

Connection Science
Vol. 17, Nos. 1–2, March–June 2005, 167–183



Competition between cue response and place response: a model of rat navigation behaviour

RICARDO CHAVARRIAGA*, THOMAS STRÖSSLIN, DENIS SHEYNIKHOVICH
and WULFRAM GERSTNER

Ecole Polytechnique Fédérale de Lausanne (EPFL), School of Computer and Communication
Sciences, and Brain Mind Institute, 1015 Lausanne, Switzerland

Different neural systems are involved in animal navigation depending on the type of task. Experimental studies support the idea that the hippocampus is necessary to learn a spatial representation required to navigate toward hidden goals (place response), whereas the dorsolateral striatum is involved in the learning of stimulus–response associations when navigating toward visible (or cued) goals. These systems compete for action selection according to the characteristics of the task, previous experience (e.g. training procedure) or endogenous factors. This paper reviews both experimental data on the theory of multiple memory systems involved in navigation, and a recent computational model of action selection based on the competition of place and cue responses learnt during training. The model implements separately the two types of response, i.e. place response and stimulus response. Furthermore, competition takes place to select which behaviour will actually be performed. The model was tested in a simulated environment using a protocol analogous to those used in experiments with animals.

Keywords: Action selection; Place response; Cue response; Reinforcement learning; Hippocampus; Striatum; Biometric agents

1. Introduction

Experimental evidence supports the idea of multiple memory systems in the rat brain mediating different forms of learning (Packard *et al.* 1989, Devan and White 1999, Ragozzino *et al.* 2001, Packard and Knowlton 2002, White and McDonald 2002, Gold 2004). Some of these studies aim to dissociate the role of different brain areas by observing how localized lesions reduce the performance in one type of tasks, while leaving other tasks unaffected. In the context of navigation, it has been shown that rats with hippocampal lesions are unable to navigate toward hidden goals (Morris 1981, Redish 1999, Pearce *et al.* 1998), whereas lesions in the dorsal striatum (Caudato–putamen, CPu) affect the learning of stimulus–response (S–R) associations required to navigate toward visible cues (Packard and Knowlton 2002).

*Corresponding author. Email: ricardo.chavarriaga@a3.epfl.ch

Navigating toward hidden goals requires the use of a representation of space. Such a representation is the basis of the cognitive map theory (Tolman 1948), whose anatomical locus in rats has been proposed to be the hippocampus (O'Keefe and Nadel 1978). The finding of place sensitive cells (*place cells*) in this structure and neurobehavioural experiments support the idea that the hippocampus maintains an allocentric representation of space, which can be used for navigation (see Redish (1999) for a review). The use of such a representation to solve a navigation task is often termed *place response* or *locale* strategy.

On the other hand, the dorsal striatum has been usually related to the acquisition of S-R associations (Packard *et al.* 1989, Packard and McGaugh 1996, Packard and Knowlton 2002, White and McDonald 2002). This S-R association can be used to navigate toward cued goals (*cue response*), which can be classified as a *taxon* navigation strategy (Redish 1999). The S-R learning has also been associated with the incremental acquisition of habits (Packard and McGaugh 1996, Jog *et al.* 1999), where a stimulus elicits a stereotypic response or sequence of movements (*praxic navigation*).

Interestingly, lesions in the dorsomedial striatum, although they do not prevent spatial learning, produce a preference for cue responses (Whishaw *et al.* 1987, Devan and White 1999), suggesting that the dorsomedial striatum is, to some extent, mediating the expression of spatial responses. This speaks in favour of a further segregation of the striatum, as opposed to the widely used distinction between ventral versus dorsal (Voorn *et al.* 2004).

Models of animal behaviour have considered the role of the basal ganglia (including the striatum) in action selection (Gurney *et al.* 2001, Baldassarre 2002) and stimulus–reward associations (Schultz *et al.* 1997, Sutton and Barto 1998, Doya 2002, Schultz 2002).

Models of action selection can be classified into two different groups depending on the type of modelled behaviour. The first type of model proposes that the basal ganglia implement a competition among different stereotypic sensory motor programmes (Gurney *et al.* 2001, Baldassarre 2002). These models are based on the existence of segregated loops going from the cortex to the basal ganglia, and then back to the cortex in a highly topographic way (Mink 1996). These loops are thought to implement separate S-R associations which may be selected by a mechanism of selective inhibition on the output structures of the basal ganglia (direct/indirect pathways) (Chevalier and Deniau 1990, Mink 1996). Note that this is a competition between the same type of associations (S-R), not between different memory systems as described above.

The second class of models focuses on actions based on spatial representation. As mentioned above, the spatial representation is believed to reside in the hippocampus, and goal-directed navigation is achieved by mapping places to local actions (Gerstner and Abbott 1996, Recce and Harris 1996, Redish and Touretzky 1998, Gaussier *et al.* 2002, Hasselmo *et al.* 2002). Besides the hippocampal structure, the nucleus accumbens (in the ventral striatum) has also been thought to play a role in reward-based spatial learning. Several models of hippocampal navigation have proposed this structure to be responsible for the selection of actions (direction of movement) based on place coding information from the hippocampus and reward information from the ventral tegmental area (Burgess *et al.* 1994, Brown and Sharp 1995, Foster *et al.* 2000, Arleo *et al.* 2004).

Along these lines, we have previously proposed a model of hippocampal spatial learning that combines internal and external information to build a representation of space in a population of place cells (Arleo and Gerstner 2000). Place-based (i.e. locale) navigation was performed by a population of action cells whose activity was coded for the direction of movements leading to a (hidden) goal location. Simulated and real robots were used to test the performance of the model using both simple (one-dimensional vision) (Arleo and Gerstner 2000) and more complex external inputs (Arleo *et al.* 2004). Recently, we extended this model to allow both place-based strategies and simple cue response behaviours (S-R) to compete for

selection of actions (Chavarriga *et al.* 2005). This extended model is the one we use in this paper.

In the following, we first review experimental studies supporting the theory of multiple memory systems involved in navigation (section 2). Then we describe the model (section 3), followed by the results obtained using an experimental paradigm and the comparison with those obtained with animals (section 4).

2. Experimental approaches dissociating parallel navigation systems

Navigation tasks are often used to test mnemonic functions in rats. These tasks are designed in such a way that the rat has to use either spatial or cue information to solve them. The hidden-platform water maze (Morris 1981) is one of the most common experimental paradigms used to test spatial learning. It consists of a swimming pool filled with coloured water, in which the rat is trained to reach an escape platform submerged just below the surface of the water. As no single cue signals the location of the platform, the animal should use spatial information to solve the task.[†]

The learning of S-R associations can be tested by using a cue marking the location of the goal; this task can be solved by learning an approaching response to the cue (landmark guidance, or beaconing). The visible version of the water maze is one of these tasks. In this paradigm, the escape platform stands above the surface of the water, thus being visible. In some cases, a visual cue is added above the platform to increase its visibility. In radial mazes, a sensory cue (i.e. light) may be used to signal the correct arm (Packard *et al.* 1989). Other tasks test the ability to learn a sequence of movements, for instance, by locating the goal always at the same position relative to the starting position (*e.g.* in a radial maze, if the bait is always located in the arm to the left of the entry arm, the animal learns to repeat the same body turn, to the left, at the centre of the maze).

These paradigms have been used to validate the theory of multiple memory systems by observing the effect of localized lesions (*e.g.* in the hippocampus or dorsal striatum) in the preference for place or S-R strategies (Whishaw *et al.* 1987, Packard *et al.* 1989, Pearce *et al.* 1998, Da Cunha *et al.* 2003). As mentioned previously, substantial evidence supports the idea that the hippocampus mediates spatial learning, whereas the dorsal striatum is involved in S-R learning (see Redish 1999, Packard and Knowlton 2002, White and McDonald 2002, for reviews).

Some studies use a single experimental paradigm to test simultaneously the development of both cue and place response (Packard and McGaugh 1996, Oliveira *et al.* 1997, Devan and White 1999, Chang and Gold 2003b). Additionally, the interactions between the different systems have been studied by testing the facilitating effect of lesioning a system that is not related to the task to be solved (Packard *et al.* 1989, Schroeder *et al.* 2002, Chang and Gold 2003a). Another interesting line of research studies the correlation of neurochemical factors with the expression of different kinds of response (Ragozzino and Gold 1995, Colombo and Gallagher 1998, Janis *et al.* 1998, Bizon *et al.* 2003, Chang and Gold 2003b, McIntyre *et al.* 2003, Marriott and Korol 2003).

[†] Studies exist, however, which show that animals with hippocampal impairments can solve the hidden water maze task under certain training conditions. Some of these cases can be explained by a form of learning, different from spatial learning, as responsible for solving the task (Eichenbaum *et al.* 1990, Whishaw *et al.* 1995, Pearce *et al.* 1998, Redish 2001, Pouzet *et al.* 2002).

2.1 Simultaneous development of place and cue response

Devan and White (1999) applied an experimental procedure in which cue and place responses were developed simultaneously in the water maze (figure 1). Rats were trained over 9 days, interleaving trials in the visible (days 1, 2, 4, 5, 7 and 8) and hidden (days 3, 6 and 9) versions of the maze. The escape platform remained in the same location during the entire training phase. On the tenth day, a test (competition) trial was performed with a visible escape platform in a location different from the one used during training. This testing paradigm allows the dissociation of both types of response, as animals exhibiting cue response will swim directly toward the visible platform, whereas a place response behaviour will lead the animal toward the location where the platform was during the training phase. The experimental procedure was applied to rats that were either intact or had lesions in one of the following areas: (i) hippocampal formation (fornix/fimbria); (ii) dorsolateral striatum; or (iii) dorsomedial striatum. During training, intact animals decreased their escape latency (cf. time required to reach the escape platform) in both the visible and hidden trials. Lesioned animals also improved during training, except for animals with hippocampal lesions, which learnt the visible version of the task but were unable to reach the hidden platform in blocks 3, 6 and 9. In the competition

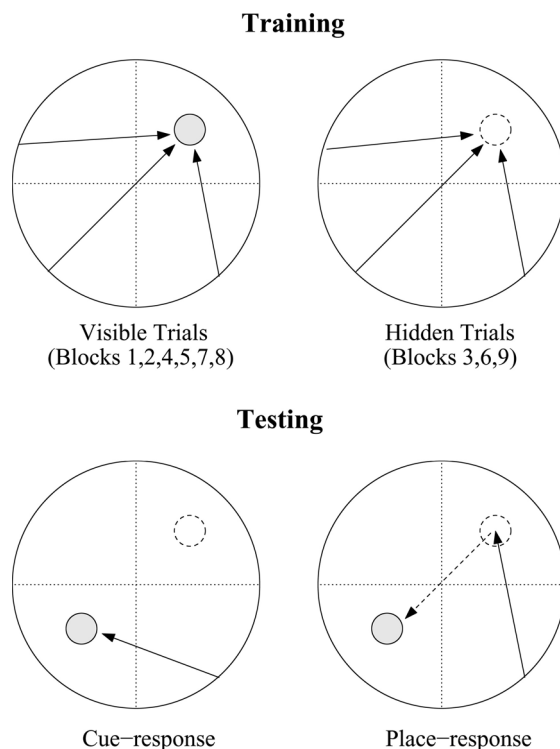


Figure 1. Water maze used by Devan and White (1999). Top: training phase. The rats were trained to reach an escape platform at a fixed location. During days 1, 2, 4, 5, 7 and 8 the platform was visible (left). In the remaining days (3, 6 and 9) the platform was below the surface of the water (right). The interleaved training in both the visible and hidden versions of the water maze allows the simultaneous development of place and cue responses. Bottom: on day 10, a competition trial is performed with a visible platform at a different location from the one used in training. The filled circle marks the new position of the platform. The dotted circle shows the location of the platform during training. Cue response will guide the animal toward the visible platform (left). In contrast, place response will lead the animal toward the location used during training (right). Animals who chose place response later changed to cue response and swam toward the visible platform (dotted arrow).

test, consistent with other studies, hippocampal and dorsomedial striatal lesions produced a preference for cue response, as opposed to lesions in the dorsolateral striatum, which biased the animal to exhibit place responses. Among the control animals, nearly half of them swam first to the location of the platform used during training (place response) and then toward the visible platform (cue response). This supports the idea of both types of response developing simultaneously in intact animals, and shows the ability to switch from a first behaviour to a second one if the first fails.

Other experiments have shown that, in some experimental conditions, rats change gradually from using a place strategy in early phases of training to a S-R behaviour, after longer training. Packard and McGaugh (1996) trained rats in a T-maze with a fixed starting position to go toward a baited arm (*e.g.* arm pointing to the west). Both the starting position and the bait location remained fixed during the training phase. A test trial was performed after a rotation of the maze by 180° to observe whether the rat followed a place response behaviour, going toward the same arm location (west), or if it repeated the same response as in training, performing the same body turn at the junction of the maze (entering the eastern arm). Figure 2 shows the experimental apparatus used in this experiment. After 8 days of training, most of the intact animals exhibited place response, but after 16 days most of them performed the same body turn as during training. Functional inactivation of the dorsolateral striatum at day 16 led the animals to exhibit place response, showing that this type of memory is preserved even if it is not behaviourally manifested.

2.2 Studying interactions between systems

The precise mechanism of the interactions between the different neural navigation systems is not yet clear. Studies have shown opposite effects of lesions in the hippocampus and the dorsolateral striatum when solving a place or cue task, respectively. The idea of a competition among those systems is supported by studies showing that lesions in one system may improve the performance in tasks associated to the unlesioned brain area (*i.e.* animals with hippocampal

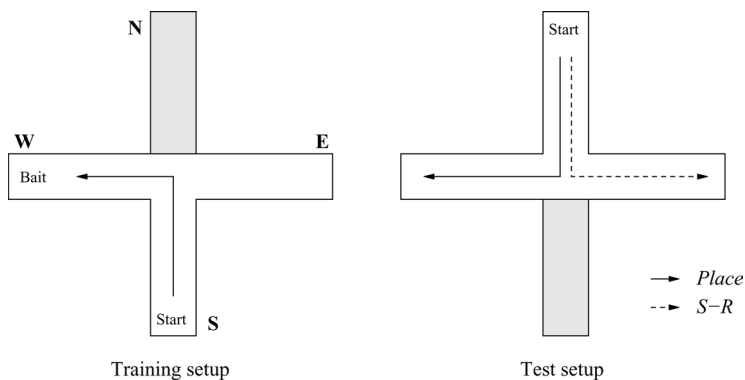


Figure 2. Plus-maze used by Packard and McGaugh (1996) and Chang and Gold (2003b). During training, access to the arm opposite the start location remains blocked (grey arm) to form a T-maze. Animals are trained to enter a consistently baited arm. Left: training set-up. Both the starting position (south) and the baited arm (west) remain fixed. Access to the north arm remains blocked during training. Right: testing set-up. The configuration is rotated by 180°, the starting point is changed to the north arm and the access to the south arm is now blocked. Animals expressing a place response enter the west arm, as they will go toward the same arm location (solid arrow). In contrast, animals eliciting S-R behaviour perform the same body turn at the intersection, entering the eastern arm (dotted arrow).

lesions outperform control animals in cue response tasks) (Packard *et al.* 1989, Devan and White 1999, Schroeder *et al.* 2002, Chang and Gold 2003a). This suggests that the intact area (eliciting appropriate behaviour) no longer has to compete with others for action selection, leading to a better performance (Gold 2004).

For instance, Chang and Gold (2003a) trained rats in a plus-maze with a varying starting point to go into a baited arm that could be either always at the same room location (place version), or defined by the same body turning, i.e. arm left of the starting point (S-R version). Hippocampal lesions, as expected, impair the learning of the place version of the task. Moreover, hippocampal lesions facilitate learning in the dorsostriatal-dependent S-R version of the maze. Similar facilitation effects have been reported in the water maze task described earlier (Devan and White 1999), where both hippocampal and dorsomedial striatal lesioned animals outperform control animals in the competition trial. Furthermore, ethanol administration, which produces memory impairments similar to those due to hippocampal lesions, also facilitates the acquisition of response tasks (Matthews *et al.* 1999).

The role of neurochemical factors mediating the competition between different systems has been studied recently (see Gold (2004) for a review). Injections of glutamate in either the hippocampus or striatum produce a dominance of spatial or S-R solution in a T-maze task (Packard 1999), respectively; similarly, the preference for spatial learning changes across the oestrous cycle of female rats such that animals are more likely to select place strategies when there are high levels of oestrogen (pro-oestrous) (Daniel and Lee 2004, Korol *et al.* 2004). A possible mechanism for oestrogen levels to affect the strategy selection is by modulation of neuromodulators such as acetylcholine (ACh) (Marriott and Korol 2003, Korol 2004).

Colombo and Gallagher (1998) reported that increases in acetylcholinesterase activity in the dorsolateral striatum are negatively correlated with the accuracy in hippocampal-dependent working memory tasks. Furthermore, alterations in the cholinergic septohippocampal system leading to enhanced or impaired cholinergic function in the hippocampus seem, respectively, to enhance or reduce learning and memory in this structure (Ragozzino and Gold 1995, Janis *et al.* 1998, Gold 2003) (but see Cahill and Baxter 2001, Bizon *et al.* 2003). Simultaneous measurement of acetylcholine in the hippocampus and the striatum showed a correlation between the changes of ACh release and the expression of place response or S-R behaviour (Chang and Gold 2003b, McIntyre *et al.* 2003). Interestingly, hippocampal ACh levels prior to the experiment reliably predicted which type of response would be preferred during training, suggesting a neurochemical bias toward a specific type of response depending on ACh levels. Changes in ACh in the dorsolateral striatum and the hippocampus were not negatively correlated, suggesting an indirect interaction, at least with respect to cholinergic activity, among the two systems (Gold 2004).

According to the studies described above, models of animal navigation should: (i) implement separately the two different types of response; and (ii) provide a competition mechanism that selects which kind of action should be performed. The outcome of the competition may depend on the characteristics of the task and the training process (Packard and McGaugh 1996, Oliveira *et al.* 1997, Chang and Gold 2003a), as well as endogenous factors such as stress, hormone status, or motivation (Kim and Baxter 2001, White and McDonald 2002, Marriott and Korol 2003). Additionally, as a consequence of the competition between place and cue response, the disruption of one of the modules should improve the performance in tasks involving the remaining, intact module.

Following these considerations, in the next section we present a model of rat navigation behaviour that is able to develop place and cue responses and selects the appropriate action to be performed.

3. Model of navigation

Inspired by the idea of multiple navigation systems acting in parallel and competing for action selection, we have proposed a top-down model of navigation in rats (Chavarriaga *et al.* 2005). The model will be reviewed in this section and a further comparison with the Devan and White (1999) experiment will be presented. A simulated robot was used to test the performance of the model.

The model is composed of two separate modules, implementing place response and cue response, respectively. Both modules use reinforcement learning to build an association between a state representation and a set of actions (Schultz *et al.* 1997, Sutton and Barto 1998). In the case of cue response, the state space codes for the location of a cue associated with the delivery of reward. In the case of place response, the state codes the location of the animal (which can be considered as a *cognitive map*). Both modules consist of two populations of rate-coded neurons: the first stores the state representation (place or presence of cues) and the second encodes the set of possible actions (direction of movements toward the goal location). Each module computes independently a direction of movement. A competition mechanism is used to select which action will actually be performed by the agent (figure 3).

In the following sections, we describe the encoding of cue and space representation (sections 1 and 2, respectively), as well as the action selection mechanism (section 3) and the learning procedure for the entire model (section 4).

3.1 Cue representation

As object recognition is out of the scope of this work, a simple representation of cues was chosen. A linear grey scale image was used to drive a population of sensory cells (SC). Each sensory cell i corresponded to a specific direction ϕ_i in the visual field and its activity corresponded to the normalized grey scale value in that direction, $r_i^{SC} = I(\phi_i)$.

In simulations, we used a dark cue on light background to mark the goal. Approaching behaviour could be achieved by turning in the direction of the cue and moving in that direction.

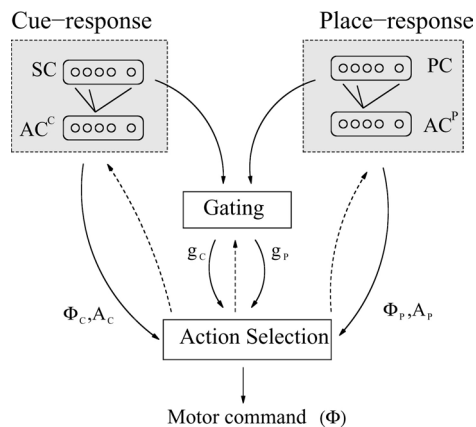


Figure 3. Block diagram of the model. Separate modules implement independently place and cue responses (Φ_P and Φ_C , respectively). A competition mechanism is used to select the action to be performed, as well as to modulate learning in both modules. SC \equiv Sensory cells forming the cue representation. PC \equiv Spatial representation stored in a population of place cells. AC \equiv Action cells coding for direction of movements. g_C , g_P \equiv Gating values for the cue and place response, respectively. A_C \equiv Action value of a movement in the direction Φ_C selected by the cue response module. A_P \equiv Action value of a movement in the direction Φ_P selected by the place response module.

3.2 Space representation

To build a representation of space, required for place response, we used a model proposed previously (Arleo and Gerstner 2000, Arleo *et al.* 2004), which consists of two separate pathways processing idiothetic and allothetic information that converge onto a population of place-sensitive cells (place cells, PC). Figure 4 shows the structure of the model.

The idiothetic pathway consists of a population of model neurons (PI cells, for path integration) with Gaussian receptive fields and preconfigured metric relations. As the agent moves, the estimated position is updated by integrating self-motion signals, which can be provided by proprioception and vestibular signals in animals or by odometry in robots.

The allothetic pathway is composed of two populations. The first population (view cells, VC) stores the features of previously perceived local views and compares it with the current visual input. In this version of the model, a simplified version of the external input to the hippocampus is used, as local views are determined by the distances to walls around the agent (Burgess *et al.* 2000, Chavarriaga *et al.* 2005). The VC population projects onto the second population (allothetic place cells, APC), where a vision-based representation of space is built (Arleo *et al.* 2004).

Both allothetic and idiothetic pathways converge on to the hippocampal population of PC. When exploring a new environment, a new PC is recruited at each new location and connected to simultaneously active APC and PI cells. Hebbian learning is applied to those connections

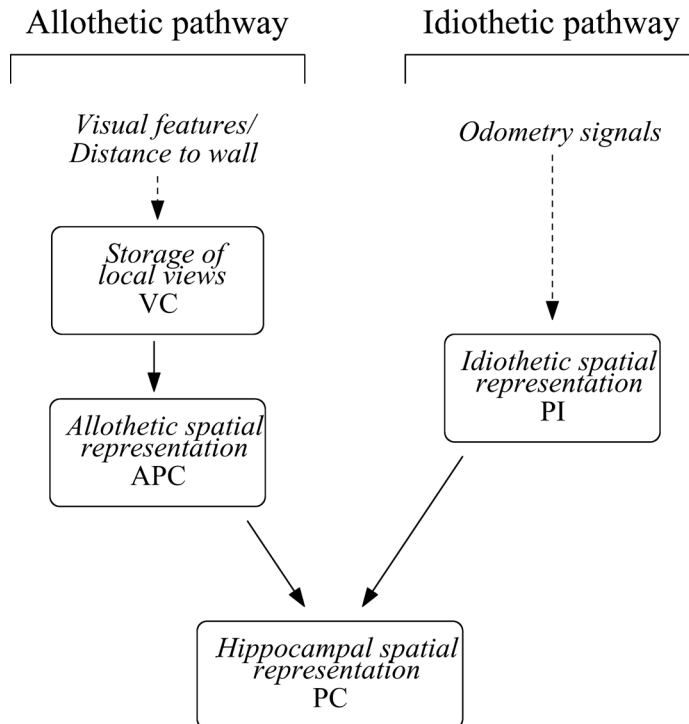


Figure 4. Functional diagram of the hippocampal model. Dashed lines denote algorithmic transformation of the sensory input, solid lines denote projections between populations of model neurons. The population of hippocampal place cells (PC) yields a representation of space used in the place response model of figure 3. PI denotes a representation of space based on path integration using proprioceptive odometry signals. APC denotes a spatial representation by allothetic place cells, driven by view cells (VC) encoding the local view. PI and APC information is combined in the final PC representation.

in order to combine external and internal information into one hippocampal representation of space.

After learning, the activity r^{PC} of a hippocampal PC correlates with the location of the agent. The population activity of PC forms a spatial representation consisting of a set of overlapping receptive fields covering the environment, which can be used to estimate the location of the agent. Figure 5(a) shows the receptive field of a typical place cell and figure 5(b) the activity of the PC population (~ 2000 PC cells) in a square environment. The cross signals the location of the agent; each dot corresponds to one cell and is located at the centre of its receptive field. Dark dots correspond to highly activated cells.

3.3 Action selection

Based on the representation of cue (section 1) or space (section 2) information, each module computes a possible action to be taken, corresponding to a cue or place response behaviour, respectively.

Actions are encoded in a population of N_{AC} action cells (AC). Each cell i codes for a specific direction of movement $\phi_i = 2\pi i / N_{\text{AC}}$, and its activity corresponds to the action value or Q -value (predicted future reward) in the sense of reinforcement learning [Sutton and Barto 1998].

Each module (cue or place response) has its own AC population driven by a presynaptic population described in earlier sections. The activity of an action cell i in the module $k \in [C, P]$ (for cue and place response, respectively) is computed:

$$a_i^k = \sum_j w_{ij}^k r_j^{\text{pre}}, \quad (1)$$

where r_j^{pre} corresponds to the firing rate of presynaptic cells coding for spatial representation ($k = P$) or location of a cue ($k = C$). Action cells in the module implementing place response (a_i^{P}) will code for the expected reward when moving toward a specific location in the environment, whereas the cue response module (a_i^{C}) encodes the reward expected when approaching a specific cue in the visual field.

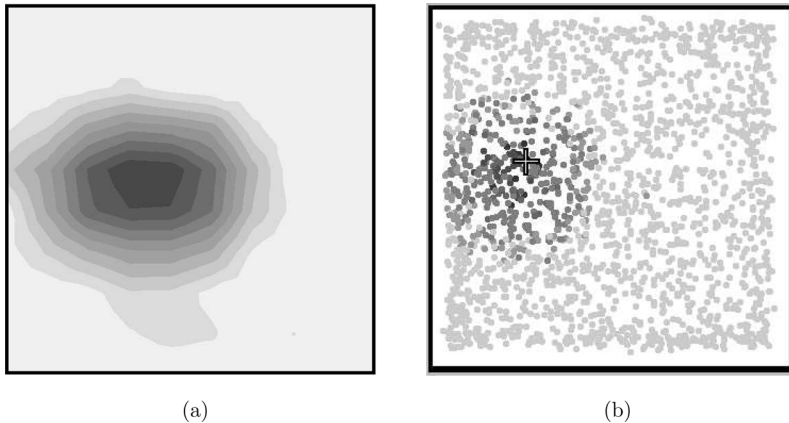


Figure 5. (a) Receptive field of a typical place cell. The dark region indicates the area where the place cell is maximally active (place field). (b) Population of hippocampal place cells after exploration of a square environment (80×80 cm). Each dot corresponds to one cell and is located at the centre of its receptive field. Dark dots correspond to highly activated cells. The location of the agent is marked by the white cross.

For each type of response the population vector Φ_k corresponds to the continuous direction of movement which predicts the maximum reward

$$\Phi_k = \arctan \left(\frac{\sum_i a_i^k \sin(\phi)}{\sum_i a_i^k \cos(\phi)} \right). \quad (2)$$

The agent must select among the directions computed by each module (Φ_C for cue response, or Φ_P for place response). To do so, we implement the selection of action by means of a competition among the modules. The probability of moving in the direction Φ_k suggested by a module $k \in [C, P]$ depends on a gating value g_k and the action value A_k of the proposed direction. Specifically,

$$P(\Phi = \Phi_k) = \frac{g_k A_k}{(g_P A_P) + (g_C A_C)}. \quad (3)$$

As AC provide the action values of a set of discrete directions ϕ_i , the action value A_k ($A_k \in [A_P, A_C]$) for the continuous direction Φ_k is computed by linear interpolation of the action values (equation (1)) of the two nearest discrete directions.

Gating values g_k (g_P and g_C , for place and cue response, respectively) depend on the hippocampal activity (PC) as well as the sensory input coding for the presence of cues (SC).

$$g_k = \sum_{j \in \text{PC}} (z_{kj}^{\text{PC}} r_j^{\text{PC}}) + \sum_{j \in \text{SC}} (z_{kj}^{\text{SC}} r_j^{\text{SC}}), \quad (4)$$

where z_{kj}^{PC} and z_{kj}^{SC} are the weights of a connection from a presynaptic cell j to a gating unit k .

The system defined by equations (1)–(4) will select among place and cue response according to the gating values g_k and the reward predicted by each module (Chavarriaga *et al.* 2005).

At this point, at each time step both modules propose a direction of movement (Φ_C , Φ_P) according to its perceptual input (equation (2)). Gating values are computed (equation (4)) and used to select the direction of the next movement. To allow a chosen response to be applied during several time steps, competition among the modules (equation (3)) is performed only when the accumulated prediction error (since the moment of the last competition) reaches a predefined threshold.

3.4 Learning

The parameters of the model should be updated in such a way that both modules can improve their performance in the current task, and the module that better predicts the reward should increase its probability of being selected (i.e. increase its gating value). In the TD(λ) algorithm (Sutton and Barto 1998) weights are updated proportional to the reward prediction error δ_k ,

$$\delta_k(t) = R_t - A^k(t-1) + \gamma A^k(t), \quad (5)$$

where $A^k(t)$ corresponds to the action value of the action taken at time t and $R(t)$ corresponds to the immediate reward received at time t . In our model the signal δ_k is scaled by a factor

h_k depending on both the gating value and the prediction error for the module k (Baldassarre 2002):

$$h_k = \frac{g_k c_k}{\sum_i (g_i c_i)}, \quad (6)$$

where $c_k = \exp(-\rho \delta_k^2)$ ($\rho > 0$). For module k weights will be updated according to

$$\Delta w_{ij}^k = \eta^k \delta_k h_k e_{ij}^k. \quad (7)$$

Here η^k is the learning rate for expert k and e_{ij}^k is the eligibility trace (Sutton and Barto 1998). The scaling factor h_k ensures that weight modification is most significant in the modules with high gating values (more likely to be selected) and best reward prediction ($c_k \rightarrow 1$). It should be noted that a module with a small gating value or a big reward prediction error will not be updated significantly. This allows the agent to preserve previously learnt responses (memories) which, if not useful in the present task, might be required in the future. In contrast, in situations with no conflict (i.e. when both tasks are equally good solving the task), both modules will tend to have small prediction errors, leading to comparable gating values for the two strategies, therefore both modules will be updated similarly.

The gating network weights are updated such that $g_k \rightarrow h_k$,

$$\Delta z_{kj}^{\text{PC,SC}} = \xi (h_k - g_k) r_j^{\text{PC,SC}} \quad (8)$$

with a learning rate ξ . This will assign higher gating values to the modules that consistently have small prediction error (Baldassarre 2002, Chavarriga *et al.* 2005).

4. Results

We tested the model following the experimental paradigm described by Devan and White (1999). This paradigm allows the simultaneous development of both place and cue responses in a navigation task. After training, the selection mechanism is probed in a conflict situation (competition trial) testing the ability to choose the most appropriate strategy to solve the task. Moreover, in some cases a change in the strategy may take place during the task when the initially selected strategy fails (i.e. such as animals switching from place to cue response in the Devan and White (1999) experiment).

The experiment consists of a combined cue–place task in the water maze. The training phase consists of nine training blocks, with four trials per block. The visible version of the maze is used in blocks 1, 2, 4, 5, 7 and 8, and the hidden version in the remaining blocks (3, 6 and 9). In our simulations, a reward signal was delivered when the agent reached the location of the platform, which remains constant during the entire training phase.

On day 10, a test (competition) trial is performed with a cued platform at a different location than during training. This paradigm dissociates the use of place and cue response: if the agent selects a place response behaviour, it will go toward the location of the platform in the training phase, whereas cue information will lead it toward the location of the cued platform.

In the simulations, at the end of the training phase (end of the ninth training block) both modules have simultaneously learnt responses that guide the agent toward the escape platform (figure 6 top). The difference between the place and cue response developed in each module can be observed in the configuration of the competition trial once the visible platform is moved to a different place (figure 6 bottom). The cue response module elicits responses guiding the agent toward the visible cue, whereas the place response module still points toward the location of the platform learned during training.

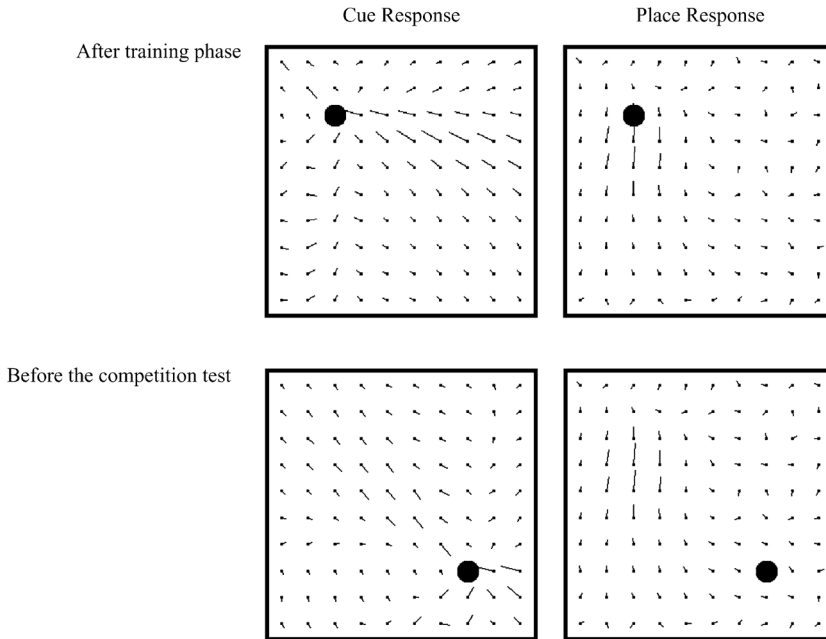


Figure 6. Navigation map after nine blocks of training in the combined cue–place task (Devan and White 1999). The navigation map shows, for every sampled position in the environment (dots), the direction of movement computed for each module. The length of each line is proportional to the action value (A_k) associated with that direction. The closed circle marks the location of the goal. Left column: cue response module. Right column: place response module. Top: platform at the learned location. Bottom: platform shifted to a new location.

In the competition trial, the trajectory followed by the agent depends on the outcome of the competition among the cue and place modules. As in the experiment with animals, cue response was chosen in some cases right from the beginning, leading to straight trajectories toward the escape platform. In other cases, the agent first exhibits place response and then switches to a cue response. Trajectories followed by the agent in both cases are shown in figure 7.

The escape latency (time steps taken by the agent to reach the goal) may be used to measure the learning of a navigational task. Across the nine sessions of training, the animal (or robot) reduced the time to reach the goal. Devan and White (1999) reported that latencies of intact animals decreased in both the visible (blocks 1, 2, 4, 5, 7 and 8) and hidden trials (blocks 3, 6 and 9). The performance in the visible trials improves faster than in the hidden trials. Figure 8 shows the averaged escape latency over several simulations of our model. As observed in animals, both cue and place responses are acquired gradually during training. Also, the short latency in the competition trial (block 10) is evidence of the ability of the model to overcome the conflict situation and reach the goal.

A facilitation effect has been proposed as a consequence of the competition among different memory systems. That is, lesions in one system improve the performance in a task associated with the unlesioned system. It has been observed that hippocampal lesions improve the performance in S-R tasks (presumably by releasing the striatum from competition) (Packard *et al.* 1989, Packard and McGaugh 1996, Devan and White 1999). According to this, an agent following exclusively a taxon strategy should have smaller latencies in the test block than an agent controlled by the competition mechanism. In figure 9 we compare the performance in the competition trial (as shown in figure 8) with the performance obtained if a cue response is

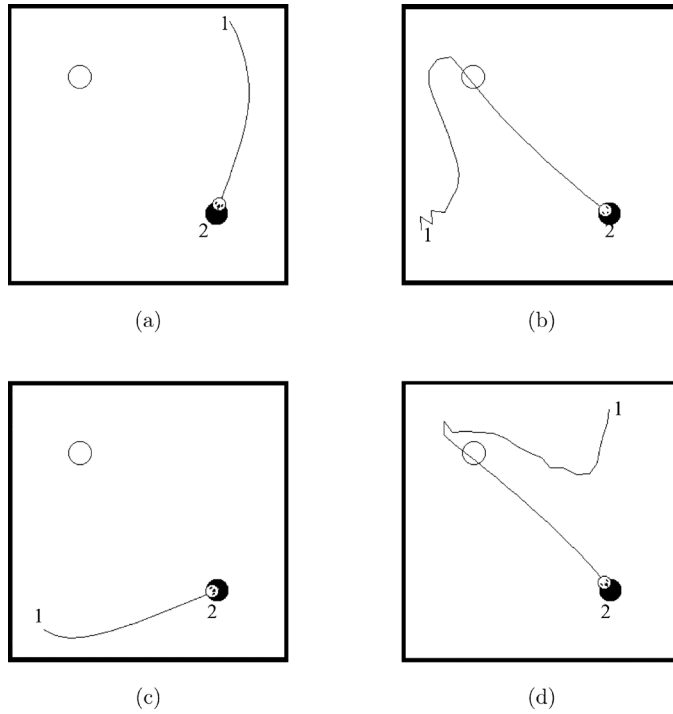


Figure 7. Examples of trajectories observed during the competition trial. In some cases the agent goes directly toward the visible goal (a) and (c) by applying cue response behaviour. In (b) and (d), a place-based strategy is selected leading the agent to the place where the platform was located during training (open circle). A cue response is then adopted allowing the animal to reach the goal (closed circle). (1) Starting position; (2) final position of the agent.

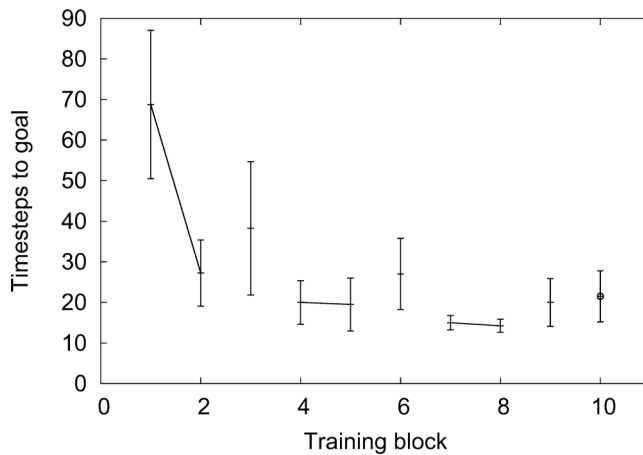


Figure 8. Escape latency (time steps taken by the agent to reach the platform). Each training block consists of four trials, and each trial starts with the agent located at a random position (selected such that it takes more than ten steps to reach the goal). *Disconnected* points correspond to the trials in the hidden water maze (blocks 3, 6 and 9) and the competition test (block 10).

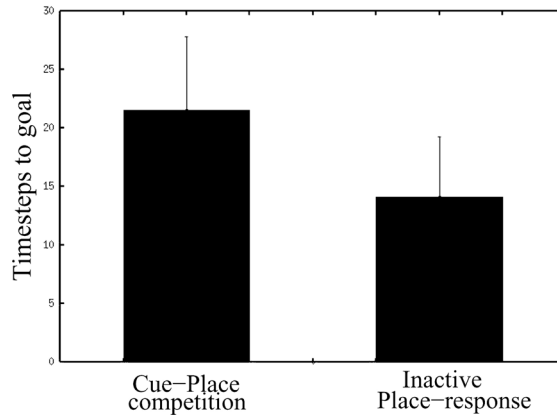


Figure 9. Deactivation of the place response module improves the performance in the competition trial. When both cue and place response modules compete for action selection, the agent takes, on average, more time to reach the visible goal than when only the cue response module is functional.

always selected. As expected, the deactivation of the place response results in shorter latencies in the competition test (a S-R task).

5. Discussion

A substantial body of experimental results reviewed in section 2 supports the theory of multiple systems acting in parallel and mediating different forms of learning. The hippocampal formation on one side and the dorsolateral striatum on the other side have been dissociated as responsible for spatial learning (place response) and S-R behaviour (cue response), respectively. Different interactions among those systems have been observed as they are able simultaneously to learn a task. This is consistent with the concept of latent learning, and it has been suggested as the basis for habit formation (Packard and McGaugh 1996).

We have presented a system-level computational model for navigation based on the competition among different modules implementing independently place and cue responses. Consistent with experimental data, the modules are able to develop simultaneously a proper strategy to solve the task (figures 6 and 8). Moreover, the model develops a pure cue response if the location of a cued platform varies from one trial to the next and a pure place response if cues vary while the location of the hidden platform remains stable (Chavarriaga *et al.* 2005).

As a result of the competition, the agent is able to change from one type of response to another when it is no longer suitable for solving the task (figure 7). Additionally, experiments have shown that the performance in a given task can be improved by lesioning unrelated competing neural systems (Packard *et al.* 1989, Devan and White 1999, Chang and Gold 2003a). In our model, deactivation of the place response module yielded smaller escape latencies in the competition test (figure 9).

Both modules apply reinforcement learning to learn associations between state (place or cue) representations and actions (Sutton and Barto 1998). The update of the parameters of the system is modulated by the gating system (equation (7)) such that modules with better performance (small prediction error, δ) and high gating values (g_k) will be more strongly modified.

How different systems in the rat brain interact in the competition for action selection is still an open question. One possibility for the competition is that a brain structure, outside the

striatum and hippocampus, receives the information processed by those systems and controls the expression of learned responses (Gold 2004, Mizumori *et al.* 2004). Also, an indirect way of communication between systems via the cerebral cortex has been proposed by White (2004). He suggested that output coming from one of the systems could change the representation of information stored in the cortex that will, in turn, be sent to all other competing systems. Currently, experimental support for these hypotheses is still missing.

In the water maze the medial region of the dorsal striatum seems to be involved in the competition process (Whishaw *et al.* 1987, Devan and White 1999) as lesions in this area did not prevent spatial learning, but produced a preference for cue response when cue information was available. This suggests a form of interaction via the dorsomedial striatum, in which hippocampal-mediated behaviour can override cue responses. The role of the basal ganglia in action selection has been modelled as a competition among different sensory motor programmes (implemented by parallel cortico-striatal loops) (Gurney *et al.* 2001). This competition may involve place responses based on spatial information from the hippocampus transferred via the ventral and dorsomedial striatum as a possible mechanism of interaction between place and cue responses.

Our model selects appropriately between place and cue responses depending on the performance of those strategies during training. According to this, if both strategies successfully solve the task they will have equal probability of being selected. However, animals gradually shift from place to cue response after extensive training (Packard and McGaugh 1996, Chang and Gold 2003b). At this stage, despite the preference for cue responses, the hippocampal structure seems to keep engaged in the task as it is indicated by the high levels of ACh release in the hippocampus (Chang and Gold 2003b) and the fact that place response can be restated by reversible striatal lesions (Packard and McGaugh 1996). The assumption that the strategy selection depends exclusively on the expected performance of each system implies that, after extensive training, the striatal system will consistently yield smaller errors than the hippocampal system, which is unlikely. An alternative hypothesis is that animals are biased toward cue responses when both responses are equally good at solving the task. This bias can be due to the fact that the dorsolateral striatum can take advantage of direct connections to motor and somatosensory cortices, while the hippocampus does not have such connections. In consequence, after extensive training, when both responses accurately solve the task, rapidly elicited stimulus-responses will overcome hippocampal-dependent place responses, producing the shift from place to cue responses.

The extension of the model to test this hypothesis, as well as tests in further experimental conditions and the implementation on a real robot, are the next steps in our modelling efforts. The model can also be extended to allow several S-R associations to compete with place responses, and to study possible mechanisms of competition between multiple modules.

Acknowledgement

This work was supported by the Swiss National Science Foundation under grant no. 200020-100265/1.

References

- A. Arleo and W. Gerstner, "Spatial cognition and neuro-mimetic navigation: a model of hippocampal place cell activity", *Biol. Cybernet.*, 83, pp. 287–299, 2000.
- A. Arleo, F. Smeraldi and W. Gerstner, "Cognitive navigation based on nonuniform Gabor space sampling, unsupervised growing networks, and reinforcement learning", *IEEE Trans. Neural Net.*, 15, pp. 639–652, 2004.
- G. Baldassarre, "A modular neural-network model of the basal ganglia's role in learning and selecting motor behaviours", *Cognitive Syst. Res.*, 3, pp. 5–13, 2002.

- J.L. Bizon, J.S. Han, C. Hudon and M. Gallagher, "Effects of hippocampal cholinergic deafferentation on learning strategy selection in a visible platform version of the water maze", *Hippocampus*, 13, pp. 676–684, 2003.
- M.A. Brown and P.E. Sharp, "Simulation of spatial-learning in the Morris water maze by a neural network model of the hippocampal-formation and nucleus accumbens", *Hippocampus*, 5, pp. 171–188, 1995.
- N. Burgess, A. Jackson, T. Hartley and J. O'Keefe, "Predictions derived from modelling the hippocampal role in navigation", *Biol. Cybernet.*, 83, pp. 301–312, 2000.
- N. Burgess, M. Recce and J. O'Keefe, "A model of hippocampal function", *Neural Net.*, 7, pp. 1065–1081, 1994.
- J.F. Cahill and M.G. Baxter, "Cholinergic and noncholinergic septal neurons modulate strategy selection in spatial learning", *Eur. J. Neurosci.*, 14, pp. 1856–1864, 2001.
- Q. Chang and P.E. Gold, "Intra-hippocampal lidocaine injections impair acquisition of a place task and facilitate acquisition of a response task in rats", *Behav. Brain Res.*, 144, pp. 19–24, 2003a.
- Q. Chang and P.E. Gold, "Switching memory systems during learning: changes in patterns of brain acetylcholine release in the hippocampus and striatum in rats", *J. Neurosci.*, 23, pp. 3001–3005, 2003b.
- G. Chevalier and J.M. Deniau, "Disinhibition as a basic process in the expression of striatal functions", *Trends Neurosci.*, 13, pp. 277–280, 1990.
- P.J. Colombo and M. Gallagher, "Individual differences in spatial memory and striatal chat activity among young and aged rats", *Neurobiol. Learn. Mem.*, 70, pp. 314–327, 1998.
- C. Da Cunha, S. Wietzikoski, E.C. Wietzikoski, E. Miyoshi, M.M. Ferro, J.A. Anselmo-Franci and N.S. Canteras, "Evidence for the substantia nigra pars compacta as an essential component of a memory system independent of the hippocampal memory system", *Neurobiol. Learn. Mem.*, 79, pp. 236–242, 2003.
- J.M. Daniel and C.D. Lee, "Estrogen replacement in ovariectomized rats affects strategy selection in the Morris water maze", *Neurobiol. Learn. Mem.*, 82, pp. 142–149, 2004.
- B.D. Devan and N.M. White, "Parallel information processing in the dorsal striatum: relation to hippocampal function", *J. Neurosci.*, 19, pp. 2789–2798, 1999.
- K. Doya, "Metalearning and neuromodulation", *Neural Net.*, 15, pp. 495–506, 2002.
- H. Eichenbaum, C. Stewart and R.G. Morris, "Hippocampal representation in place learning", *J. Neurosci.*, 10, pp. 3531–3542, 1990.
- D.J. Foster, R.G.M. Morris and P. Dayan, "A model of hippocampally dependent navigation, using the temporal difference learning rule", *Hippocampus*, 10, pp. 1–16, 2000.
- P. Gaussier, A. Revel, J.P. Banquet and V. Babeau, "From view cells and place cells to cognitive map learning: processing stages of the hippocampal system", *Biol. Cybernet.*, 86, pp. 15–28, 2002.
- W. Gerstner and L.F. Abbott, "Learning navigational maps through potentiation and modulation of hippocampal place cells", *J. Comput. Neurosci.*, 4, pp. 79–94, 1996.
- P.E. Gold, "Acetylcholine modulation of neural systems involved in learning and memory", *Neurobiol. Learn. Mem.*, 80, pp. 194–210, 2003.
- P.E. Gold, "Coordination of multiple memory systems", *Neurobiol. Learn. Mem.*, 82, pp. 230–242, 2004.
- K. Gurney, T.J. Prescott and P. Redgrave, "A computational model of action selection in the basal ganglia. I. A new functional anatomy", *Biol. Cybernet.*, 84, pp. 401–410, 2001.
- M.E. Hasselmo, J. Hay, M. Ilyn and A. Gorchetnikov, "Neuromodulation, theta rhythm and rat spatial navigation", *Neural Net.*, 15, pp. 689–707, 2002.
- L.S. Janis, M.M. Glasier, Z. Fulop and D.G. Stein, "Intraseptal injections of 192 igg saporin produce deficits for strategy selection in spatial-memory tasks", *Behav. Brain Res.*, 90, pp. 23–34, 1998.
- M.S. Jog, Y. Kubota, C.I. Connolly, V. Hillegaart and A.M. Graybiel, "Building neural representations of habits", *Science*, 286, pp. 1745–1749, 1999.
- J.J. Kim and M.G. Baxter, "Multiple brain-memory systems: the whole does not equal the sum of its parts", *Trends Neurosci.*, 24, pp. 324–330, 2001.
- D.L. Korol, "Role of estrogen in balancing contributions from multiple memory systems", *Neurobiol. Learn. Mem.*, 82, pp. 309–323, 2004.
- D.L. Korol, E.L. Malin, K.A. Borden, R.A. Busby and J. Couper-Leo, "Shifts in preferred learning strategy across the estrous cycle in female rats", *Horm. Behav.*, 45, pp. 330–338, 2004.
- L.K. Marriott and D.L. Korol, "Short-term estrogen treatment in ovariectomized rats augments hippocampal acetylcholine release during place learning", *Neurobiol. Learn. Mem.*, 80, pp. 315–322, 2003.
- D.B. Matthews, M. Ilgen, A.M. White and P.J. Best, "Acute ethanol administration impairs spatial performance while facilitating nonspatial performance in rats", *Neurobiol. Learn. Mem.*, 72, pp. 169–179, 1999.
- C.K. McIntyre, L.K. Marriott and P.E. Gold, "Patterns of brain acetylcholine release predict individual differences in preferred learning strategies in rats", *Neurobiol. Learn. Mem.*, 79, pp. 177–183, 2003.
- J.W. Mink, "The basal ganglia: focused selection and inhibition of competing motor programs", *Prog. Neurobiol.*, 50, pp. 381–425, 1996.
- S.J. Mizumori, O. Yeshenko, K.M. Gill and D.M. Davis, "Parallel processing across neural systems: implications for a multiple memory system hypothesis", *Neurobiol. Learn. Mem.*, 82, pp. 278–298, 2004.
- R.G.M. Morris, "Spatial localization does not require the presence of local cues", *Learn. Motiv.*, 12, pp. 239–260, 1981.
- J. O'Keefe and L. Nadel, *The Hippocampus as a Cognitive Map*, Oxford: Clarendon Press, 1978.
- M.G. Oliveira, O.F. Bueno, A.C. Pomarico and E.B. Gugliano, "Strategies used by hippocampal- and caudate-putamen-lesioned rats in a learning task", *Neurobiol. Learn. Mem.*, 68, pp. 32–41, 1997.

- M.G. Packard, "Glutamate infused posttraining into the hippocampus or caudate-putamen differentially strengthens place and response learning", *Proc. Natl. Acad. Sci. USA*, 96, pp. 12881–12886, 1999.
- M.G. Packard, R. Hirsh and N.M. White, "Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems", *J. Neurosci.*, 9, pp. 1465–1472, 1989.
- M.G. Packard and B.J. Knowlton, "Learning and memory functions of the basal ganglia", *Ann. Rev. Neurosci.*, 25, pp. 563–593, 2002.
- M.G. Packard and J.L. McGaugh, "Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning", *Neurobiol. Learn. Mem.*, 65, pp. 65–72, 1996.
- J.M. Pearce, A.D.L. Roberts and M. Good, "Hippocampal lesions disrupt navigation based on cognitive maps but not heading vectors", *Nature*, 396, pp. 75–77, 1998.
- B. Pouzet, W.N. Zhang, J. Feldon and J.N. Rawlins, "Hippocampal lesioned rats are able to learn a spatial position using non-spatial strategies", *Behav. Brain Res.*, 133, pp. 279–291, 2002.
- K.E. Ragozzino, S. Leutgeb and S.J. Mizumori, "Dorsal striatal head direction and hippocampal place representations during spatial navigation", *Exp. Brain Res.*, 139, pp. 372–376, 2001.
- M.E. Ragozzino and P.E. Gold, "Glucose injections into the medial septum reverse the effects of intraseptal morphine infusions on hippocampal acetylcholine output and memory", *Neuroscience*, 68, pp. 981–988, 1995.
- M. Recce and K.D. Harris, "Memory for places: a navigational model in support of Marr's theory of hippocampal function", *Hippocampus*, 6, pp. 735–748, 1996.
- A. Redish, *Beyond the Cognitive Map, From Place Cells to Episodic Memory*, London: MIT Press-Bradford Books, 1999.
- A.D. Redish, "The hippocampal debate: Are we asking the right questions?", *Behav. Brain Res.*, 127, pp. 81–98, 2001.
- A.D. Redish and D.S. Touretzky, "The role of the hippocampus in solving the Morris water maze", *Neural Comput.*, 10, pp. 73–111, 1998.
- J.P. Schroeder, J.C. Wingard and M.G. Packard, "Post-training reversible inactivation of hippocampus reveals interference between memory systems", *Hippocampus*, 12, pp. 280–284, 2002.
- W. Schultz, "Getting formal with dopamine and reward", *Neuron*, 36, pp. 241–263, 2002.
- W. Schultz, P. Dayan and P.R. Montague, "A neural substrate of prediction and reward", *Science*, 275, pp. 1593–1599, 1997.
- R. Sutton and A.G. Barto, *Reinforcement Learning – An Introduction*, Cambridge, MA: MIT Press, 1998.
- E.C. Tolman, "Cognitive maps in rats and men", *Psychol. Rev.*, 55, pp. 189–208, 1948.
- P. Voorn, L.J. Vanderschuren, H.J. Groenewegen, T.W. Robbins and C.M. Pennartz, "Putting a spin on the dorsal-ventral divide of the striatum", *Trends Neurosci.*, 27, pp. 468–474, 2004.
- I.Q. Whishaw, J.C. Cassel and L.E. Jarrad, "Rats with fimbria-fornix lesions display a place response in a swimming pool: a dissociation between getting there and knowing where", *J. Neurosci.*, 15, pp. 5779–5788, 1995.
- I.Q. Whishaw, G. Mittleman, S.T. Bunch and S.B. Dunnett, "Impairments in the acquisition, retention and selection of spatial navigation strategies after medial caudate-putamen lesions in rats", *Behav. Brain Res.*, 24, pp. 125–138, 1987.
- N.M. White, "The role of stimulus ambiguity and movement in spatial navigation: a multiple memory systems analysis of location discrimination", *Neurobiol. Learn. Mem.*, 82, pp. 216–229, 2004.
- N.M. White and R.J. McDonald, "Multiple parallel memory systems in the brain of the rat", *Neurobiol. Learn. Mem.*, 77, pp. 125–184, 2002.

