Available online at www.sciencedirect.com
ScienceDirect
Behavioural Brain Research 189 (2008) 1–8
BEHAVIOURAL BRAIN RESEARCH
www.elsevier.com/locate/bbr

Review

Memory formation by refinement of neural representations: The inhibition hypothesis

Nikolai Axmacher[a,b,*], Christian E. Elger[a,b], Juergen Fell[a]

[a] *Department of Epileptology, University of Bonn, Sigmund-Freud-Street 25, 53105 Bonn, Germany*
[b] *Life and Brain Center of Academic Research, Sigmund-Freud-Street 25, 53127 Bonn, Germany*

Received 4 December 2007; accepted 20 December 2007
Available online 28 December 2007

Abstract

There is no reasonable doubt that the hippocampus plays an important role in memory processing. A virtually uncountable number of studies in animals and humans have revealed changes in neural activity in this structure during memory formation [Squire LR. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. Psychol Rev 1992;99:195–231; Squire LR, Stark CE, Clark RE. The medial temporal lobe. Annu Rev Neurosci 2004;27:279–306], and hippocampal damage reliably leads to impairments in a large number of memory tests. However, while several correlates of successful memory formation have been found in the hippocampus, it is still an open question *why* specific neural processes support encoding of a particular item. An answer to this question would help to resolve current debates about which memory functions are actually supported by the hippocampus, and why activity in the neural networks of the hippocampus is involved in, or even necessary for, some memory processes but not for others. In this review, we first summarize findings on the electrophysiological activity within the hippocampus during different memory processes. We try to differentiate whether the hippocampus is merely involved in these processes, or whether the hippocampus appears to be necessary for them. Based on a distinction between a more general "encoding state" and the more specific process of "content-specific memory formation", we review data on neural representations within hippocampus and neocortex. We suggest that during memory formation, the hippocampus renders neural representations more sparse by providing an inhibitory signal to the neocortex.

Keywords: Memory formation; Neural representations; Sparse coding; Hippocampus; Intracranial EEG

Contents

1. Memory systems supported by the hippocampus

In this section, we present data indicating that the hippocampus is involved in or necessary for (a) episodic long-term memory (LTM), (b) semantic LTM, (c) non-declarative

* Corresponding author at: Department of Epileptology, University of Bonn, Sigmund-Freud-Street 25, 53105 Bonn, Germany. Tel.: +49 228 287 19347; fax: +49 228 287 19351.
E-mail address: nikolai.axmacher@ukb.uni-bonn.de (N. Axmacher).

0166-4328/$ – see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbr.2007.12.018

Table 1
Memory effects in the hippocampus

	Long-term memory			Working memory
	Declarative		Non-declarative	
	Episodic	Semantic		
Hippocampus involved	Yes	Maybe	Yes [17]	Yes [59,50,5]
Hippocampus necessary	Yes [53]	Probably not [64]	Maybe [79]	Maybe [1,30,50,52]
Electrophysiological signature: ERPs	Subsequent memory related positive late potential [26]	–	–	Load-dependent negative DC shift [5]
Electrophyisological signature: oscillations	Power decrease in lower gamma band [23], increase in high gamma band [67]	–	–	Power increase in gamma band activity [34,5]
Electrophysiological signature: single units	Decrease in firing rate [11]	–	–	–

(implicit) forms of LTM, and (d) working memory (WM). We also describe a necessary role of the hippocampus in episodic memory and some forms of implicit and WM involving multiple items or associations. An overview is given in Table 1.

Episodic (or autobiographic) memory refers to the ability to perform a "mental time travel" [73] and consciously access one's own experiences, together with the role one has played in these experiences. Together with semantic memory, which allows for the conscious retrieval of facts and general knowledge without associated autobiographical information, episodic memory is also termed declarative, i.e. consciously accessible, memory [69]. The best-known example of a clear episodic memory impairment following hippocampal lesions in human subjects is patient H.M., who underwent bilateral removal of the hippocampus due to pharmacoresistant epilepsy [53]. Even patients who have unilateral hippocampal damage, e.g., due to ammons horn sclerosis, often have severe material-dependent memory impairments: Patients with damage to the language-dominant hemisphere often have deficits in verbal memory, whereas spatial memory is typically impaired in patients with lesions contralateral to the dominant hemisphere [32]. The role of the hippocampus for semantic memory is still debated; while H.M. had extensive deficits, e.g., in the detection of ambiguity in sentences [45], these deficits were not significant in other patients with more selective hippocampal deficits [64], suggesting that semantic memory predominantly requires lateral temporal regions. This explanation is in line with findings that semantic memory depends on consolidation, which is related to information transfer into the neocortex and decreasing hippocampal engagement (e.g., [77,4]).

Initial studies with H.M. and other patients indicated that these patients were not impaired in a variety of tasks involving WM, defined as the ability to maintain and manipulate information over a short period of time [6]. Recently, however, this seemingly clear distinction between LTM processes in the hippocampus and extra-hippocampal WM processes has been questioned [58]. Doubts have arisen from several sources. First, a variety of functional magnetic resonance imaging (fMRI) studies in healthy subjects showed that the hippocampus was activated in WM experiments. While some of these studies found increased activation even during maintenance of a single item [59,65,50], others reported activation only during maintenance of inter-item associations [54] or simultaneous maintenance of multiple items [5]. The different results with respect to hippocampal activation during maintenance of a single item may in part be explained by the use of different baseline conditions in these experiments, and further studies are needed to clarify the role of the hippocampus in WM tasks.

However, these studies provide converging evidence that the hippocampus is activated during maintenance of multiple items in WM. Similarly, patients with hippocampal damage did show impairments in WM tasks if these involved maintenance of multiple items [1] or maintenance of associations between items [30,52]. The apparent discrepancy to earlier clinical findings that patients with hippocampal damage are not impaired in WM tasks is likely due to the use of different paradigms: whereas early studies used simple delayed matching to sample tasks involving maintenance of a single item, these recent studies utilized more complex paradigms requiring maintenance of multiple items or inter-item associations. Similarly, early clinical investigations found no impairments in patients with hippocampal lesions in tasks involving non-declarative (implicit) forms of LTM formation such as procedural learning and priming [14], but more recent neuroimaging studies did find activation of the hippocampus in implicit learning tasks involving associations [17]. Again, patients with hippocampal damage showed impairments in these tasks [79], suggesting that the criterion for hippocampal involvement is mostly the use of associative material, and not the particular type of memory.

2. Functional methods for studying memory processes

In addition to the observation in neuroimaging studies that the hippocampus is activated during some WM tasks, patient data suggest that the hippocampus plays a causal role in WM tasks. However, these studies cannot explain *why* this structure is so important for these tasks. Similarly, the well-known studies relating hippocampal damage to LTM impairments cannot explain why the hippocampus is crucial for LTM. How can the question of the hippocampal involvement in specific memory

processes be resolved? It appears crucial to understand why the particular neural architecture of this structure makes it well suited to support specific memory processes (and to be even necessary for them).

Neuroimaging findings reporting hippocampal activation during these tasks are thus insufficient: detecting activation in a brain structure does not explain why this brain structure supports a specific cognitive process. Thus, the specific spatiotemporal activity patterns in the hippocampus during different tasks need to be investigated. Unfortunately, while fMRI has the advantage of being readily usable in healthy subjects and is thus the method of choice if the localization of activity in the entire brain is investigated, its relatively low temporal resolution severely restricts the distinction of specific temporal patterns. In some cases, sustained activations, which are considered to underlie the continuous maintenance of information during WM tasks, can be distinguished from more transient patterns (e.g., [13,15]), but neural activity on a subsecond scale cannot currently be investigated. Moreover, the neural processes correlated with the BOLD response are far from evident and difficult to investigate. The first results on this issue were obtained by simultaneous electrophysiological and fMRI recordings in animals. These studies indicated that a BOLD increase in the visual system correlated stronger with local field potentials, likely reflecting synaptic inputs, than with single-unit activity [44,37]. In humans, simultaneous intracranial EEG (iEEG) and fMRI recordings are currently not possible due to safety considerations. However, several recent investigations used the same paradigm subsequently in epilepsy patients with iEEG electrodes and healthy control subjects [9,49,5]. This approach is particularly interesting because the two methods can be considered complimentary: while fMRI allows the identification of significantly active regions in the entire brain of healthy subjects, iEEG recordings provide direct insights into the neural processes in these regions.

These studies indicate that BOLD responses are better correlated with activity in high rather than low frequency ranges, although low-frequency activity is the most important contributor to event-related potentials (ERPs). For the hippocampus, data are still rather scarce. An increase in gamma power and a negative shift of the direct current (DC) potential was found to accompany an increased BOLD response during WM maintenance of an increasing number of items [5], suggesting that negative potentials in the hippocampus and increased gamma activity correlate with an increased BOLD response. However, studies of LTM formation in a word list-learning paradigm suggest that the polarity of hippocampal components which correlate with an increased BOLD response may be variable: while subsequent memory in this task correlated with an increased late *positive* hippocampal potential in iEEG recordings [26], it induced an increased BOLD response in an fMRI study using the same paradigm [71]. This relationship is particularly interesting because positive potentials may correspond to inhibition of activity [68], suggesting that LTM formation is linked to an inhibitory signal in the hippocampus.

3. Intracranial EEG data of memory processes: activation of inhibition?

Given the variable relationship between BOLD response and neural activity, fMRI recordings cannot explain why processes within the hippocampus are required for some memory processes but not for others. Electrophysiological experiments, on the other hand, allow for a much higher temporal resolution and can be performed in animals and (during presurgical invasive EEG recordings) in human epilepsy patients.

As mentioned above, subsequent memory to individual words correlated with a more pronounced late positive potential in the hippocampus of epilepsy patients [26]. However, subsequent memory effects on event-related potentials might be different for other types of stimuli: for example, in a recent study using rare and abundant items from two different categories (faces and houses) as stimuli, we found a pronounced subsequent memory effect on the P300 component (which has a negative polarity in the hippocampus), but no late positive component (unpublished data). The P300 component (though with a negative polarity in the hippocampus) has been linked to inhibitory processes [66], again suggesting a link between memory formation and inhibition. Time–frequency analyses of these data allow to investigate the effect of subsequent memory on oscillatory activity in the hippocampus, which is particularly interesting because it links these human data to more mechanistic studies in reduced systems such as intracellular recordings from animals or from brain slices. Several effects of subsequent memory on activity in the hippocampus have been described (see Table 1). First, subsequent memory correlated with a decreased activity in the lower gamma frequency range (32–48 Hz; [23]), but with an increased activity in the high gamma band (44–64 Hz; [67]). Second, apart from effects on power, we observed an increased inter-trial phase coherence during subsequently remembered as opposed to forgotten items (unpublished data). These findings may be related to data from animals, where stimulus presentation and direct electrical stimulation induces a reset of ongoing activity in the theta range [47]. As a result, stimulus representations are locked to a specific phase of theta activity, which has been shown to induce long-term potentiation or depression, the probable cellular correlates of memory formation ([35,33]; for a recent review, see [3]). Third, phase synchronization of the hippocampus with the adjacent rhinal cortex was found to be enhanced in the gamma band (32–48 Hz; [23]) and in the low-frequency range (1–19 Hz; [24]) during presentation of subsequently remembered as opposed to subsequently forgotten words.

In addition to these findings from iEEG recordings, the neural processes underlying memory formation have also been investigated using microelectrodes, which allows for the identification of individual action potentials from single cells. Single-unit recordings thus provide the most specific information about cellular neural processes in the human hippocampus. In a series of studies, Itzhak Fried and colleagues investigated changes in firing rate of individual neurons within different subregions of the medial temporal lobe (MTL) during various memory tasks. In a first study, subjects were exposed to a large number of faces with varying emotional expressions and to different objects and later

underwent an old/new recognition task with the same and distractor stimuli [28]. Whereas activity in many brain regions was significantly altered (either increased or decreased) when an item had been presented previously, it did not significantly depend on the subject's response. A similar experiment by Rutishauer et al. [63] confirmed the finding that firing rate during recall predicted previous exposure to a stimulus—even more accurately than the subject's response.

To investigate subsequent memory effects on neural firing rate, Cameron et al. [11] recorded single-unit activity from the medial temporal regions during a paired associate word-learning task. They found that the discharge frequency of single units from the hippocampus during encoding was correlated with subsequent memory of the pairs; importantly, in most neurons the firing rate was significantly lower during presentation of subsequently remembered than forgotten pairs. Also during recognition, hippocampal neurons showed an inhibition of firing rate below baseline upon repeated presentation of the same stimulus, while parahippocampal neurons did not respond to the stimulus anymore at all [74], consistent with the effect of repetition suppression in this region (e.g., [72]). These findings suggest an intriguing hypothesis: inhibition might be a mechanism used by hippocampal neurons to support sparse stimulus encoding in adjacent regions. In other words, successful memory formation might be linked to activation of inhibitory units in the hippocampus, which subsequently refine the neural representation of stimuli in the neocortex [74]. The decrease in firing rate observed with microelectrodes fits well with the observed increase in higher gamma power, because oscillations in the gamma band are likely due to activity in inhibitory interneuron networks, which restrict firing in hippocampal pyramidal cells to specific time windows [76]. Hippocampal activity during states with abundant gamma (and theta) oscillations is relatively low, probably due to an inhibition of hippocampal pyramidal cells by cholinergic inputs [10,31]. However, it should be noted that although synchronization in the gamma band is linked to inhibitory interneurons, it requires activation as well.

A further line of evidence supporting the hypothesis that the hippocampus transmits an inhibitory signal to the neocortex during LTM encoding comes from studies on event-related potentials (ERPs). As described above, subsequent memory effects on hippocampal ERPs have both been observed on a late positive component during encoding of words [26] and on the hippocampal P300 component during encoding of novel faces and houses (unpublished data). In the hippocampus, there are no clear data so far showing whether a positive component corresponds to an increase or a reduction of cellular activity. Such a relationship has been established in the neocortex, however: here, slow positive potentials (with a similar latency as the late component in the hippocampus) are related to inhibition (e.g., [8]), as for instance demonstrated by a disfacilitation of the startle reflex [66]. In general, subsequent memory effects in ERPs recorded with scalp EEG (e.g., increased P300 and late positivity) are usually positive [25], suggesting that neocortical ERP correlates of memory formation correspond to a decreased neural activity in the neocortex.

These data show that similar potentials in the hippocampus and neocortex are enhanced due to successful LTM encoding, and suggest that at least in the neocortex, this correlates with a reduction of cellular activity. But is there any evidence that the hippocampus actually causes these neocortical subsequent memory effects? First, patients with hippocampal lesions show impaired LTM encoding, indicating that the integrity of the hippocampus is required for these tasks (see above). Second, a direct relationship between the neocortical P300 component (which is larger for subsequently remembered items) and the hippocampus has been shown by Knight [38] who investigated patients with hippocampal lesions and found a clear reduction of the P300 response to novel items recorded with scalp EEG. This suggests that the proper functioning of the hippocampus is actually required for a memory-related attenuation of activity in the neocortex.

Taken together, there is some evidence that the subsequent memory effects in the hippocampus are linked to increased inhibition in this region, which is then transmitted to the neocortex. This hypothesis is explained further below.

4. Content-specific memory formation versus the encoding state

Why do these neural processes support memory formation? Some conceptual clarification is necessary here. In his article "What is a neural correlate of consciousness", Chalmers [12] argues that two different meanings of "correlate of consciousness" can be distinguished. On the one hand, general neural processes supporting consciousness (e.g., as compared to coma) can be investigated. On the other hand – and according to Chalmers, this is probably the more interesting case – the neural representation of a specific content while it is consciously processed can be compared to the neural representation of the same content when it is *not* processed consciously. This distinction can also be drawn in the context of memory formation: a central question concerns the neural processes during encoding of a particular item, not the neural processes favoring memory formation or an "encoding state" in general. Among the latter are also the neural bases of psychological processes such as attention, motivation (or consciousness) which support memory formation in one way or another, but are not specifically involved in successful encoding of a particular item—they rather facilitate encoding of a large group of similar items. The neural correlate of a specific content, on the other hand, is the neural representation of this content, and the correlate of *content-specific memory formation* refers to the modification of this representation by memory formation. As this approach requires first to find out the neural representation of this particular content, the next paragraph will give a brief overview of the principles underlying neural representations within neocortex and hippocampus.

5. Neural representations within neocortex and hippocampus

In principle, two different ways of stimulus representation, or "coding schemes", have been proposed: "Sparse coding" by

a small group of neurons, and "distributed coding" by neural assemblies. According to a conceptually relatively simple idea, the central nervous system responds to a stimulus in an increasingly specific ("sparse") way when a stimulus is being processed and analyzed in consecutive perceptual stages [39]. According to this idea, in higher processing stages only a very small number of highly selective neurons, or even a single neuron, responds to each particular stimulus. Indeed, stimulus representations in the visual system appear to be consistent with this idea. For example, a complex visual stimulus (e.g., a face) elicits increased activity in a high percentage of cells in the primary visual cortex. This activity does not depend on the identity of the face or on complex properties such as its gender or emotional expression, but on low-level properties such as contrasts and the angle of the visual field covered by the stimulus. Moreover, adjacent neurons in this region respond to adjacent positions in the visual field, i.e. this region shows a topographical organization with respect to spatial position (accordingly, in the primary auditory cortex, adjacent neurons respond to adjacent frequencies). When the stimulus is consecutively analyzed in higher-order visual areas, its representation depends on increasingly more complex stimulus properties both concerning the identity of the stimulus, which is being processed in the "what-" stream of the temporal cortex, and concerning spatial features, which are being processed in the parietal "where-" stream. Neurons in the inferior temporal cortex as the final processing stage of the "what-" stream respond specifically to stimulus categories such as faces, houses, places, or objects (e.g., [2]).

During these later processing stages, stimulus representations become increasingly specific, i.e. they evoke responses in a decreasing number of cells; this is linked to the fact that the topographical organization of higher order regions follows increasingly complex stimulus dimensions (e.g., [18]). When stimulus processing reaches the hippocampus, individual neurons were shown to respond to the identity of a stimulus, even if this stimulus is being presented in perceptually completely different ways [55]. For example, the same neuron increases its firing rate when pictures of the actress Jennifer Aniston are presented and upon presentation of the written or spoken words "Jennifer Aniston", but not upon presentation of the pictures of other actresses. Calculations based on the number of neurons in the hippocampus, the number of electrode contacts, etc. suggest that only a relatively small proportion (<1%) of all cells in this region respond to each specific stimulus [75], consistent with the idea of "sparse coding" in the hippocampus. However, there are still a considerable absolute number of neurons in this region which respond to a particular stimulus (around one million; [75]). Thus, while hippocampal representations are sparse as compared to neocortical stimulus representations, they are still distributed because in absolute numbers, a large quantity of cells responds to each stimulus.

Are these neural representations in the hippocampus organized in a topographical manner, i.e. do spatially adjacent neurons respond to "similar" inputs? This question is important because only in this case, field potentials (such as ERPs) from the hippocampus could be specific for a given stimulus, and subsequent memory effects on field potentials could be specific modifications of a given stimulus representation ("content-specific memory formation"): field potentials arise from the averaged activity of spatially adjacent neurons and are thus a measure for the similarity of synaptic inputs to these neurons. In humans, there is some evidence for category, but not stimulus specificity of hippocampal field potentials [40]; a recent study by Quiroga et al. [56] shows that cells recorded by adjacent microelectrodes do not respond to similar stimuli and stimulus decoding did not improve when correlations between multiple neurons were taken into account. Similar results were obtained in rodents, where activity of closely adjacent neurons did not show increased correlations as compared to more remote neurons [61], in contrast to earlier findings [29]. Thus, although it is still possible that a topographical organization in the hippocampus according to more abstract stimulus dimensions exists, there is currently no experimental evidence for this. Theoretical considerations even suggest that a non-topographical organization of stimulus representations in an intermediate storage is advantageous for efficient information encoding [46,42].

Taken together, while modality-specific regions of the ventral visual stream show a clear topographical organization, there is no evidence for such an organization in the hippocampus. Hippocampal subsequent memory effects obtained by field potential recordings are therefore most likely not related to learning-related modifications of a specific item, but rather support memory by a more general mechanism. In other words, these effects are not the neural correlates of content-specific memory formation, but rather related to an encoding state. However, they may serve to provide an inhibitory signal to neocortical representations which renders them more sparse. The next paragraph will turn to the relationship of encoding-related electrophysiological activity in the hippocampus and neural representation patterns in the neocortex, i.e. the neural correlates of content-specific memory formation.

6. Learning modifies stimulus representations

As described in Section 4, content-specific memory formation, i.e. memory encoding of a particular event, corresponds to the modification of the neural representation of this event due to memory formation. For example, when you meet a person the first time and encode her face into episodic memory, the neural firing patterns elicited by this face are modified and the neural representation of the face is being embedded into a certain context. There is indeed considerable evidence on the effect of learning on stimulus representations. However, it should be noted that these studies were performed in monkeys, which does not allow for a fine distinction between different types of LTM; it is possible that specific kinds of LTM induce different modifications in stimulus representation (see next section). In monkeys, neural responses in the inferior temporal cortex were more selective to learned than to unlearned items [7]. Interestingly, while the maximal response to the "optimal" stimulus remained stable during learning, the response to other stimuli decreased significantly, indicating that learning involves suppression of irrelevant (and not amplification of relevant) activity. A more recent study by Freedman et al. [27] found that even

the response to the optimal stimulus decreased during learning. Activity of stimulus-specific neurons in many brain regions including the prefrontal cortex [57] and the inferior temporal cortex [43] decreased with repeated stimulus exposure, again consistent with a sharpening of the representations of individual items. Even in the hippocampus, stimulus-specific neurons showed modifications of their firing rates in tasks involving the formation of new associations [78], which became more selective after extensive exposure [80].

In a recent review on representations in the ventral visual stream, Reddy and Kanwisher [62] suggested that familiarity with an item increases both spatial clustering of item representations and the specificity of neural responses. An increase of specificity related to memory formation is consistent with the findings of Viskontas et al. [74], who showed that neural responses in the parahippocampal cortex decrease with learning. Indeed, results from various sources suggest that subsequent memory effects in the hippocampus mainly serve as an inhibitory signal which blocks irrelevant stimulus representations: the increased hippocampal power in the high gamma frequency range [67] is likely related to increased activation of inhibitory interneurons [76]; and the ERP studies reported above show that potentials depending on the hippocampus such as the P300 component [38] are linked to neocortical inhibition.

These findings suggest that the specificity of stimulus representations increases due to an inhibitory signal from the hippocampus. In addition, there is an evidence that learning increases associations within an assembly, probably by Hebbian synaptic plasticity [41]. Within the MTL, successful memory formation was linked to gamma-band synchronization between rhinal cortex and hippocampus [23]. While we elsewhere have extensively described the possible mechanisms linking synchronization and spike-timing dependent plasticity in the hippocampus [23], similar processes might support the strengthening of associations during learning in the neocortex: it has been shown that learning of stimulus associations increased correlations between the firing rates of cells representing each individual stimulus, suggesting increased spatial clustering of familiar items [21,22,48]. Thus, it might be speculated that learning is not only supported by rhinal–hippocampal synchronization, but by synchronization in the neocortex as well, although this idea needs to be tested directly. Taken together, these results suggest that associative learning (or the learning of an item content together with its context) involves a particular learning signal, which does not only sharpen stimulus representations but modifies the network organization as well.

7. Different modifications during different forms of LTM?

In the previous section, we have argued that the hippocampus promotes learning through an inhibitory signal that interacts with the neural representations of a specific stimulus and sharpens these representations. However, as these results were mostly derived from monkeys, a fine distinction between specific subprocesses of LTM formation was not possible. There is good evidence today that declarative memory formation and retrieval rely on at least two processes, familiarity and recollection (for a recent review, see [20]; an alternative view is expressed in [70]). While familiarity describes the feeling that an item has already been encountered before (and is thus the opposite of the subjective impression of novelty), recollection refers to the simultaneous remembering of the context, i.e. of associated items or of the autobiographical situations when the item was first experienced (e.g., [19]). The existence of two separate processes underlying familiarity and recollection has been shown in several double dissociation experiments in patients with selective damage of subregions of the MTL [81] and in fMRI studies in healthy human subjects (e.g., [16,60]). These data present converging evidence that the hippocampus is necessary for the association of an item and its context, i.e. for recollection, while the adjacent perirhinal cortex supports familiarity-based recognition and parahippocampal regions provide the hippocampus with context-relevant information. The results from animal studies reported in the last paragraph indicate that stimulus specificity increases with learning both when individual stimuli are presented and when stimulus associations are learned. It is thus possible that both the perirhinal cortex and the hippocampus promote learning via an inhibitory signal to upstream neocortical regions, but that only the hippocampus involves "clustering" of multiple responses of cells which are spatially remote initially [22] and induces correlations between neurons representing each individual stimulus [21,48]. Thus, the function of the hippocampus during recollection based learning, or generally during learning of associations, might be to provide a twofold signal to modality-specific regions—both inhibitory, to increase the selectivity of stimulus representations, and associative, modifying the network organization in modality-specific neocortical regions.

8. Conclusions and future work

To summarize, we suggest that during memory formation, neural representations in the MTL (perirhinal cortex in the case of single items; hippocampus in the case of associations) become more specific due to increased inhibitory activity. Subsequently, also neocortical representations become the target of this inhibitory signal, so that stimulus-specific responsiveness in these regions decreases. On a cellular level, it might be assumed that inhibition modifies synaptic connections via long-term depression (LTD); this would be in line with recent observations that learning-related exploration of novel objects-environment associations induces hippocampal LTD [36], which might refine stimulus representations. For the future, it will be desirable to directly test this idea. The dependence of learning-related changes in neocortical stimulus representations on the integrity of the hippocampus can be tested, for instance, in clinical studies of human patients, e.g., with hippocampal sclerosis. Furthermore, the idea needs to be more directly established that the hippocampus sends an inhibitory signal; this requires simultaneous measurements of local field potentials and single neurons in the hippocampus during learning. Another idea would be to measure the concentration of the inhibitory neurotransmitter GABA in the hippocampus during memory formation,

either using microdialysis or by MR spectroscopy [51]. Finally, it would be ideal to correlate hippocampal activity with learning-related changes in neocortical representations by simultaneous recordings in neocortical regions and in the hippocampus. This would also allow to test the idea that the hippocampus both sends an inhibitory signal to the neocortex and involves synchronization within neocortical assemblies.

Acknowledgments

The authors would like to thank Michael Cohen and Rodrigo Quian-Quiroga for critically reading the manuscript and helpful suggestions. We gratefully acknowledge a grant from the Volkswagen Foundation (I-79878).

References

[1] Aggleton JP, Shaw C, Gaffan EA. The performance of postencephalitic amnesic subjects on two behavioural tests of memory: concurrent discrimination learning and delayed matching-to-sample. Cortex 1992;28:359–72.
[2] Allison T, McCarthy G, Nobre A, Puce A, Belger A. Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. Cereb Cortex 1994;4:544–54.
[3] Axmacher N, Mormann F, Fernandez G, Elger CE, Fell J. Memory formation by neuronal synchronization. Brain Res Rev 2006;52:170–82.
[4] Axmacher N, Haupt S, Fernandez G, Elger CE, Fell J. The role of sleep in declarative memory consolidation—direct evidence by intracranial EEG. Cereb Cortex 2007:14 [Epub ahead of print].
[5] Axmacher N, Mormann F, Fernandez G, Cohen MX, Elger CE, Fell J. Sustained neural activity patterns during working memory in the human medial temporal lobe. J Neurosci 2007;27:7807–16.
[6] Baddeley A. Working memory. Science 1992;255:556–9.
[7] Baker CI, Behrmann M, Olson CR. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. Nat Neurosci 2002;5:1210–6.
[8] Birbaumer N, Elbert T, Canavan AG, Rockstroh B. Slow potentials of the cerebral cortex and behavior. Physiol Rev 1990;70:1–41.
[9] Brazdil M, Dobsik M, Mikl M, Hlustik P, Daniel P, Pazourkova M, et al. Combined event-related fMRI and intracerebral ERP study of an auditory oddball task. Neuroimage 2005;26:285–93.
[10] Buzsaki G. Two-stage model of memory trace formation: a role for "noisy" brain states. Neuroscience 1989;31:551–70.
[11] Cameron KA, Yashar S, Wilson CL, Fried I. Human hippocampal neurons predict how well word pairs will be remembered. Neuron 2001;30:289–98.
[12] Chalmers DJ. What is a neural correlate of consciousness? In: Metzinger T, editor. Neural correlates of consciousness. Cambridge, MA: MIT Press; 2000. p. 17–39.
[13] Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, et al. Temporal dynamics of brain activation during a working memory task. Nature 1997;386:604–8.
[14] Cohen NJ, Squire LR. Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. Science 1980;210:207–10.
[15] Courtney SM, Ungerleider LG, Keil K, Haxby JV. Transient and sustained activity in a distributed neural system for human working memory. Nature 1997;386:608–11.
[16] Davachi L, Mitchell JP, Wagner AD. Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. Proc Natl Acad Sci USA 2003;100:2157–62.
[17] Degonda N, Mondadori CR, Bosshardt S, Schmidt CF, Boesiger P, Nitsch RM, et al. Implicit associative learning engages the hippocampus and interacts with explicit associative learning. Neuron 2005;46:505–20.
[18] Desimone R, Schein SJ, Moran J, Ungerleider LG. Contour, color and shape analysis beyond the striate cortex. Vis Res 1985;25:441–52.
[19] Diana RA, Reder LM, Arndt J, Park H. Models of recognition: a review of arguments in favor of a dual-process account. Psychon Bull Rev 2006;131:1–21.
[20] Eichenbaum H, Yonelinas AP, Ranganath C. The medial temporal lobe and recognition memory. Annu Rev Neurosci 2007;30:123–52.
[21] Erickson CA, Desimone R. Responses of macaque perirhinal neurons during and after visual stimulus association learning. J Neurosci 1999;19:10404–16.
[22] Erickson CA, Jagadeesh B, Desimone R. Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. Nat Neurosci 2000;3:1143–8.
[23] Fell J, Klaver P, Lehnertz K, Grunwald T, Schaller C, Elger CE, et al. Human memory formation is accompanied by rhinal–hippocampal coupling and decoupling. Nat Neurosci 2001;4:1259–64.
[24] Fell J, Klaver P, Elfadil H, Schaller C, Elger CE, Fernández G. Rhinal–hippocampal theta coherence during declarative memory formation: interaction with gamma synchronization? Eur J Neurosci 2003;17:1082–8.
[25] Fernandez G, Weyerts H, Tendolkar I, Smid HG, Scholz M, Heinze HJ. Event-related potentials of verbal encoding into episodic memory: dissociation between the effects of subsequent memory performance and distinctiveness. Psychophysiology 1998;35:709–20.
[26] Fernandez G, Effern A, Grunwald T, Pezer N, Lehnertz K, Dumpelmann M, et al. Real-time tracking of memory formation in the human rhinal cortex and hippocampus. Science 1999;285:1582–5.
[27] Freedman DJ, Riesenhuber M, Poggio T, Miller EK. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. Cereb Cortex 2006;16:1631–44.
[28] Fried I, MacDonald KA, Wilson CL. Single neuron activity in human hippocampus and amygdala during recognition of faces and objects. Neuron 1997;18:753–65.
[29] Hampson RE, Simeral JD, Deadwyler SA. Distribution of spatial and nonspatial information in dorsal hippocampus. Nature 1999;402:610–4.
[30] Hannula DE, Tranel D, Cohen NJ. The long and the short of it: relational memory impairments in amnesia, even at short lags. J Neurosci 2006;26:8352–9.
[31] Hasselmo ME. Neuromodulation: acetylcholine and memory consolidation. Trends Cogn Sci 1999;3:351–9.
[32] Helmstaedter C, Kurthen M. Memory and epilepsy: characteristics, course, and influence of drugs and surgery. Curr Opin Neurol 2001;14:211–6.
[33] Holscher C, Anwyl R, Rowan MJ. Stimulation on the positive phase of hippocampal theta rhythm induces long-term potentiation that can be depotentiated by stimulation on the negative phase in area CA1 in vivo. J Neurosci 1997;17:6470–7.
[34] Howard MW, Rizzuto DS, Caplan JB, Madsen JR, Lisman J, Aschenbrenner-Scheibe R, et al. Gamma oscillations correlate with working memory load in humans. Cereb Cortex 2003;13:1369–74.
[35] Huerta PT, Lisman JE. Heightened synaptic plasticity of hippocampal CA1 neurons during a cholinergically induced rhythmic state. Nature 1993;364:723–5.
[36] Kemp A, Manahan-Vaughan D. Hippocampal long-term depression and long-term potentiation encode different aspects of novelty acquisition. Proc Natl Acad Sci USA 2004;101:8192–7.
[37] Kim DS, Ronen I, Olman C, Kim SG, Ugurbil K, Toth LJ. Spatial relationship between neuronal activity and BOLD functional MRI. Neuroimage 2004;21:876–85.
[38] Knight R. Contribution of human hippocampal region to novelty detection. Nature 1996;383:256–9.
[39] Konorski J. Integrative activity of the brain: an interdisciplinary approach. University of Chicago Press; 1967.
[40] Kreiman G, Koch C, Fried I. Category-specific visual responses of single neurons in the human medial temporal lobe. Nat Neurosci 2000;3:946–53.
[41] Letzkus JJ, Kampa BM, Stuart GJ. Does spike timing-dependent synaptic plasticity underlie memory formation? Clin Exp Pharmacol Physiol 2007;34:1070–6.
[42] Leutgeb S, Leutgeb JK, Moser MB, Moser EI. Place cells, spatial maps and the population code for memory. Curr Opin Neurobiol 2005;15:738–46.

[43] Li L, Miller EK, Desimone R. The representation of stimulus familiarity in anterior inferior temporal cortex. J Neurophysiol 1993;69:1918–29.
[44] Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature 2001;412:150–7.
[45] MacKay DG, Stewart R, Burke DM. H.M. revisited: relations between language comprehension, memory, and the hippocampal system. J Cogn Neurosi 1998;10:377–94.
[46] Marr D. A theory of cerebellar cortex. J Physiol 1969;202:437–70.
[47] McCartney H, Johnson AD, Weil ZM, Givens B. Theta reset produces optimal conditions for long-term potentiation. Hippocampus 2004;14:684–7.
[48] Messinger A, Squire LR, Zola SM, Albright TD. Neuronal representations of stimulus associations develop in the temporal lobe during learning. Proc Natl Acad Sci USA 2001;98:12239–44.
[49] Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. Science 2005;309:951–4.
[50] Nichols EA, Kao YC, Verfaellie M, Gabrieli JD. Working memory and long-term memory for faces: evidence from fMRI and global amnesia for involvement of the medial temporal lobes. Hippocampus 2006;16:604–16.
[51] Northoff G, Walter M, Schulte RF, Beck J, Dydak U, Henning A, et al. GABA concentrations in the human anterior cingulate cortex predict negative BOLD responses in fMRI. Nat Neurosci 2007;10:1515–7.
[52] Olson IR, Page K, Moore KS, Chatterjee A, Verfaellie M. Working memory for conjunctions relies on the medial temporal lobe. J Neurosci 2006;26:4596–601.
[53] Penfield W, Milner B. Memory deficit produced by bilateral lesions in the hippocampal zone. AMA Arch Neurol Psychiatry 1958;79:475–97.
[54] Piekema C, Kessels RP, Mars RB, Petersson KM, Fernandez G. The right hippocampus participates in short-term memory maintenance of object-location associations. Neuroimage 2006;33:374–82.
[55] Quiroga RQ, Reddy L, Kreiman G, Koch C, Fried I. Invariant visual representation by single neurons in the human brain. Nature 2005;435:1102–7.
[56] Quiroga RQ, Reddy L, Koch C, Fried I. Decoding visual inputs from multiple neurons in the human temporal lobe. J Neurophysiol 2007;98:1997–2007.
[57] Rainer G, Miller EK. Effects of visual experience on the representation of objects in the prefrontal cortex. Neuron 2000;27:179–89.
[58] Ranganath C, Blumenfeld RS. Doubts about double dissociations between short- and long-term memory. Trends Cogn Sci 2005;9:374–80.
[59] Ranganath C, D'Esposito M. Medial temporal lobe activity associated with active maintenance of novel information. Neuron 2001;31:865–73.
[60] Ranganath C, Yonelinas AP, Cohen MX, Dy CJ, Tom SM, D'Esposito M. Dissociable correlates of recollection and familiarity within the medial temporal lobes. Neuropsychologia 2003;42:2–13.
[61] Redish AD, Battaglia FP, Chawla MK, Ekstrom AD, Gerrard JL, Lipa P, et al. Independence of firing correlates of anatomically proximate hippocampal pyramidal cells. J Neurosci 2001;21:RC134.
[62] Reddy L, Kanwisher N. Coding of visual objects in the ventral stream. Curr Opin Neurobiol 2006;16:408–14.
[63] Rutishauser U, Mamelak AN, Schuman EM. Single-trial learning of novel stimuli by individual neurons of the human hippocampus–amygdala complex. Neuron 2006;49:805–13.
[64] Schmolck H, Kensinger EA, Corkin S, Squire LR. Semantic knowledge in patient H.M. and other patients with bilateral medial and lateral temporal lobe lesions. Hippocampus 2002;12:520–33.
[65] Schon K, Hasselmo ME, Lopresti ML, Tricarico MD, Stern CE. Persistence of parahippocampal representation in the absence of stimulus input enhances long-term encoding: a functional magnetic resonance imaging study of subsequent memory after a delayed match-to-sample task. J Neurosci 2004;24:11088–97.
[66] Schupp HT, Lutzenberger W, Rau H, Birbaumer N. Positive shifts of event-related potentials: a state of cortical disfacilitation as reflected by the startle reflex probe. Electroencephalogr Clin Neurophysiol 1994;90:135–44.
[67] Sederberg PB, Schulze-Bonhage A, Madsen JR, Bromfield EB, McCarthy DC, Brandt A, et al. Hippocampal and neocortical gamma oscillations predict memory formation in humans. Cereb Cortex 2007;17:1190–6.
[68] Speckmann EJ, Elger CE. Introduction to the neurophysiological basis of the EEG and DC potentials. In: Niedermeyer E, Lopes da Silva F, editors. Electroencephalography. 4th ed. Baltimore: Williams & Wilkins; 1999. p. 15–27.
[69] Squire LR. Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. Psychol Rev 1992;99:195–231.
[70] Squire LR, Stark CE, Clark RE. The medial temporal lobe. Annu Rev Neurosci 2004;27:279–306.
[71] Strange BA, Otten LJ, Josephs O, Rugg MD, Dolan RJ. Dissociable human perirhinal, hippocampal, and parahippocampal roles during verbal encoding. J Neurosci 2002;22:523–8.
[72] Suzuki WA, Miller EK, Desimone R. Object and place memory in the macaque entorhinal cortex. J Neurophysiol 1997;78:1062–81.
[73] Tulving E. How many memory systems are there? Am Psychol 1985;40:385–98.
[74] Viskontas IV, Knowlton BJ, Steinmetz PN, Fried I. Differences in mnemonic processing by neurons in the human hippocampus and parahippocampal regions. J Cogn Neurosi 2006;18:1654–62.
[75] Waydo S, Kraskov A, Quian Quiroga R, Fried I, Koch C. Sparse representation in the human medial temporal lobe. J Neurosci 2006;26:10232–4.
[76] Whittington MA, Traub RD, Jefferys JG. Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. Nature 1995;373:612–5.
[77] Wiltgen BJ, Brown RA, Talton LE, Silva AJ. New circuits for old memories: the role of the neocortex in consolidation. Neuron 2004;44:101–8.
[78] Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, Suzuki WA. Single neurons in the monkey hippocampus and learning of new associations. Science 2003;300:1578–81.
[79] Yang J, Weng X, Guan L, Kuang P, Zhang M, Sun W, et al. Involvement of the medial temporal lobe in priming for new associations. Neuropsychologia 2003;41:818–29.
[80] Yanike M, Wirth S, Suzuki WA. Representation of well-learned information in the monkey hippocampus. Neuron 2004;42:477–87.
[81] Yonelinas AP, Kroll NE, Quamme JR, Lazzara MM, Sauve MJ, Widaman KF, et al. Effects of extensive temporal lobe damage or mild hypoxia on recollection and familiarity. Nat Neurosci 2002;5:1236–41.