Theta theory: Requirements for encoding events and task rules explain theta phase relationships in hippocampus and neocortex

Michael E. Hasselmo
Department of Psychology Center for Memory and Brain,
Program in Neuroscience and Center for BioDynamics,
Boston University, 64 Cummington St., Boston,, MA 02215
hasselmo@bu.edu. http://people.bu.edu/hasselmo

Abstract.
Memory tasks require processes for encoding associations and retrieving them to guide goal-directed behavior (Hasselmo et al., 2002). Simulations of the interaction of hippocampus and neocortex demonstrate specific dynamical requirements required for encoding episodic memory and learning task rules. 1.) Encoding of new information without interference from previously encoded information requires cyclical transitions between encoding and retrieval states, 2.) Context dependent retrieval of task relevant information from episodic memory requires mechanisms for timing and synchronization of inputs from entorhinal cortex and region CA3 to hippocampal region CA1 during retrieval. Theta theory provides a framework for describing how performance is enhanced by specific phase relationships of theta rhythm in different regions of the rat hippocampus and neocortex.

I INTRODUCTION

The hippocampal theta rhythm is a 3-10 Hz oscillation which appears prominently in the electroencephalograph (EEG) recorded at the hippocampal fissure. This theta rhythm shows clear dependence upon behavior, with a large increase in the amplitude of theta rhythm during locomotion and exploration (Buzsaki et al., 1983), and also during attention to behaviorally relevant stimuli such as predators or the tone in appetitive or eye blink conditioning paradigm (Seager et al., 2001). Lesions which block theta rhythm impair performance in a range of memory tasks, including spatial alternation, 8-arm radial maze and spatial reversal.

Recent research in my laboratory has focused on the possible functional role of the theta rhythm within the hippocampus. Theta rhythm results from oscillatory changes in synaptic currents in different layers of region CA1 of the hippocampus, as shown by current source density analysis (Brankack et al., 1993; Bragin et al., 1998). This current source density analysis is summarized in Figure 1. In particular, at the trough the EEG recorded in the hippocampal fissure, there is a strong current sink in stratum lacunosum-moleculare, the layer which receives excitatory synaptic input from the entorhinal cortex. In contrast, during this phase, there is a strong current source in the cell body layer, stratum pyramidale, reflecting the outward currents resulting from inhibition at the cell bodies. At the opposite phase of the theta rhythm, there is a current sink in stratum radiatum, associated with strong synaptic input from region CA3.

The mechanisms of generation of hippocampal theta rhythm have been studied extensively by a number of researchers (Stewart and Fox, 1990; Buzsaki et al., 1983). However, few researchers have described models of the functional role of hippocampal theta rhythm. In recent work, I have described a model of the functional role of hippocampal theta rhythm (Hasselmo et al., 2002a; 2002b; 2002c; Cannon et al., 2002). This theta theory provides a functional framework in which a behavioral performance measure is maximized by the relative phases of different oscillatory dynamics within different regions of the hippocampus and associated structures. This paper will extend this theta theory to address further phase relationships between different sets of physiological data on the hippocampus.

II FUNCTION OF THETA RHYTHM

A. A simple example. To make the functional role of theta rhythm clear, I will provide an extremely simple example of its possible role in reducing interference during learning of new paired associates. In a typical paired associate memory task, subjects are given a set of word pairs A-B, such as leather-holster, and are tested on retrieval (e.g. what word was associated with leather?). They are then trained repeatedly on a second set of overlapping word pairs A-C, such as leather-boot, and are tested on a second retrieval (e.g. what word was most recently associated with leather?). If we consider each word A, B and C to be represented by column vectors a_A a_B and a_C and retrieval activity by a_{out}, the performance on the second retrieval phase can be measured by:

$$M = \sum[a_C^T a_{post} - a_B^T a_{post}]\,. \tag{1}$$

Now consider a simple heteroassociative memory (e.g. Kohonen, 1984) which stores the word pairs and is then cued with the input a_A to generate the response a_{out}. In most models, interference during encoding is prevented by using specialized dynamics for encoding and retrieval. The simplified example presented here focuses on showing the interference effects which could occur in a neural system, and how theta rhythmicity could solve these problems. Note that in this example, we will always assume vectors of unit length.

Consider a network with two populations a_{post} and a_{pre} with all initial weights W(0)=0. The network receives afferent input representing the two words and undergoes Hebbian synaptic modification analogous to Hebbian long-term potentiation:

$$a_{post} = a_B + W a_A \qquad a_{pre} = a_A \tag{2}$$

$$\Delta W = a_{post} a_{pre} = a_B a_A^T \tag{3}$$

Note that when a second overlapping paired associate is stored in the network, synaptic modification will include interference from retrieval of the first paired associate:

$$a_{post} = a_C + W a_A = a_C + a_B a_A^T a_A = a_C + a_B \tag{4}$$

$$\Delta W = (a_C + a_B) a_A^T, W = (a_C + a_B) a_A^T + a_B a_A^T \tag{5}$$

Subsequently, retrieval cued by input of word A will not effectively differentiate the recent word, resulting in poor performance shown by the performance measure.

$$a_{post} = W a_A = a_C + a_B + a_B \tag{6}$$

$$M = \sum[a_C^T a_{post} - a_B^T a_{post}] = -1 \tag{7}$$

B. Oscillatory dynamics. This extremely simple example can be used to illustrate the potential functional role of theta rhythm in providing appropriate dynamics for encoding and retrieval within neural circuitry of the hippocampal formation. The interference from previous retrieval during encoding can be prevented by assuming oscillatory functions for each component of this model.

$$\theta_{IN}(t) = x/2 * \sin(t + \phi_{IN}) + (1 - x/2) \tag{8}$$

$$\theta_{LTP}(t) = \sin(t + \phi_{LTP}) \tag{9}$$

$$\theta_{RET}(t) = x/2 * \sin(t + \phi_{RET}) + (1 - x/2) \tag{10}$$

Learning of the first pattern takes the form:

$$a_{post} = \theta_{IN} a_B + \theta_{RET} W a_A \quad a_{pre} = a_A \tag{11}$$

$$\Delta W = \int_0^{2\pi} \theta_{LTP}(\theta_{IN} a_B + \theta_{RET} W a_A) a_A^T \tag{12}$$

We will assume first pattern learning reaches a maximum strength of W(0)=k. Then learning of the second pattern takes the form:

$$a_{post} = \theta_{IN} a_C + \theta_{RET} k a_B a_A^T a_A \tag{13}$$

$$\Delta W = \int_0^{2\pi} \theta_{LTP}(\theta_{IN} a_C + \theta_{RET} k a_B) a_A^T \tag{14}$$

Subsequently, we will evaluate retrieval cued by input of word A (note that we will take retrieval at the maximum of the oscillatory phase). $a = \max[\theta_{RET} W a_A]$

When the integral over full cycles is computed, the performance measure takes the form.

$$M = (x/2)\pi \cos(\phi_{LTP} - \phi_{IN}) - k - (x/2)\pi \cos(\phi_{LTP} - \phi_{RET})$$

This function is plotted in Figure 1. As can be seen from the figure, the best function of the network is obtained when input is in phase with the phase of long-term potentiation, and when retrieval RET is 180 degrees out of phase with long-term potentiation (Hasselmo et al., 2002a).

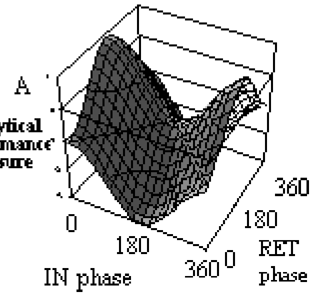


Figure 1. Magnitude of the performance measure M (z axis) for different phase relationships between the input (IN) and the learning rate (LTP) and between the retrieval (RET) and the learning rate (LTP).

In the description above, the encoding of the second paired associate C is accompanied by forgetting of the first paired associate B. However, in an actual paired associate task subjects would be able to retrieve both B and C in a context-dependent manner, suggesting that B is not forgotten. In this case, correct retrieval of C or B depends upon an association of the experimental context with the specific paired associate. The full details of this will not be presented here, but we can assume that this experimental context involves modified connections providing input which is selective for one pattern. Previously, this interaction was used for goal-directed selection of next desired location (Hasselmo et al., 2002b). Here we will just represent context dependent cue activity as oscillatory input biased toward one response $\theta_{IN} a_C$. The retrieval activity of the network will also be regulated by oscillations in the activation of the postsynaptic population θ_{DEP} as follows:

$$a_{post} = \int_0^{2\pi} \theta_{DEP}[\theta_{IN} a_C + \theta_{RET} W a_A] \qquad (15)$$

With this activation function the performance measure changes with phase relationships in the following manner:

$$\begin{aligned} M &= (x^2/4)\pi\cos(\phi_{post} - \phi_{IN}) + 2\pi(1 - x/2)^2 + \\ &[(x^2/4)\pi\cos(\phi_{post} - \phi_{RET}) + 2\pi(1 - x/2)^2] \\ &\times[(x/2)\pi\cos(\phi_{LTP} - \phi_{IN}) - k - (x/2)\pi\cos(\phi_{LTP} - \phi_{RET})] \end{aligned} \qquad (16)$$

This form of the equation will reach a maximum when postsynaptic depolarization is in phase with the retrieval spread RET. The phase relationships of post=RET are plotted in the top of Figure 2, which demonstrates their general consistency with the timing of synaptic currents recorded electrophysiologically in the hippocampus (See Figure 2).). If we assume that that the oscillatory modulation of LTP induction must stay positive (no LTD), then the phase relationships for maximum function are shifted slightly relative to each other similar to the data at the bottom of Figure 2. Note that the modulation of postsynaptic depolarization is constrained to stay positive, though spiking activity in the dendrites during encoding might cause an explicit suppression of repeated patterns during retrieval due to a refractory period after dendritic spiking (Golding et al., 2002

The cues for retrieval are presumed to be held in working memory and delivered from neocortical areas such as prefrontal cortex. The same steps described above can be used to demonstrate that this would require rhythmic activity in prefrontal cortex which is in phase with the hippocampal retrieval activity. Specifically, phasic activity in prefrontal cortex working memory would appear in equations as θ_{PFC} with the same maximal phase as demonstrated above for θ_{DEP}.

$$a_{post} = \int_0^{2\pi} [\theta_{IN}\theta_{PFC} a_A + \theta_{RET} W \theta_{PFC} a_A] \qquad (17)$$

Ongoing experiments in our laboratory have shown that at least in some periods prefrontal and hippocampal theta rhythm are synchronized (Hyman et al., 2002).

C. Electrophysiological data. This use of oscillatory variables to regulate dynamics is directly motivated by experimental evidence from electrophysiological recordings within the hippocampal formation of awake behaving rats. In fact, the oscillatory variables described above appear to resemble evidence from current source density showing phasic changes in magnitude of synaptic input in different layers of the hippocampal region CA1 (Brankack et al., 1993). In particular, the phase of input IN could correspond to the excitatory afferent input from entorhinal cortex causing current sinks in stratum lacunosum moleculare of region CA1, the phase of retrieval RET could correspond to afferent input from region CA3 causing current sinks in stratum radiatum of region CA1(Brankack et al., 1993), and the phase of postsynaptic depolarization DEP could correspond to the phase of rhythmic somatic depolarization (Fox, 1989; Kamondi et al., 1995) appearing when inhibition is weak within stratum pyramidale of region CA1. The phase relationships for maximum performance are summarized in Figure 2, and the bottom of the figure shows how these phase relationships match the current source density data showing the relative timing of current sinks determined by current source density analysis in region CA1 of the hippocampal formation.

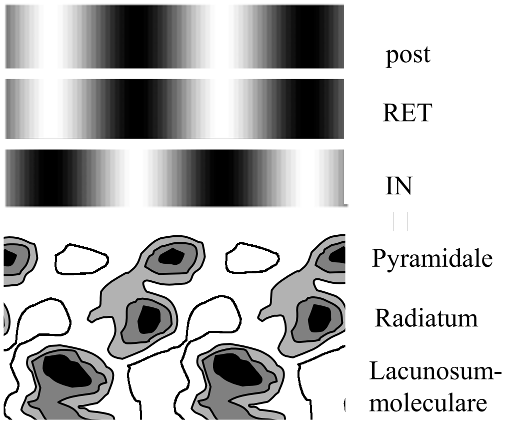


Fig. 2 – Top: Phases for maximum function in equation 16. Bottom: Current source density data showing phases in region CA1 of hippocampus (Brankack et al., 1993).

These phase relationships are behaviorally relevant, as lesions of region CA1 of the hippocampus have been shown to selectively impair performance in paired associate memory tasks in humans (Rempel-Clower et al., 1996). In rats, lesions of the fornix which remove much of the amplitude of theta rhythm impair the ability to learn reversal tasks, in which a previously rewarded behavior must be replaced with an opposite behavior. For example, rats with fornix lesions have difficult learning to make a right turn response after initially learning a left turn response (M'Harzi et al., 1987; Markowska et al., 1989).

These phasic changes require that long-term potentiation (LTP) at synapses arising from region CA3 should be strongest at the phase when synaptic transmission at these synapses is weakest. Consistent with these requirements, our recent data (Wyble et al., 2000) shows that synaptic transmission in stratum radiatum is weakest at the positive peak of local theta, which is when previous data shows induction of LTP is strongest in this layer (Holscher et al., 1997; Orr et al., 2001; Hyman et al., 2002). LTP induction could be due to induction of dendritic spikes by entorhinal input at a time when the soma is hyperpolarized to prevent spiking (Golding et al., 2002).

III GOAL DIRECTED SELECTION OF NEXT MOVEMENT

A The hippocampal circuit. As noted above, it is desirable to maintain representations of multiple previous associations, rather than forgetting them. In this case, it is necessary to regulate retrieval so that only individual associations are retrieved based on the specific context or goal of one phase of a task. In the description above, we simply assumed context dependent input could bias the retrieval toward one output. Here I will present the circuitry which could perform this context dependent retrieval in more detail. This has been addressed in studies modeling the role of hippocampus in goal directed spatial navigation. Recent papers from my laboratory have addressed various aspects of this goal directed spatial navigation (Hasselmo et al., 2002a, 2002b; Cannon et al., 2002; Koene et al., 2003; Gorchetchnikov et al., 2003). The overall circuitry of such models is summarized in Figure 3, which schematically illustrates the connectivity of different subregions of entorhinal cortex and hippocampus modeled in these simulations.

This basic structure has been used to guide the movements of a virtual rat through a virtual environment. The virtual rat learns pathways through the environment by encoding associations between sequentially activated place cells – that is neurons responding selectively when the rat is in specific spatial locations (Muller et al., 1986; Skaggs et al., 1996). These pathways can then be used to guide movement in an interactive manner. Activity in EC III spreads in a reverse direction from goal location until it converges with current location, at which time it activates forward associations in region CA3. These two regions send convergent input to region CA1, where neurons are selectively activated on the basis of the next desired location of the rat.

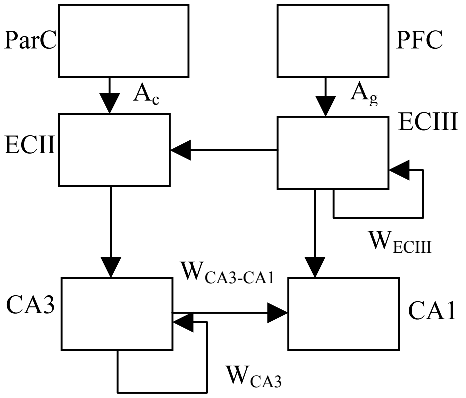


Figure 3. Schematic diagram of the hippocampal formation. The entorhinal cortex provides most input to the hippocampus, receiving convergent input from many regions of association cortex. This figure highly simplifies the input, showing the connections bringing spatial input from parietal cortex (ParC) to entorhinal cortex layer II (ECII) and goal dependent input from prefrontal cortex (PFC) to layer III (ECIII). These regions provide input to region CA3 and region CA1 of the hippocampus.

This basic functional framework raises a number of dynamical issues which have been addressed in a number of publications from this laboratory (Hasselmo et al., 2002a, 2002b, Cannon et al., 2002; Koene et al., 2003; Gorchetchnikov et al., 2003). Among other things, these models require a buffer in entorhinal cortex which can maintain spiking activity so that slow behavioral transitions can cause sequential spike within the time course of spike timing dependent synaptic plasticity (Jensen and Lisman, 1996; Fransen et al., 2002; Hasselmo et al., 2002c).

B. Theta phase precession. Among other phenomena, these circuits underlie the experimental phenomenon of theta phase precession. Theta phase precession is a characteristic of place cells in linear environments – for example when a rat must run on along an elevated triangular track in a single direction (O'Keefe and Recce, 1993; Skaggs et al., 1996). In such tasks, many place cells show characteristic changes in phase, firing late in the theta cycle when first entering that cells place field, and then gradually firing at earlier and earlier phases as the rat traverses the place field. This phenomenon has been modeled in a number of different ways, including models based on activity spreading along sequences of place cell representations. Two of these models are summarized in Figs. 4 and 5. Both of these models can be thought of as feedforward models of phase precession. They depend upon current location, but do not depend upon the goal or context of behavior.

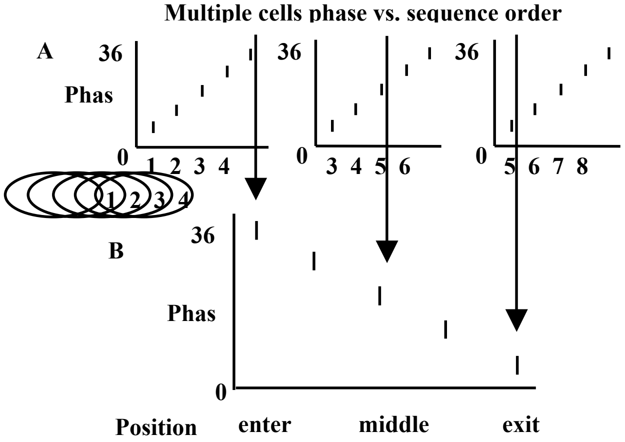


Fig. 4. Simple sequential model of phase precession (Jensen and Lisman, 1996; Tsodyks et al., 1996). A. Associations between place cells allow the current location to cause readout of a sequence of future locations. When observing a single cell (5), this initially appears as firing late in cycle, which moves to earlier phases of theta.

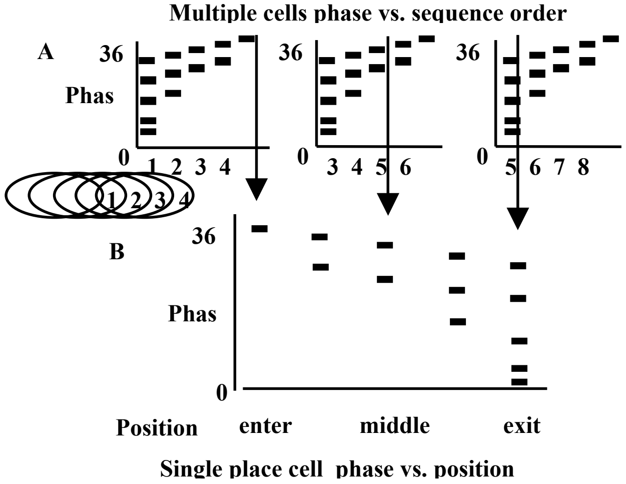
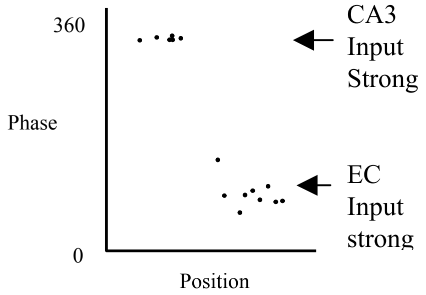


Fig. 5. Theory of phase precession based on changes in length of retrieval sequence (Wallenstein and Hasselmo, 1997). A. In this model, sequences are read out to different lengths depending on strength of synaptic transmission or depolarization at different phases of theta. B. When observing a single cell (5), this appears as a wide, slightly scalloped distribution of spiking.

In more recent modeling (Hasselmo et al., 2002b, 2002c), theta phase precession has been obtained which depends upon the interaction of associations from the goal in entorhinal cortex with associations from current location in region CA3. This generates predictive spiking in the late phases of theta, and spiking driven by input in early phases of theta, as shown in Figure 6. This figure shows spiking of a single place cell in the simulations, with spike times plotted relative to theta phase in the model and the position of the virtual rat in the virtual environment. In initial models, this predictive phase precession has a biphasic nature to it, but recent simulations show a smoother transition of phase precession (Koene et al., 2003).



Cortical rule learning. Performance of most behavioral tasks requires learning of rules associated with these tasks. Thus, effective simulation of hippocampal dependent behavior requires a sophisticated model of rule learning in neocortical structures as well. Recent simulations have extended a similar framework from Fig. 3 to modeling of neocortical interactions between goal-dependent feedback and forward associations from current location. This allows learning of complex relationships between elements of sensory input and motor output, but continues to require separation of sensory input phases from goal-dependent feedback activity. In particular, this simulation has been used to model generation of task dependent behavior in the standard spatial alternation task. The spatial alternation task shows strong impairments with lesions of the hippocampus, probably due to loss of episodic memory of previous responses in the task. Recent electrophysiological studies have demonstrated selective neuronal spiking activity dependent upon prior or future responses in this task (Wood et al., 2000) and ongoing simulations and experiments are addressing specific mechanisms of this selective spiking activity.

Acknowledgements: This work supported by NIMH MH60013, MH61492, MH60450 and NIDA DA16454.

References.

Bragin A, Jando, G, Nadasdy, Z, Hetke, J, Wise, K, Buzsaki, G. Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *Journal of Neuroscience*, 15: 47-60 (1995).

Brankack, J, Stewart, M, Fox, SE (1993). Current source density analysis of the hippocampal theta rhythm: associated sustained potentials and candidate synaptic generators. *Brain Research*, 615(2): 310-327

Buzsaki G, Leung LW, Vanderwolf CH. (1983) Cellular bases of hippocampal EEG in the behaving rat. Brain Res. 287(2):139-71.

Cannon, R.C., Hasselmo, M.E. and Koene, R.A. (2002) From biophysics to behavior: Catacomb2 and the design of biologically plausible models for spatial navigation. Neuroinformatics 1(1): 3-42.

Fox, S. E. (1989). Membrane potential and impedence changes in hippocampal pyramidal cells during theta rhythm. *Exp. Brain Res.*, **77**, 283-294.

Givens B (1996) Stimulus-evoked resetting of the dentate theta rhythm: relation to working memory. Neuroreport 8: 159-163.

Golding NL, Staff NP, Spruston N (2002) Dendritic spikes as a mechanism for cooperative long-term potentiation. Nature 418(6895): 326-331.

Gustafsson B, Wigstrom H, Abraham WC, Huang YY (1987) Long-term potentiation in the hippocampus using depolarizing current pulses as the conditioning stimulus to single volley synaptic potentials. J Neurosci. 7(3):774-80

Hasselmo, M.E., Bodelon, C. and Wyble, B.P. (2002a) A proposed function for hippocampal theta rhythm: Separate phases of encoding and retrieval enhance reversal of prior learning. Neural Computation 14(4): 793-817.

Hasselmo, M.E., Hay, J., Ilyn, M. and Gorchetchnikov, A. (2002b) Neuromodulation, theta rhythm and rat spatial navigation. Neural Networks 15: 689-707.

Hasselmo, M.E. and McClelland, J.L. (1999) Neural models of memory. Curr. Opinion Neurobiol. 9: 184-188.

Hasselmo, M.E., Wyble, B.P. and Cannon, R.C. (2002c) From spike frequency to free recall: How neural circuits perform encoding and retrieval. In: A. Parker, T.J. Bussey, E. Wilding (eds.) The cognitive neuroscience of memory:
Encoding and retrieval. London: Psychology Press

Holscher, C, Anwyl, R, Rowan, MJ (1997). 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*, 17(16): 6470-6477.

Huerta PT, Lisman JE (1995). Bidirectional synaptic plasticity induced by a single burst during cholinergic theta oscillation in CA1 in vitro. *Neuron*, 15(5): 1053-1063.

Hyman JM, Wyble BP, Goyal V, Hasselmo ME (2001) Phase relationship of LTP inducation and behavior to theta rhythm in the rat hippocampus. Soc. Neurosci. Abstr. 27.

Hyman, J., Wyble, B.P. Rossi, C.A., Hasselmo, M.E. (2002) Coherence between theta rhythm in rat medial prefrontal cortex and hippocampus. Soc. Neurosci. Abstr. 28: 477.6.

Jensen O, Lisman JE (1996) Novel lists of 7 +/- 2 known items can be reliably stored in an oscillatory short-term memory network: interaction with long-term memory. Learn Mem. 3(2-3):257-63.

Kamondi, A, Acsady, L, Wang, XJ, Buzsaki, G (1998). Theta oscillations in somata and dendrites of hippocampal pyramidal cells in vivo: activity-dependent phase-precession of action potentials. *Hippocampus*, 8(3): 244-61

Koene, R.A., Cannon, R.C. and Hasselmo, M.E. (2003) Goal-directed spatial navigation of the rat depends on phases of theta oscillation in hippocampal circuitry. Proc. IJCNN 2003.

Kohonen, T. (1984) Self-organization and Associative Memory. Springer-Verlag.

Macrides F., Eichenbaum, HB, Forbes, W.B (1982) Temporal relationship between sniffing and the limbic theta rhythm during odor discrimination reversal learning. *J. Neurosci.* 2: 1705-1717.

Markowska AL, Olton DS, Murray EA, Gaffan D (1989) A comparative analysis of the role of fornix and cingulate cortex in memory: rats. Exp Brain Res. 74(1):187-201.

M'Harzi, M., Palacios, A., Monmaur P, Willig F, Houcine O and Delacour J. (1987) Effects of selective lesions of fimbria-fornix on learning set in the rat. Physiol. & Behav. 40: 181-188.

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

Orr G, Rao G, Stevenson GD, Barnes CA, McNaughton BL (1999) Hippocampal synaptic plasticity is modulated by the theta rhythm in the fascia dentata of freely behaving rats. Soc. Neurosci. Abstr. 25: 2165 (864.14).

Rempel-Clower NL, Zola SM, Squire LR, Amaral DG (1996) Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. J. Neurosci. 16(16): 5233-5255.

Skaggs, W.E., McNaughton, B.L., Wilson, M.A. and Barnes, C.A. (1996) Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6, 149-172.

Sohal, V.S. and Hasselmo, M.E. (1998a) Changes in GABAB modulation during a theta cycle may be analogous to the fall of temperature during annealing. *Neural Computation* 10: 889-902.

Tsodyks MV, Skaggs WE, Sejnowski TJ and McNaughton BL (1996) Population dynamics and theta rhythm phase precession of hippocampal place cell firing: a spiking neuron model. Hippocampus 6(3): 271-280.

Wallenstein, G.V. and Hasselmo, M.E. (1997a) GABAergic modulation of hippocampal population activity: Sequence learning, place field development and the phase precession effect. *J. Neurophysiol.* 78(1): 393-408.

Wiener SI, Paul CA, Eichenbaum H. (1989) Spatial and behavioral correlates of hippocampal neuronal activity. J Neurosci. 9(8):2737-63.

Wood ER, Dudchenko PA, Robitsek RJ, Eichenbaum H. (2000) Hippocampal neurons encode information about different types of memory episodes occurring in the same location. Neuron. 27(3):623-33.

Wyble BP, Linster C, Hasselmo ME. (2000) Size of CA1-evoked synaptic potentials is related to theta rhythm phase in rat hippocampus. J. Neurophysiol. 83(4):2138-2144.