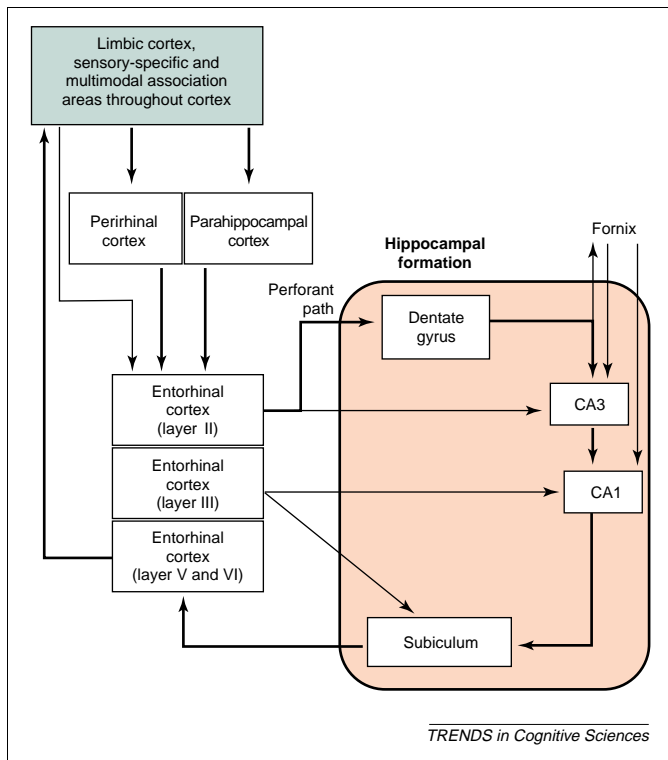


Fig. 1. The hippocampal region consists of several interconnected subregions; major pathways are shown as arrows. Terminology is not always consistent. The name 'hippocampus proper' is usually reserved for the fields CA1 to CA3; the fields shown in the figure are called the 'medial temporal lobe', the 'hippocampus' or, as in this article, the 'hippocampal region'.

and humans depend on this brain region. By simplifying and isolating core principles of brain design, computational models help us understand which aspects of brain anatomy, circuitry and neural function are responsible for particular types of behavior [2].

We review here two main classes of theories and computational models of the hippocampal region: those that seek to explain incremental learning and those that address episodic memory. We argue that a potential synthesis is emerging wherein the incremental learning models describe representational transformations that might be localized to the input regions to the hippocampus, whereas the declarative memory models describe the

storage and recall of these previously processed representations by the CA3 and CA1 regions. According to this view, both the representational and episodic approaches are right, but incomplete.

Models of hippocampal-region function in incremental learning

The influence of hippocampal-region function on incremental learning can be seen through the wide empirical literature demonstrating impaired or altered behavior following damage to some or all of the hippocampal-region structures (Table 1). These data have suggested to several modelers that the hippocampal region performs an information-processing function to modify stimulus representations that are then adopted by other brain regions. This approach, like much computational neuroscience, traces its earliest roots to the models of David Marr [9]. Here we review two such computational models. In both, one network module representing the hippocampal region interacts with other network modules representing other brain regions. Damage to the hippocampal region is simulated by disabling the hippocampal-region module and observing the behavior of the remaining modules. These models can implement many aspects of associative learning, particularly classical conditioning, and they are useful for understanding how the hippocampal region might interact with the rest of the brain to facilitate certain kinds of learning.

The cortico-hippocampal model of Gluck and Myers

Gluck and Myers [1,2] approached hippocampal functioning top-down by beginning with a broad and abstract description of the computations that depend on the hippocampal region in classical conditioning. In their initial model, the hippocampal region was treated as an information-processing system that transformed stimulus representations according to specified rules within a series of linked connectionist networks. In particular, the model argued that the hippocampal region compresses (or makes more similar) the representations of inputs that co-occur or are otherwise redundant, and differentiates (or makes less similar) the representations of inputs that predict

Table 1. Hippocampal-region involvement in incremental learning

Paradigm	Details*	Sample references
Sensory preconditioning	AB + A + B?	Port and Patterson [32]
Latent inhibition	A - A +	Solomon and Moore, Shohamy <i>et al.</i> [33,48]
Learned irrelevance	A - , + A +	Allen <i>et al.</i> [21]
Discrimination reversal	A + , B - A - , B +	Berger and Orr [34]
Trace conditioning	A → +	Moyer <i>et al.</i> [35]
Context-shift	A + in X A? in Y	Penick and Solomon, Honey and Good [36,37]
Negative patterning	A + , B + , AB -	Rudy and Sutherland [38]
Generalization gradient	A + test variants on A	Solomon and Moore [33]
Occasion setting	A → B + , C → B -	Han <i>et al.</i> [39]
Blocking	A + AB + B?	Allen <i>et al.</i> , Baxter <i>et al.</i> [31,40]
Overshadowing	A + , AB +	Garrud <i>et al.</i> , Schmajuk <i>et al.</i> [41,42]
Long ISI conditioning	A + , where A is long	Port <i>et al.</i> [43]
Contextual conditioning	In X, A + , in Y, A -	Good <i>et al.</i> [44]
US magnitude shifts	A + , then weaken or strengthen +	Han <i>et al.</i> [45]
Acquired equivalence	A + , B + A - B?	Coutreau <i>et al.</i> [46]

*Details listed are representative of canonical classical conditioning paradigms. A, B, C = stimuli. X, Y = conditioning contexts; ISI = interstimulus interval; US = reward. A + = A followed by reward; A - = A not followed by reward; A? (or B?) = test responding to A (or B). Comma separates trial types that are interleaved within a phase, vertical bar separates phases. Arrow indicates temporal delay between events.

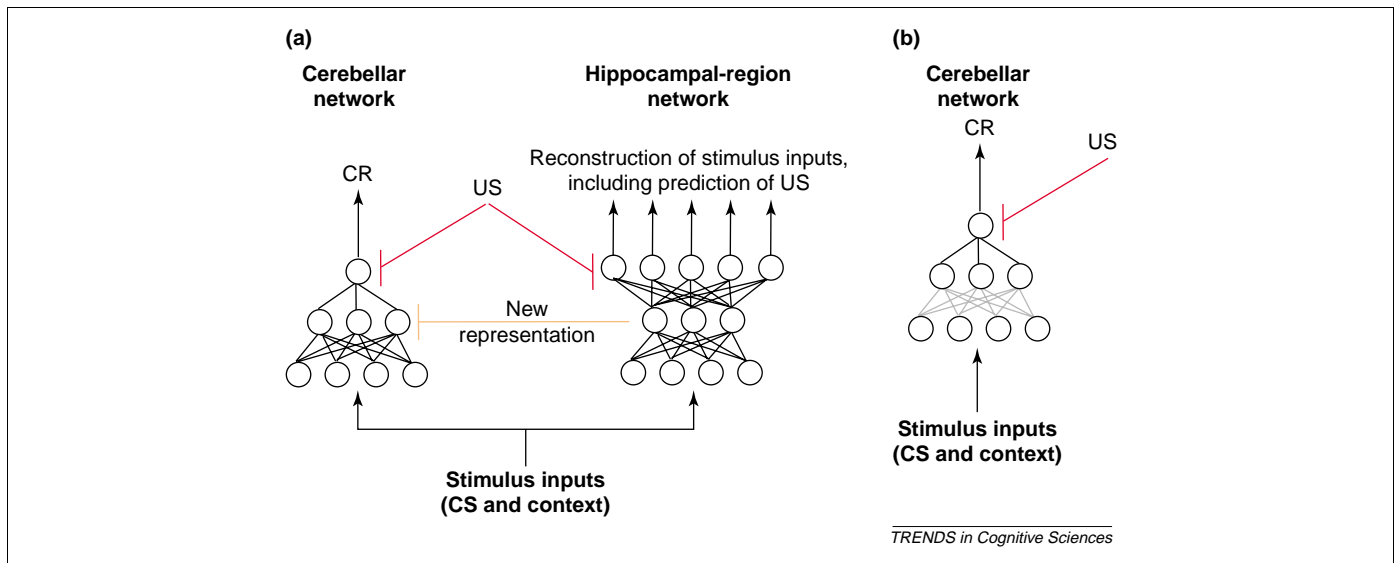


Fig. 2. The cortico-hippocampal model (Gluck and Myers [1]). (a) In the intact model, the hippocampal region provides representational information to long-term memory sites, such as the cerebellum (illustrated here) and cortex; these representations are incorporated into ongoing learning to map from stimuli to responses. (CR, conditioned response; US, unconditioned stimulus) (b) After damage to the hippocampal region, the representational information is eliminated, although simple learning to map stimuli to responses is still possible.

different future events. As a simple analogy, if thunder and lightning always co-occur, they should be treated as analogous and part of the same broader event. On the other hand, if two mushrooms look roughly alike but one is edible and one is poisonous, then their representations should be made more distinct, exaggerating the subtle differences between them. The compressed and differentiated representations formed in the hippocampal region develop over multiple training trials through exposure to a range of stimuli and contextual regularities. These representations are then provided to other modules representing long-term storage in cerebral and cerebellar areas, which incorporate these new stimulus representations into their ongoing stimulus–response learning.

This information-processing theory is incorporated in the connectionist network model shown in Fig. 2a [1,2]. Processing in the hippocampal region is implemented via a predictive autoencoder [15,16], which learns to transform stimulus inputs, through a narrow internal node layer, to outputs that reconstruct those inputs and also predict future reinforcement (or other salient events). Because the internal layer in this network contains fewer nodes than the input and output layers, the network is forced to compress redundant information while at the same time preserving and differentiating information that predicts reinforcement.

This hippocampal-region network then sends the new representations to a long-term memory (LTM) network, which models storage sites in the neocortex and cerebellum. A random recoding of the hippocampal-region network's internal-layer activations becomes the 'desired output' for the internal layer of the LTM network, and the error is the difference between this and the internal layer's actual output. The LTM network then uses an error-correcting rule to adapt its lower layer weights, just as it did to adapt its upper layer weights. Over time, the internal-layer nodes of the LTM network develop

representations that are linear recombinations of those developed by the hippocampal-region network.

Within this model framework, broad hippocampal-region damage is simulated by disabling the hippocampal region network (Fig. 2b). In this lesioned model, no new hippocampal-dependent representations are formed, and the training signal to the LTM network is silenced. The LTM network can adopt no new representations, although it can still learn to map from its existing representations to new behavioral responses.

Empirical applications and implications

Although Gluck and Myers's original cortico-hippocampal model did not directly address physiological mechanisms, it was sufficient to capture a broad range of data regarding the effects of hippocampal-region damage in animals and humans [1,17–19]. The model led to novel predictions regarding lesion effects, many of which have since been confirmed in animals [20,21] and in humans with hippocampal-region damage [22–25]. In more recent computational modeling, Gluck and Myers have shown how some of this proposed information-processing function could indeed emerge from known anatomical and physiological characteristics of the brain substrate, including the entorhinal cortex [26], the dentate gyrus and the septo-hippocampal cholinergic system [27,28].

Support for this view of processing in the hippocampal region comes from functional brain imaging. In particular, the cortico-hippocampal model suggests that the medial temporal lobes should be very active early in training, when subjects are learning about stimulus–stimulus regularities and evolving a new stimulus representation, but less active later in training when other brain regions (e.g. the basal ganglia) are using these representations to perform the task. In a recent study [29], subjects were given a probabilistic category learning task in which they were exposed to hundreds of trials in which they were asked to predict the weather ('sun' or 'rain') based

on the presence or absence of four different tarot cards with various geometric features. As expected by the cortico-hippocampal model [1,17], medial temporal lobe activity was highest early in training and then decreased; by contrast, basal ganglia activity was low at first and increased during training.

Schmajuk and DiCarlo (S–D) model

An alternative view of the role of the hippocampal region in conditioning has been proposed by Schmajuk and colleagues in an evolving series of computational models [3,4,30]. The spirit and aim of these models are similar to the cortico-hippocampal model, in that they are concerned with information-processing roles for different brain regions, and with how these regions interact. Moreover, Schmajuk's models address much of the same body of empirical data as the cortico-hippocampal model. However, the particular function that Schmajuk and his collaborators assign to the hippocampus differs from that supposed by the cortico-hippocampal model; specifically, Schmajuk and colleagues assume that the hippocampal region is crucial for forming new stimulus configurations. A classic example is the negative patterning problem (known to computer scientists as 'exclusive-or' or XOR): cues A and B each mean one thing when presented alone but mean something else when presented together as the configuration AB. This is a particularly difficult task for animals and machines to learn because the correct response to AB is not merely a sum of the responses to A and B alone.

Schmajuk and DiCarlo have presented a computational model, often referred to as the S–D model [3,4], that assumes that the cortex is able to combine cue information to allow configural learning. Then, they argue, the cerebellum learns to map from this configural information to a behavioral response. They argue that the hippocampal region helps to modulate this learning both by helping to calculate an 'error' measure – the difference between what the system predicted and what actually happened – and broadcasting this error measure to the cortex and cerebellum. In addition, the hippocampal region broadcasts the system's predictions to the cerebellum.

Damage to the hippocampal system is simulated in the S–D model by disabling both of these putative hippocampal functions. As a result, the cortex can no longer form configural representations and, additionally, the cerebellum cannot make use of the hippocampal prediction to control stimulus competition. Learning paradigms involving stimulus configuration or stimulus competition will be disrupted, although simpler learning, such as associating a single cue with a single response, might still be possible. Applied to classical conditioning, the S–D model can account for a sizeable range of empirical findings [3,4].

Incremental learning: brain substrates

Schmajuk's S–D model and Gluck and Myers's cortico-hippocampal model both address the same domain (classical conditioning), so there is considerable overlap in their predictions regarding what kinds of behavior

are and are not dependent on hippocampal-region processing. However, the two models make a few divergent predictions. For example, the S–D model predicts that blocking should be abolished after hippocampal-region damage. By contrast, Gluck and Myers's cortico-hippocampal model expects that the blocking effect should be spared in lesioned animals, albeit with subtle changes in responding on the first compound-cue trial (a prediction recently confirmed [31]). However, both models have had considerable success in accounting for a large body of existing data on incremental learning.

More recently, the authors of both models have begun to consider how their putative hippocampal-region function could be mapped onto specific brain structures. Gluck and Myers argued that at least one component of their proposed hippocampal-region function, namely compression of the representations of co-occurring stimuli, could emerge naturally from the anatomy and physiology of the entorhinal cortex [26]. When this entorhinal cortex model (Fig. 3a) was connected to the LTM model [47], the resulting model performed like an animal that has a selective lesion of the hippocampus proper but spares the entorhinal cortex, as shown in Fig. 3b. This means that

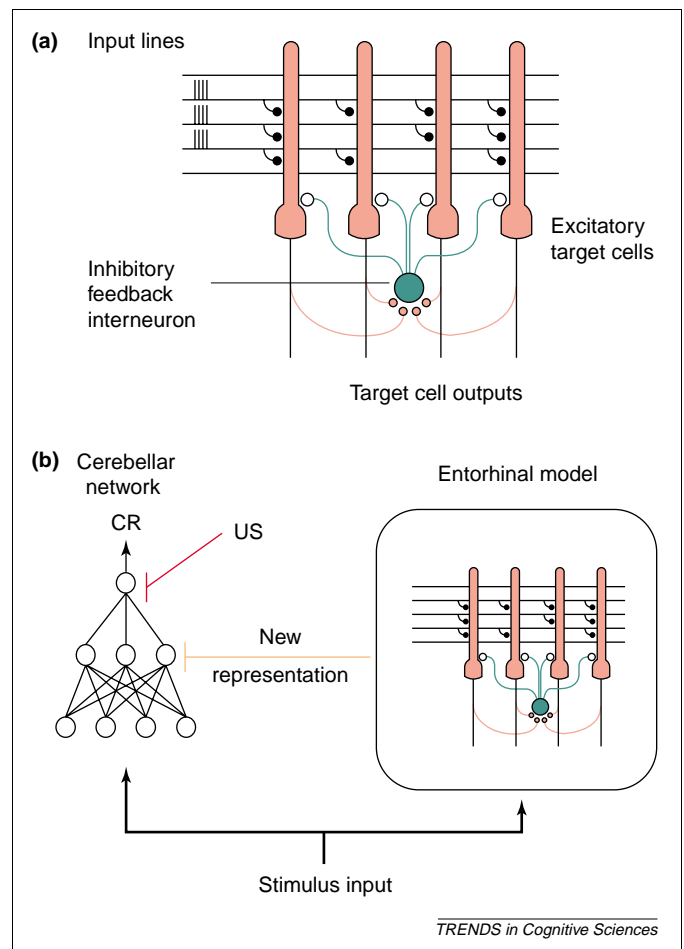


Fig. 3. (a) The anatomy and physiology of the entorhinal cortex is sufficient to allow it to perform representational compression, part of the function ascribed by Gluck and Myers to the hippocampal region as a whole. (b) This implies that selective hippocampal lesion, that spares the entorhinal cortex, should spare the ability to perform entorhinal-dependent representational changes [26].

those conditioning behaviors that depend specifically on representational compression might be abolished by entorhinal lesion but might survive a more selective hippocampal lesion that spares the entorhinal cortex. For example, the entorhinal-only model can perform latent inhibition and learned irrelevance, two behaviors that are disrupted by broad damage to the hippocampal region. These predictions were recently confirmed using ibotenic acid lesions to destroy cells in either the entorhinal cortex or the hippocampus in rabbits trained on either latent inhibition and learned irrelevance in the rabbit eyeblink conditioning preparation. As expected, only the entorhinal lesions interfered with these putative redundancy-compression behaviors [48,49]. More recently, Gluck and Myers have suggested that the other half of their putative hippocampal-region function might similarly depend on other hippocampal-region structures (see [26]). Importantly, this conceptualization leaves open the question of what function the hippocampus proper (fields CA1 and CA3) might perform.

Schmajuk and colleagues have likewise elaborated their initial model to include specific hippocampal-region substructures [50,51]. They have suggested that at least one component of their proposed hippocampal-region function – the prediction signal used to control stimulus competition – might be localized in the entorhinal cortex, although they still assume that the remaining component – the error signal that modulates configural association – takes place in the hippocampus [30].

The emerging consensus from both models could thus be that structures outside the hippocampus proper (particularly entorhinal cortex) might be the true home of some of the functionality previously ascribed to the hippocampal region as a whole, where stimulus representations are incrementally changed during associative learning so as to reflect salient stimulus–stimulus regularities that emerge over many training trials.

Models of hippocampal-region function in episodic memory

Whereas the incremental memory models reviewed above focus on how new representations are formed and shaped over multiple training trials, a second tradition of hippocampal models has focused on what happens during a

single episode or event. These models have assumed that the hippocampal region simply stores whatever pattern is presented to it by the neocortex. This view is exemplified by Marr's theory of simple memory [9]. Marr contrasted such 'simple' hippocampal memory with the neocortex, which he thought was capable of forming new representations.

This view of the hippocampal region as a photo-camera of neocortical input pervades most episodic memory models. In these models, the input to the hippocampal region consists of arbitrary vectors to be stored. The vectors are then later retrieved in their entirety based on a partial cue [7,9,11–13,52]. In essence, the hippocampal region forms a compact code that is bidirectionally linked to a neocortical representation. If part of the neocortical representation is later presented to the hippocampal system as input, it will reactivate this compact code. Subsequently, the hippocampal code can reactivate the missing parts of the neocortical representation [9]. The function of the hippocampal region is thus to complete stored patterns for reinstatement in neocortical regions, as illustrated in Fig. 4.

Representational assumptions in episodic memory models

In episodic memory models, hippocampal patterns are either assumed to be random (e.g. [12]) or are formed automatically via unsupervised learning [7,9,11–14,52]. In the latter case, such patterns are formed during one single presentation of the input [7,11,13]. This is an essential difference with incremental memory models, in which representations usually evolve over the course of many trials [1].

The view that episodic memory is about storing random vectors seen only once is also central to many mathematical models of memory [53,54]. It is not an unnatural view: episodic memories consist of random elements, such as the unrelated words on a word list or the unpredictable collection of people attending a particular party, and this information might be provided only once. The task facing episodic memory is not to discover regularities (e.g. which guests always visit the same parties) but to recall accurately who was present at a particular party. For all practical

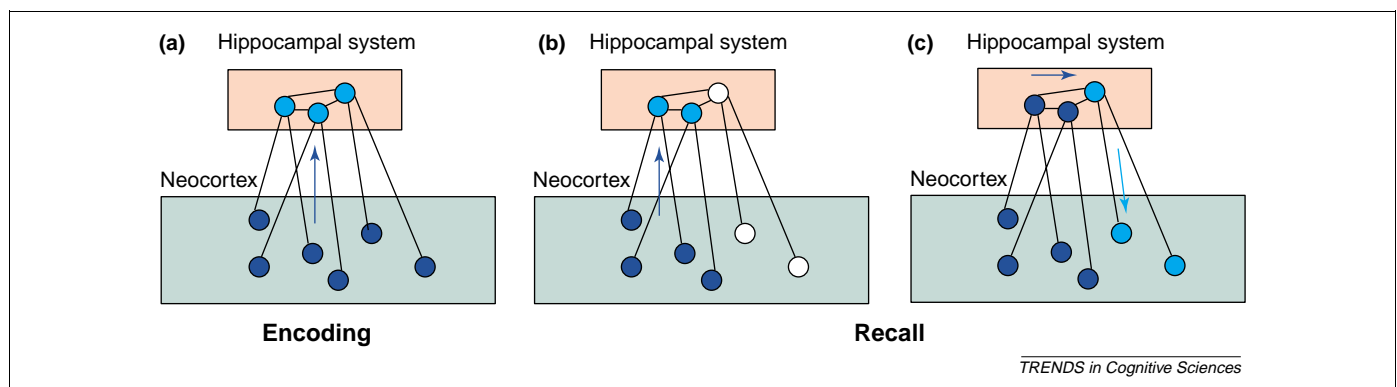


Fig. 4. (a) Formation of an episodic memory: a neocortical pattern gives rise to a hippocampal-region pattern, with connections being laid between the neocortical and hippocampal pattern. (b) When part of the pattern is active in the neocortex during recall, it will reactivate part or whole of the hippocampal pattern. (c) This pattern will in turn complete the neocortical pattern.

Table 2. Hippocampal-region involvement in episodic memory

Paradigm	Exposure trials	Sample references
Autobiographic memory	1	Vargha-Khadem <i>et al.</i> [55]
Verbal memory, recall	3–5	Reed and Squire, Rempel-Clower <i>et al.</i> [56,57]
Picture recognition	1	Reed and Squire, Rempel-Clower <i>et al.</i> [56,57]
Delayed figure reproduction	1	Reed and Squire, Rempel-Clower <i>et al.</i> [56,57]
Contextual fear	1	Kim and Fanselow [58]
Morris water maze	1–4	Cassel <i>et al.</i> [59]
Radial arm maze	1–4	Cassel <i>et al.</i> [59]

purposes, episodic memory systems thus have to deal with random, unstructured input that they do not have to ‘understand’ or process, merely file away for later retrieval. The idea that the hippocampus is such an episodic memory system stems from the effects of hippocampal-region damage on the ability to form episodic memories after a single exposure, although more typically after a few exposures; key data from the wide literature are summarized in Table 2.

In many episodic memory models, the hippocampal system is assumed to form relatively sparse patterns that overlap less than the input patterns from which they are formed [13,52]. These patterns are sometimes referred to as ‘conjunctive’ [52,60] because they code for a conjunction of features (i.e. an episode) and not for individual features separately. The lesser overlap is referred to as ‘orthogonalization’ or ‘pattern separation’: patterns that overlap in neocortex become non-overlapping (orthogonal) in the hippocampal system. Orthogonalization is essential for an episodic memory system because it reduces interference between similar memories. Sparseness of patterns helps in this orthogonalization: two sparse patterns in which, for example, two neurons out of 100 are active, are less likely to overlap with each other than two patterns in which 50 neurons are active [9,60,61].

Extensions of the basic theme

Models of the hippocampal system and episodic memory still have many elements in common with Marr’s basic view of hippocampal functioning, but recent efforts have added considerable elaboration and refinement.

Marr’s work on calculating the capacity of a ‘simple memory’ [9] was extended using sophisticated information theory measures [61]. As another example, Marr suggested for computational reasons that the hippocampus was only a temporary store, with older memories being transferred from the hippocampus to the neocortex [9]. In the 1980s, Squire and colleagues developed a similar hypothesis to explain a neuropsychological mystery, namely that recent memories seem more vulnerable to hippocampal damage than more remote memories [62,63]. In the 1990s, Buzsáki hypothesized that hippocampal function could be linked to two distinct processing modes: during ‘theta state’, when hippocampal neurons fire in rhythmic synchrony, memories are stored in the hippocampus, whereas during the less structured ‘sharp wave state’, memories are transferred from the hippocampus to the neocortex [64,65]. Both of these ideas have been incorporated in several computational models [8,10,11,66].

A new research question is how new patterns could be stored in the hippocampus without retrieval of old patterns disturbing the formation of new patterns. The solution Hasselmo and coworkers came to was that neuromodulation might set conditions suitable for either retrieval or learning, so that both happen at separate times and do not disturb each other. In their earlier work, this balance between storage and retrieval was modulated by acetylcholine [7] but septal GABAergic projections have recently emerged as an alternative or complementary means of modulation [67].

Another emerging theme in recent models of episodic memory is the core role given to sequence learning in the hippocampal region [68–70]. In sequence learning, the goal is not to complete old patterns from an incomplete cue but to retrieve the next member of a sequence of patterns. Metaphorically, the hippocampus does not shoot photographs but stores sequences of images like a video camera. Again, this is a quite natural conceptualization of episodic memory: an episode is usually seen as not a single image but as a story unfurling in time. Moreover, this idea connects episodic memory to another field in which the hippocampus clearly plays a role, namely that of spatial navigation [71,72].

One episodic memory model has also been extended to tasks requiring slow, incremental learning, such as non-linear discriminations in rats [14]. However, this was achieved by adding error-correction mechanisms that do not play a role in the rest of the model.

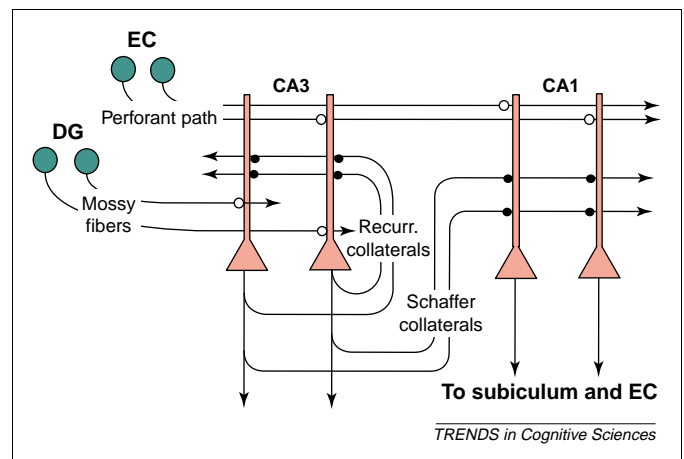


Fig. 5. Circuitry of the CA3 and CA1 fields of the hippocampus proper, typically taken to underlie episodic memory. Inputs to the hippocampus proper arrive over the perforant path of entorhinal cortex (EC) and the mossy fibers of the dentate gyrus (DG). Patterns are linked in CA3 via recurrent collaterals, and then forwarded to CA1 via the Schaffer collaterals. CA1 reintroduces the patterns to the neocortex via the subiculum and the deep layers of the entorhinal cortex.

Episodic memory: brain substrates

Some episodic memory models have treated the hippocampal system as one undifferentiated module [10,11]; others have delved more into anatomical complexities [7,9,13,14,52].

CA3 is often the central module in more detailed models of the hippocampus (Fig. 5). The anatomy of CA3 is marked by neurons with a high degree of recurrent collaterals, which would be ideal either for the autoassociative storage (binding different parts of one pattern) necessary for pattern completion [7,9,61] or for the heteroassociative storage (binding different patterns in one sequence) necessary for sequence learning [68–71]. CA1 is often assumed to play a role in ‘decoding’ of the hippocampal pattern, allowing it to be dissociated with the cortical pattern that gave rise to it [14,52]. Alternatively, CA1 has also been identified with a pattern separator [8]. The dentate gyrus is not always modeled with precision but in recent models it has received an important role as a ‘sparsifier’ – making patterns sparse and therefore enabling pattern separation [14,60].

In episodic memory models, the entorhinal cortex is generally seen as the input region to the hippocampal formation and is usually ignored or grouped with other extrahippocampal cortex. For example, O’Reilly and Norman [13] considered a parahippocampal cortical region (including entorhinal cortex) that computed a familiarity signal from a slowly formed, integrated memory, which seems very compatible with the view emerging from the incremental learning strain of hippocampal models. In both views, the entorhinal cortex slowly forms representations that do not incorporate all details in a single episode (as the episodic memory implemented in CA3 would do) but instead extracts regularities over longer time intervals. These slowly forming representations could form the basis for both a familiarity signal [13] and performance in incremental learning tasks (as suggested by [26]).

Conclusion

Two very different strains of theories and models of the hippocampal region have evolved to explain two largely nonoverlapping sets of behavioral phenomena: stimulus–outcome associations that emerge slowly from exposure to multiple training trials and episodic memories that are instantiated quickly from a single exposure. However, recent advances in anatomical, surgical and physiological techniques have suggested that many functions previously ascribed to the hippocampus proper (especially fields CA3 and CA1) are likely to be subserved by other areas in and around the hippocampus, including the entorhinal cortex, the dentate gyrus and the subiculum. This has led to a potential rapprochement between the two camps of models and theories. A synthesis is emerging wherein the incremental learning models describe representational transformations that might be localized to the input regions of the hippocampus (especially the

entorhinal cortex but possibly also the dentate gyrus) whereas episodic memory models describe the storage and recall of these previously processed representations by the CA3 and CA1 regions. According to this view, both the representational and episodic approaches are both partially correct but each is also incomplete. Future computational models will help to elaborate how these different substructures interact within the broader context of hippocampal-region function. Neither a snake, nor a wall, nor a tree, the hippocampal elephant appears to be all these things, depending on where we look.

References

- 1 Gluck, M.A. and Myers, C. (1993) Hippocampal mediation of stimulus representation: a computational theory. *Hippocampus* 3, 491–516
- 2 Gluck, M.A. and Myers, C.E. (2001) *Gateway to Memory: An Introduction to Neural Network Modeling of the Hippocampus in Learning and Memory*, MIT Press
- 3 Schmajuk, N. and DiCarlo, J. (1990) *Backpropagation, Classical Conditioning and Hippocampal Function*, Northwestern University
- 4 Schmajuk, N. and DiCarlo, J. (1992) Stimulus configuration, classical conditioning and hippocampal function. *Psychol. Rev.* 99, 268–305
- 5 Eichenbaum, H. and Bunsey, M. (1995) On the binding of associations in memory: clues from studies on the role of the hippocampal region in paired associate learning. *Curr. Dir. Psychol. Sci.* 4, 19–23
- 6 Eichenbaum, H. *et al.* (1989) Further studies of hippocampal representation during odor discrimination learning. *Behav. Neurosci.* 103, 1207–1216
- 7 Hasselmo, M.E. and Wyble, B.P. (1997) Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. *Behav. Brain Res.* 89, 1–34
- 8 Lörincz, A. and Buzsáki, G. (2000) Two-phase computational model training long-term memories in the entorhinal–hippocampal region. In *The Parahippocampal Region: Implications for Neurological and Psychiatric Diseases* (Scharfman, H.E. and Witter, M.P., eds) pp. 83–111, New York Academy of Sciences
- 9 Marr, D. (1971) Simple memory: a theory for archicortex. *Philos. Trans. R. Soc. B. Biol. Sci.* 262, 23–81
- 10 McClelland, J.L. *et al.* (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
- 11 Murre, J.M.J. (1996) TraceLink: a model of amnesia and consolidation of memory. *Hippocampus* 6, 675–684
- 12 Nadel, L. *et al.* (2000) Multiple trace theory of human memory: computational, neuroimaging and neuropsychological results. *Hippocampus* 10, 352–368
- 13 Norman, K.A. O’Reilly, R.C. (2003) Modeling hippocampal and neocortical contributions to recognition memory: a complementary learning systems approach. *Psychol. Rev.* (in press)
- 14 O’Reilly, R.C. and Rudy, J.W. (2001) Conjunctive representations in learning and memory: principles of cortical and hippocampal function. *Psychol. Rev.* 108, 311–345
- 15 Baldi, P. and Hornik, K. (1989) Neural networks and principal component analysis: learning from examples without local minima. *Neural Netw.* 2, 53–58
- 16 Hinton, G. (1989) Connectionist learning procedures. *Artif. Intell.* 40, 185–234
- 17 Gluck, M. *et al.* (1996) Late-training amnesic deficits in probabilistic category learning: a neurocomputational analysis. *Learn. Mem.* 3, 326–340
- 18 Myers, C.E. and Gluck, M.A. (1994) Context, conditioning and hippocampal re-representation. *Behav. Neurosci.* 108, 835–847
- 19 Myers, C. and Gluck, M. (1996) Cortico-hippocampal representations in simultaneous odor discrimination learning: a computational

- interpretation of Eichenbaum, Mathews & Cohen (1989). *Behav. Neurosci.* 110, 685–706
- 20 Shohamy, D. *et al.* (1999) Dissociating entorhinal and hippocampal function in latent inhibition of the classically conditioned eyeblink response. *Abstr. Soc. Neurosci. Ann. Meet. (Miami, FL)* 25, 40
 - 21 Allen, M. *et al.* (2003) Entorhinal lesions, but not hippocampal lesions, disrupt learned irrelevance in rabbit eyeblink conditioning. *Cogn. Affect. Behav. Neurosci.* 2, 214–226
 - 22 Myers, C.E. *et al.* (2000) Localized hippocampal damage impairs acquisition, but not reversal, of a spatial discrimination in humans. *Psychobiology* 28, 275–282
 - 23 Myers, C. *et al.* (2000) Latent learning in medial temporal amnesia: evidence for disrupted representational but preserved attentional processes. *Neuropsychology* 14, 3–15
 - 24 Myers, C. *et al.* (2002) Hippocampal atrophy disrupts transfer generalization in non-demented elderly. *J. Geriatr. Psychiatry Neurol.* 15, 82–90
 - 25 Myers, C. *et al.* (2003) Dissociating hippocampal vs. basal ganglia contributions to learning and transfer. *J. Cogn. Neurosci.* 15, 1–9
 - 26 Myers, C.E. *et al.* (1995) Dissociation of hippocampal and entorhinal function in associative learning: a computational approach. *Psychobiology* 28, 116–138
 - 27 Myers, C. *et al.* (1996) A computational model of the effects of septohippocampal disruption on classical eyeblink conditioning. *Neurobiol. Learn. Mem.* 66, 51–66
 - 28 Myers, C. *et al.* (1998) Further implications of a computational model of septohippocampal cholinergic modulation in eyeblink conditioning. *Psychobiology* 26, 1–20
 - 29 Poldrack, R.A. *et al.* (2001) Interactive memory systems in the human brain. *Nature* 414, 546–550
 - 30 Schmajuk, N. and Buhusi, C. (1997) Stimulus configuration, occasion setting and the hippocampus. *Behav. Neurosci.* 111, 235–258
 - 31 Allen, M.T. *et al.* (2002) Selective hippocampal lesions disrupt a novel cue effect but fail to eliminate blocking in rabbit eyeblink conditioning. *Cogn. Affect. Behav. Neurosci.* 2, 318–328
 - 32 Port, R. and Patterson, M. (1984) Fimbrial lesions and sensory preconditioning. *Behav. Neurosci.* 98, 584–589
 - 33 Solomon, P. and Moore, J. (1975) Latent inhibition and stimulus generalization of the classically conditioned nictitating membrane response in rabbits (*Oryctolagus cuniculus*) following dorsal hippocampal ablation. *J. Comp. Physiol. Psychol.* 89, 1192–1203
 - 34 Berger, T. and Orr, W. (1983) Hippocampectomy selectively disrupts discrimination reversal learning of the rabbit nictitating membrane response. *Behav. Brain Res.* 8, 49–68
 - 35 Moyer, J. *et al.* (1990) Hippocampectomy disrupts trace eye-blink conditioning in rabbits. *Behav. Neurosci.* 104, 243–252
 - 36 Penick, S. and Solomon, R. (1991) Hippocampus, context and conditioning. *Behav. Neurosci.* 105, 611–617
 - 37 Honey, R. and Good, M. (1993) Selective hippocampal lesions abolish the contextual specificity of latent inhibition and conditioning. *Behav. Neurosci.* 107, 23–33
 - 38 Rudy, J.W. and Sutherland, R. (1995) Configural association theory and the hippocampal formation: an appraisal and reconfiguration. *Hippocampus* 5, 375–398
 - 39 Han, J-S. *et al.* (1998) Hippocampal lesions enhance configural learning by reducing proactive interference. *Hippocampus* 8, 138–146
 - 40 Baxter, M. *et al.* (1999) Blocking can occur without losses in attention in rats with selective removal of hippocampal cholinergic input. *Behav. Neurosci.* 113, 881–890
 - 41 Garrud, P. *et al.* (1984) Successful overshadowing and blocking in hippocampectomized rats. *Behav. Brain Res.* 12, 39–53
 - 42 Schmajuk, N. *et al.* (1983) Absence of overshadowing in rats with hippocampal lesion. *Physiol. Psychol.* 11, 59–62
 - 43 Port, R. *et al.* (1985) Differential effects of hippocampectomy on classically conditioned rabbit nictitating membrane response related to interstimulus interval. *Behav. Neurosci.* 99, 200–208
 - 44 Good, M. *et al.* (1998) Contingent versus incidental context processing during conditioning: dissociation after excitotoxic hippocampal plus dentate gyrus lesions. *Hippocampus* 8, 147–159
 - 45 Han, J-S. *et al.* (1995) Hippocampal lesions disrupt decrements but not increments in conditioned stimulus processing. *J. Neurosci.* 15, 7323–7329
 - 46 Coutreau, E. *et al.* (2002) Acquired equivalence and distinctiveness of cues: II. Neural manipulations and their implications. *J. Exp. Psychol. Anim. Behav. Proc.* 28, 378–387
 - 47 Gluck, M. *et al.* (2001) Cerebellar substrates for error correction in motor conditioning. *Neurobiol. Learn. Mem.* 76, 314–341
 - 48 Shohamy, D. *et al.* (2000) Dissociating entorhinal and hippocampal involvement in latent inhibition. *Behav. Neurosci.* 114, 867–874
 - 49 Allen, M. *et al.* (2002) Selective entorhinal and nonselective cortical–hippocampal region lesions, but not selective hippocampal lesions, disrupt learned irrelevance in rabbit eyeblink conditioning. *Cogn. Affect. Behav. Neurosci.* 2, 214–226
 - 50 Schmajuk, N. and Blair, H. (1993) Stimulus configuration, spatial learning and hippocampal function. *Behav. Brain Res.* 59, 1–15
 - 51 Schmajuk, N. (1994) *Stimulus configuration, classical conditioning, and spatial learning: role of the hippocampus* World Congress on Neural Networks, INNS Press
 - 52 McClelland, J.L. and Goddard, G.V. (1996) Considerations arising from a complementary learning systems perspective on hippocampus and neocortex. *Hippocampus* 6, 654–665
 - 53 Humphreys, M.S. *et al.* (1989) Different ways to cue a coherent system: a theory for episodic, semantic and procedural tasks. *Psychol. Rev.* 96, 208–233
 - 54 Shiffrin, R.M. and Steyvers, M. (1997) A model for recognition memory: REM–retrieving effectively from memory. *Psychonomic Bull. Rev.* 4, 145–166
 - 55 Vargha-Khadem, F. *et al.* (1997) Differential effects of early hippocampal pathology on episodic and semantic memory. *Science* 277, 376–390
 - 56 Reed, J.M. and Squire, L.R. (1998) Retrograde amnesia for facts and events: findings from four new cases. *J. Neurosci.* 18, 3943–3954
 - 57 Rempel-Clower, N.A. *et al.* (1996) Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *J. Neurosci.* 16, 5233–5255
 - 58 Kim, J.J. and Fanselow, M.S. (1992) Modality-specific retrograde amnesia for fear. *Science* 256, 675–677
 - 59 Cassel, J.C. *et al.* (1998) Fimbria–fornix vs selective hippocampal lesions in rats: effects on locomotor activity and spatial learning and memory. *Neurobiol. Learn. Mem.* 69, 22–45
 - 60 O'Reilly, R.C. and McClelland, J.L. (1994) Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682
 - 61 Treves, A. and Rolls, E.T. (1994) Computational analysis of the role of the hippocampus in memory. *Hippocampus* 4, 374–391
 - 62 Ribot, T. (1881) *Les Maladies de la Memoire*, Germer Baillare, Paris
 - 63 Squire, L.R. *et al.* (1984) The medial temporal region and memory consolidation: a new hypothesis. *Memory Consolidation* (Weingarter, H., Parker, E., *et al.* eds), pp. 185–210, Lawrence Erlbaum Associates
 - 64 Buzsaki, G. (1989) Two-stage model of memory-trace formation: a role for 'noisy' brain states. *Neuroscience* 31, 551–570
 - 65 Buzsaki, G. (1996) The hippocampo-neocortical dialogue. *Cereb. Cortex* 6, 81–92
 - 66 Alvarez, R. and Squire, L.R. (1994) Memory consolidation and the medial temporal lobe: a simple network model. *Proc. Natl. Acad. Sci. U. S. A.* 91, 7041–7045
 - 67 Hasselmo, M.E. *et al.* (2002) A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Comput.* 14, 793–817
 - 68 Jensen, O. and Lisman, J.E. (1996) Hippocampal CA3 region predicts memory sequences: accounting for the phase precession of place cells. *Learn. Mem.* 3, 279–287
 - 69 Levy, W.B. (1996) A sequence predicting CA3 is a flexible associator that learns and uses context to solve hippocampal-like tasks. *Hippocampus* 6, 579–590
 - 70 Sohal, V.S. and Hasselmo, M.E. (1998) GABA-b modulation improves sequence disambiguation in computational models of hippocampal region CA3. *Hippocampus* 8, 171–193
 - 71 Hasselmo, M.E. *et al.* (2002) Neuromodulation, theta rhythm and rat spatial navigation. *Neural Netw.* 15, 689–707
 - 72 O'Keefe, J. and Nadel, L. (1978) *The Hippocampus as a Cognitive Map*, Clarendon Press