Biol. Cybern. 67, 535–544 (1992)

Biological Cybernetics

Modeling the dishabituation hierarchy: The role of the primordial hippocampus*

DeLiang Wang and Michael A. Arbib

Center for Neural Engineering, University of Southern California, Los Angeles, CA 90089-2520, USA

Received January 30, 1992/accepted in revised form April 6, 1992

Abstract. We present a neural model for the organization and neural dynamics of the medial pallium, the toad's homolog of mammalian hippocampus. A neural mechanism, called cumulative shrinking, is proposed for mapping temporal responses from the anterior thalamus into a form of population coding referenced by spatial positions. Synaptic plasticity is modeled as an interaction of two dynamic processes which simulates acquisition and both short-term and long-term forgetting. The structure of the medial pallium model plus the plasticity model allows us to provide an account of the neural mechanisms of habituation and dishabituation. Computer simulations demonstrate a remarkable match between the model performance and the original experimental data on which the dishabituation hierarchy was based. A set of model predictions is presented, concerning mechanisms of habituation and cellular organization of the medial pallium.

1 Biological basis

Habituation is a decrease in the strength of a behavioral response that occurs when an initially novel stimulus is presented repeatedly. It is probably the most elementary and ubiquitous form of plasticity, and its underlying mechanisms may well provide bases for understanding other forms of plasticity and more complex learning behaviors. Amphibia show stimulus specific habituation. After repeated presentation of the same prey dummy in their visual field, toads reduce the strength of orienting repsonses toward the moving stimulus. This visual habituation has the following characteristics: *(1) Locus specificity.* After habituation of an orienting response to a certain stimulus applied at a given location, the response can be released by the same stimulus applied at a different retinal locus (Eikmanns 1955; Ewert and Ingle 1971). *(2) Hierarchical stimulus specificity.* Another stimulus given at the same locus may restore the response habituated by a previous stimulus. Only certain stimuli can dishabituate a previously habituated response. Ewert and Kehl (1978) demonstrated that this dishabituation forms a hierarchy (Fig. 1A), where a stimulus can dishabituate the habituated responses of another stimulus if the latter is lower in the hierarchy.

A model developed by Wang and Arbib (1991a) showed how a group of cells in the anterior thalamus (AT) could respond to stimuli higher in the hierarchy with larger AT responses. The present paper studies where and how the response activities are stored, and investigates the learning mechanisms involved in habituation and dishabituation.

In addition to locus and hierarchical stimulus specificity, prey-catching behavior in toads exhibits the following typical properties of habituation (for a review see Ewert 1984).

- After the same prey dummy is presented at the same retinal locus repetitively, the response intensity decreases exponentially.
- Spontaneous recovery occurs after the stimulus is withheld. The time course of recovery from habituation exhibits two phases: a short-term process that lasts for a few minutes, and a long-term process which lasts for at least 6 h.
- Habituation is faster and lasts longer with an increase in the number of training series.
- The response to the dishabituating stimulus decreases with repeated presentation. The decrease of the response to a repeated dishabituating stimulus also follows a typical course of habituation.
- Dishabituation does not counteract the effects of previous habituation, but rather establishes a separate neuronal process (Wang and Ewert 1992).

Cells with adaptation properties can be found in different visual neural structures. However, in retinal,

* The research described in this paper was supported in part by grant no. 1RO1 NS 24926 from the National Institutes of Health (M.A.A., Principal Investigator)

Correspondence to: D. Wang, Department of Computer and Information Science, Ohio State University, 2036 Neil Avenue Mall, Columbus, OH 43210-1277, USA

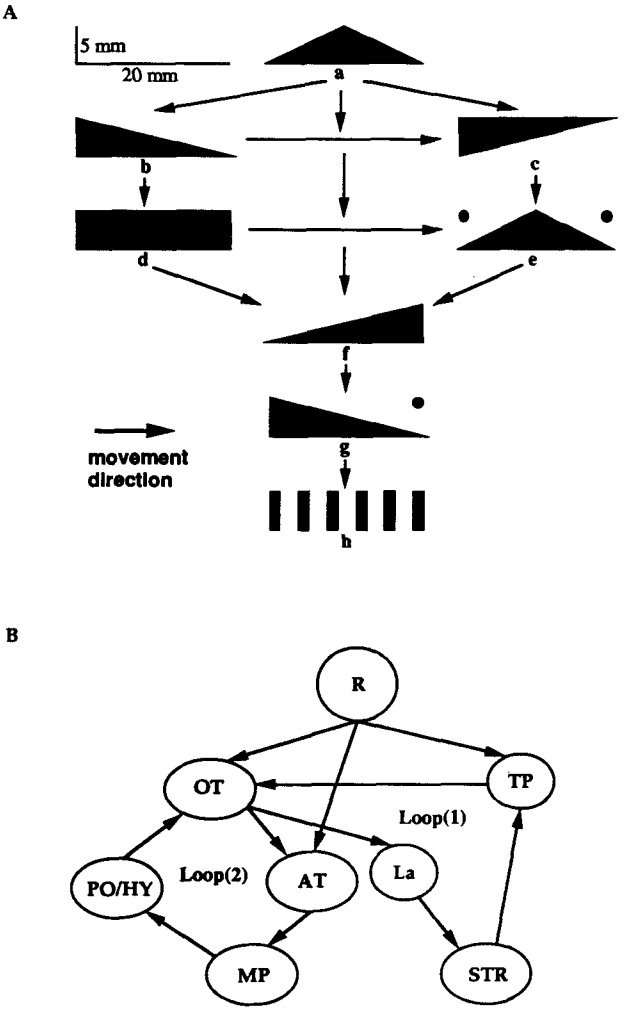

Fig. 1. **A** Dishabituation hierarchy for worm stimuli in stimulus-specific habituation. One stimulus can dishabituate all the stimuli below it. On the same level the left stimulus can slightly dishabituate the right one (redrawn with permission from Ewert and Kehl 1978). **B** Two neural loops mediating and modulating prey-catching behaviors in toads. Abbreviations are: R: retina; OT: optic tectum; TP: thalamic-pretectal area; PO: preoptic region; HY: hypothalamus; AT: anterior thalamus; La: lateral anterior thalamic nucleus; MP medial pallium; and STR: striatum

tectal or TP (thalamic-pretectal) neurons, neither neuronal adaptation nor facilitation effects lasted longer than 90 to 120 s. Since behavioral studies show that habituation effects may last 24 h or longer, habituation substrates must be located elsewhere. Anatomically, the medial pallium (MP) has been thought to be a homolog of the mammalian hippocampus ("primordium hippocampi" by Herrick 1933). Its location in the dorsomedial wall of the hemisphere is in a position similar to that of the mammalian hippocampal region. Finkenstädt (1989a) identified three types of visual sensitive neurons in MP which exhibit spontaneous firing activities. MP1 neurons strongly increase, MP3 neurons decrease, and MP2 neurons do not alter their discharge rates in response to 30 min of repetitive stimulation with a visual moving object traversing their excitatory receptive fields. After bilateral lesions of MP, toads showed no progressive decrease of the orienting activity to repetitive stimulation (Finkenstädt and Ewert 1988a). In addition, both effects of conditioning and associative learning ability in naive toads are abolished due to MP-lesions (Finkenstädt and Ewert 1988b; Finkenstädt 1989b).

Lara and Arbib (1985), in their computational model, postulate that habituation is coded in telencephalic neurons which receive input from the tectum and modulate prey-catching behavior through their projection to TP neurons, which inhibit pyramidal neurons in the tectum, thus modulating the toad's orienting response. Their modeling of habituation processes follows the general idea of the comparator model of stimulus-specific habituation (Sokolov 1960). However, direct projections from the tectum to the telencephalon lack anatomical support. Furthermore, their assumption of telencephalic modulation via TP is not consistent with the 2DG-uptake data that TP shows no change while the tectum shows a strong decrease after long-term habituation training (Finkenstädt and Ewert 1988a).

Ewert (1987) proposes the idea of neural loop interaction to explain amphibian prey-catching behavior and its modulation (Fig. 1B). Loop (1) involves the tectum, the lateral anterior thalamic nucleus (La), the striatum, and the TP area, and is supposed to mediate prey-catching behavior; the so-called loop (2) starts with the retina and the tectum which send axons to AT and from there to MP, which then projects descendingly to PO and from there to HYP which in turn sends efferents to the tectum (*retina → tectum → AT → MP → PO → HYP → tectum*). Loop (2) is supposed to modulate prey-catching behavior initiated in the tectum. All data obtained to date strongly suggest that MP is the neural structure where learning occurs. For stimulus specific habituation in anurans, MP modulates prey-catching behavior via the PO/HYP pathway, inhibiting tectal activities (Ewert 1987; Finkenstädt and Ewert 1988a). Having modeled AT in Wang and Arbib (1991a), we now turn to a model of MP which completes our explanation of hierarchical dishabituation.

2 A neural model of the medial pallium

Preliminary anatomical studies of the medial pallium indicate that this structure is organized in an orientation vertical to the telencephalic ventricle, with cells mainly projecting in this direction (Hoffman 1963; Kicliter and Ebbesson 1976). This leads us to suggest that the medial pallium processes information by means of functional units, in a form of vertical columns, consistent with the locus specificity of habituation in toad's prey-catching behavior. In what follows, we will model only one such functional unit to demonstrate most of the properties of hierarchical stimulus-specific habituation.

For stimulus-specific habituation, the medial pallium receives inputs from the anterior thalamus, a cell type of which has been modeled in Wang and Arbib

(1991a). In the model, an MP column receives input from an AT neuron, representing a specific visual location. The MP column model includes arrays of five types of neurons, MP1, MP2, and MP3 (Finkenstädt 1989a) and hypothetical cell types P1 and P2 (Fig. 2). P1 receives projections from AT, and projects to MP2. MP2 projects to P2 and MP3. MP3 inhibits MP1, which further inhibits P2. The final outcome of the functional column is reflected by a hypothetical neuron OUT that integrates activities from the P2 layer. The connections from MP2 to P2 and MP3 are the only plastic connections, that is, their efficacies are modifiable by training, which underlies stimulus-specific habituation. For modeling simplicity, we assume that during presentation of a given stimulus, a stable activity of average firing rate from AT is continuously fed to the P1 layer.

P1 layer

Grossberg and Kuperstein (1986) proposed a network, called the position-threshold-slope map, the converts different input intensities to different positions in an array of neurons which have covarying weights and thresholds. Using an array of neurons with different spans for temporal summation, we are also able to convert firing trains of different frequencies into different positions in a subsequent array (Wang and King 1988; Wang and Arbib 1991b). The layers P1 and MP2 together transform different temporal activities into spatial activity distribution centered at different positions. Suppose that each layer in the column has n neurons, with $m_{p1}(i, t)$ the membrane potential of the ith P1 cell, $P1_i$,

$$\frac{dm_{p1}(i, t)}{dt} = -A_{p1}m_{p1}(i, t) + B_{p1}I(t) + \varrho \tag{1}$$

where A_{p1} is a relaxation parameter, and B_{p1} is a connection weight from the AT input, $I(t)$, which is assumed to be constant during a stimulus presentation. ϱ represents the amplitude of an uncorrelated white noise term introduced to the AT input. The noise is introduced to compensate to some extent for our use of a constant AT input. The output of cell $P1_i$, $N_{p1}(i, t)$, is formed by

$$N_{p1}(i, t) = \begin{cases} m_{p1}(i, t) & \text{if } m_{p1}(i, t) > \theta_i \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where θ_i is the threshold chosen to make cells with larger i harder to fire. In implementation, the choice of θ_i is a linear function of position i: $\theta_i = 46.5i/n + 7.75$ (any monotonic function of i would do the job).

MP2 layer

In order to map temporal activity into different spatial locations, we adapt Grossberg's (1976) idea that shunting inhibition fits normalization better than does subtractive inhibition. We propose that the membrane potential of the ith MP2 cell

$$\frac{dm_{mp2}(i, t)}{dt} = -A_{mp2}m_{mp2}(i, t) + (B_{mp2} - m_{mp2}(i, t))I_{ii}(t) + m_{mp2}(i, t)\sum_{j>i} I_{ij}(t) \tag{3}$$

with $0 \leqslant m_{mp2}(i, 0) \leqslant B_{mp2}$. A_{mp2} and B_{mp2} are parameters, and $I_{ij}(t)$ represents the input from neuron $P1_j$, equal to $W_{ij}N_{p1}(j, t)$ for $j \geqslant i$ with $W_{ii} = 1$ and $W_{ij} < 0$ $(j > i)$. At equilibrium,

$$m_{mp2}(i) = \frac{B_{mp2}I_{ii}}{A_{mp2} + I_{ii} - \sum_{j>i} I_{ij}} \tag{4}$$

With the unilateral shunting inhibition of (3) and the P1 layer, the network exhibits a phenomenon called *cumulative shrinking*, as described in the following. Let an arbitrary stimulus activate the neuron group $P1_1$, $P1_2, \ldots, P1_i$, that is, $N_{p1}(1, t) = N_{p1}(2, t) = \cdots = N_{p1}(i, t) > 0$. At equilibrium of the MP2 layer,

$$m_{mp2}(k) = \frac{B_{mp2}N_{p1}(i)}{A_{mp2} + N_{p1}(i)\left(1 - \sum_{j=k+1}^{i} W_{kj}\right)} \tag{5}$$

for $k = 1, \ldots, i$. If A_{mp2} is small compared to $N_{p1}(i)$, the output of neuron $P1_i$, $m_{mp2}(i) \approx B_{mp2}$ (Note that the summation in the denominator vanishes for $k = i$) no matter how large $N_{p1}(i)$ becomes, and the smaller is k

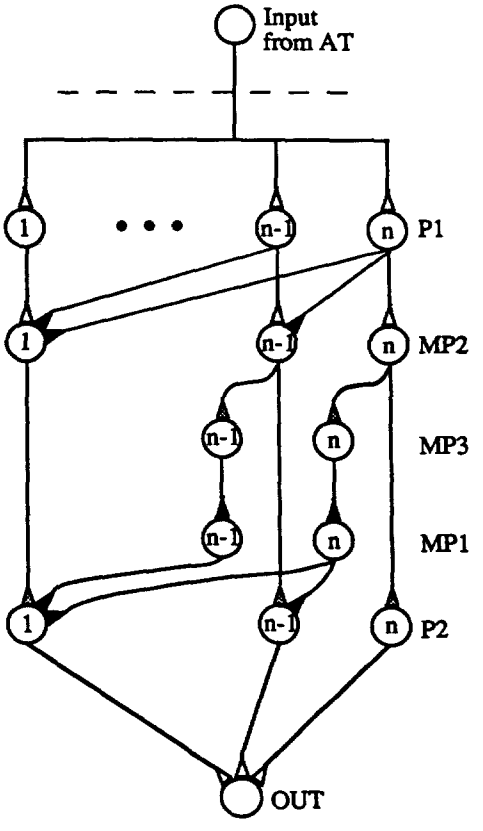

Fig. 2. Diagram of an MP column model. Each cell type is a layer of cells numbered from left to right as 1, 2, . . . , n. Synapses are indicated by triangles. Empty triangles indicate excitatory, black triangles inhibitory, and filled triangles habituation synapses

the larger is $1 - \sum_{j=k+1}^{i} W_{kj}$ (the W_{kj} are negative) and so the smaller is the value $m_{mp2}(k)$ due to cumulative inhibition. The cumulative shrinking normalizes and shrinks the activity in the MP2 layer along one direction, whereas the shunting inhibition as demonstrated by Grossberg (1976) only provides normalization. This mechanism also performs with non-equal values of inputs $N_{p1}(1), \ldots, N_{p1}(i)$.

The output of cell $MP2_i$, $N_{mp2}(i, t)$, is formed by

$$N_{mp2}(i, t) = m_{mp2}(i) + h_1 \tag{6}$$

where parameter h_1 is the base activity, yielding spontaneous firing in the MP2 layer.

Because of cumulative shrinking, the stimulus pattern is not evenly represented by the neuron group $P1_1$, $P1_2, \ldots, P1_i$, but by a distribution of normalized cell activities maximized at $MP2_i$ (referred to as the *representative* cell of the stimulus). In this cell group, the more a cell is to the left (see Fig. 2) the less it contributes to the representation due to shrunk activity. This effect can be tuned by choosing an appropriate function for W_{ij}, and in implementation we chose $W_{ij} = (i - j)/3$ for $j > i$. Figure 3 illustrates the effect of the cumulated shrinking with the AT inputs for the eight worm-like stimuli.

MP3 layer

The MP3 layer receives a sole, excitatory, input from MP2. the membrane potential of the ith MP3 cell

$$\frac{dm_{mp3}(i, t)}{dt} = -A_{mp3} m_{mp3}(i, t) + y_i(t) N_{mp2}(i, t) \tag{7}$$

where A_{mp3} is a relaxation parameter and $y_i(t)$ is the weight of the projection from cell $MP2_i$. Weight $y_i(t)$ is habituatable with initial value $y_i(0) = y_0$, as described below. Due to spontaneous activity in layer MP2, there is also spontaneous activity in layer MP3 following (7). The output of cell $MP3_i$ is formed by $N_{mp3}(i, t) = m_{mp3}(i, t)$.

MP1 layer

The MP1 layer receives its sole, inhibitory, input from MP3. The membrane potential of the ith MP1 cell is formed by

$$\frac{dm_{mp1}(i, t)}{dt} = -m_{mp1}(i, t) + h_2 - B_{mp1} N_{mp3}(i, t) \tag{8}$$

where B_{mp1} is a weight parameter of the inhibitory projection from the MP3 layer. Without external input, the membrane stabilizes at h_2, generating spontaneous activity. The inhibition from MP3 reduces the level of MP1 spontaneous activity. The output of cell $MP1_i$ is produced by $N_{mp1}(i, t) = [m_{mp1}(i, t)]^+$, with $[x]^+ = \text{Max}(0, x)$ to ensure a non-negative response.

P2 layer

This layer receives excitatory input from layer MP2 through habituatable projections, and unilateral inhibition from layer MP1. This unilateral projection integrates habituation effects to different stimuli, whereas the unilateral projection from layer P1 to layer MP2 forms group coding of the stimuli. The membrane potential of the ith P2 cell,

$$\frac{dm_{p2}(i, t)}{dt} = -A_{p2} m_{p2}(i, t) + y_i(t)(N_{mp2}(i, t) - h_1) - B_{p2} \sum_{j>i} [N_{mp1}(i, t) - C_{mp1}]^+ \tag{9}$$

where A_{p2} is a relaxation parameter, and $y_i(t)$ as appeared in (7) is the modifiable weight of the projection from cell $MP2_i$, to be described below. The term $(N_{mp2}(i, t) - h_1)$ detects external stimulation from the MP2 layer, which is above the level of spontaneous firing. In other words, only external stimulation from layer MP2 can pass through the plastic pathway. B_{p2} is the strength of left inhibition (Fig. 2), and C_{mp1} is the resting activity of an MP1 neuron before any habituation occurs. From (6)–(8), $C_{mp1} = h_2 - h_1 * y_0 * B_{mp1}$. The output of this P2 cell is generated by $N_{p2}(i, t) = [M_{p2}(i, t)]^+$.

This layer combines habituation of a stimulus pattern with unilateral inhibition, and it is where learning effects are reflected. Both layer P2 and layer MP3 receive plastic inputs from the MP2 layer. One difference between the P2 and MP3 layers is that the latter is affected by the spontaneous activity in layer MP2 (7) whereas the former is not (9). In fact, it is crucial for the function of the MP3 layer to reduce its spontaneous activity following habituation training, as outlined before. Finally, as the sole efferent neuron of the MP column, cell OUT simply integrates activities from the P2 layer.

Now let us describe synaptic plasticity. Current neurobiological studies suggest that short-term habituation operates on presynaptic terminals as a result of reduced neurotransmitter release (Kandel 1976;

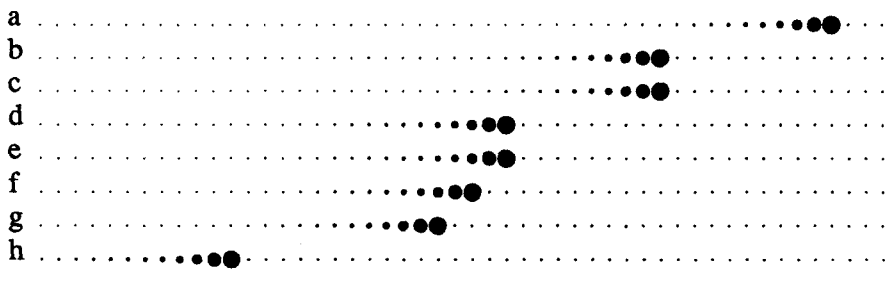

Fig. 3. Effect of cumulative shrinking. Each row shows the activity distribution of an array of 50 MP2 cells to a specific input I (1). The membrane potential of each cell is proportional to the diameter of a circle, but with zero activity represented by a dot. In the plot, inputs are the AT response to the eight worm-like stimuli in the dishabituation hierarchy (Fig. 1A), which are indicated by a name at the begining of each row. The parameter values are: $A_{p1} = 1.0$; $B_{p1} = 1.0$; $\varrho = 0.05$; $A_{mp2} = 0.1$; and $B_{p1} = 1.1$

Thompson 1986). Long-term habituation, however, may be accompanied by structural changes (Bailey and Chen 1983). The decrease of synaptic efficacy is mostly modeled by a first-order differential equation that the rate of reduction of the membrane potential is proportional to external stimulation, simulating exponential curve of habituation and spontaneous recovery (see among others Lara and Arbib 1985; Gluck and Thompson 1987). Wang and Hsu (1988; 1990) used an *S*-shaped curve to model the build-up of habituation. Since the *S*-shaped curve (see Fig. 4) has two varying courses depending on the sign of the second-order derivative and a single turning point, the inflection point, both short-term and long-term habituation are represented and the transfer of short-term to long-term memory corresponds to the switch from below to above the inflection point. Adapting their scheme to our model of the MP column, we now provide the dynamics for the synaptic weights $y_i(t)$ introduced in (7) and (9) as

$$\tau \frac{dy_i(t)}{dt} = \alpha z_i(t)(y_0 - y_i(t)) - \beta y_i(t)[N_{mp2}(i, t) - h_1]^+ \tag{10}$$

$$\frac{dz_i(t)}{dt} = \gamma z_i(t)(z_i(t) - 1)[N_{mp2}(i, t) - h_1]^+ \tag{11}$$

where τ is the time constant for controlling the rate of habituation (reduction of the weight $y_i(t)$). The first term in (10) regulates recovery towards the initial value (before any habituation training occurs) y_0 of y_i. The product $\alpha z_i(t)$ has an activity dependent control on the rate of forgetting. The second term regulates habituation, and parameter β controls the speed of habituation. The term $[N_{mp2}(i, t) - h_1]^+$ is the effective input from layer MP2 (the spontaneous activity h_1 in the MP2 layer does not affect habituation), and it multiplies $y_i(t)$ to form activity gated input. The intuition behind the activity gated input is that a habituation stimulus is more effective in an early stage of habituation when habituation starts to grow than in a late stage when habituation becomes profound. We can see from this term that habituation is regulated only by presynaptic input, which corresponds to the hypothesis that habituation is presynaptic.

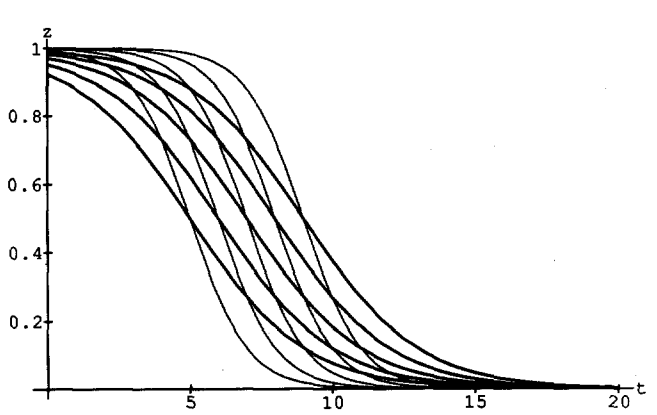

Fig. 4. Ten $z_i(t)$ functions with different parameter values. Parameter $t_0 = 5, 6, 7, 8, 9$ respectively, and $\gamma = 0.5$ for the *thick curves* and 1.0 for the *thin curves*

We want to adjust $z_i(t)$ so that if the column is activated for a short term, habituation rapidly disappears. We need $z_i(t)$ large for habituation to be short term. On the other hand, if the column is activated for a long time, we want to achieve long-term habituation, and for this we need $z_i(t)$ small. Equation (11), which essentially exhibits an *S*-shaped curve, naturally achieves both these conditions. To study the behavior of (11), let us assume a constant input $[N_{mp2}(i, t) - h_1]^+ = 1$ to the presynaptic terminal. Then the right-hand side of (11) becomes $\gamma z_i(t)(z_i(t) - 1)$. With the initial condition $z_i(t_0) = 0.5$, the solution to the equation is $z_i(t) = 1/(1 + \mathrm{Exp}(\gamma(t - t_0)))$. This *S*-shaped curve of $z_i(t)$ has an inflection point at $t = t_0$. The transition speed from the first fast-decreasing stage to the second slow-decreasing one is controlled by the slope at the inflection point, which is equal to $-\gamma/4$. The larger is γ, the quicker is the transition. The overall speed of the decrease of the curve $z_i(t)$ is controlled by the value of t_0, and the larger is t_0 the slower is the speed of decrease. Furthermore, a larger t_0 corresponds to a larger $z_i(0)$, the initial value of z_i. Therefore, the initial value of z_i in (11) governs the speed of decrease of the curve. Figure 4 shows two groups of z_i curves with different values of γ and t_0.

The effect of $z_i(t)$ on the synaptic weight is to control the speed of forgetting of habituation. As seen from Fig. 4, before the inflection point $z_i(t)$ holds a relatively large value, and forgetting as manifested by the first term of (10) is relatively fast; after the inflection point, $z_i(t)$ holds a relatively small value, and forgetting is relatively slow. These two phases are used to model two phases of memory: short-term memory (STM) and long-term memory (LTM). In an extreme case, when $z_i(t)$ equals its asymptote, 0, habituation as accumulated by training in (10) never involves forgetting. It is imaginable that with very small value of $z_i(t)$, forgetting may take longer than the time span of a certain animal. This is equivalent to say that long-term memory is never lost. The time course of transition from STM to LTM can be fully controlled by the model parameters γ and t_0 (t_0 is converted into $z_i(0)$ in real simulations).

3 Computer simulation

The single column model of the medial pallium presented before has been simulated. Before the results are presented, the following points need to be clarified: (a) The stimuli were originally presented to the model of the toad retina, and all the calculations from the retina to the anterior thalamus have been carried out by the Wang and Arbib model (1991a); (b) For comparison with the prey-catching response of the toad (see Ewert and Kehl 1978; Wang and Ewert 1992), the output of the medial pallium needs to affect the tectum where prey-catching behavior is generated. Modeling of this interaction is beyond the scope of this paper. Rather,

the response from MP is viewed as modifying the initial orienting activity.

For each array of the MP cell type, 50 cells have been simulated ($n = 50$). The results are shown by the activities of the single output cell of the column: OUT. The result seen from cell OUT is interpreted as a coefficient that modifies the number of initial orienting turns made by the animal in the first min interval when the stimulus is seen. Time was measured like this: a basic time step 0.05 (discretization step of the differential equations) corresponded to 1 s. Each stimulus was continuously presented for 60 min, and then it was switched to another stimulus. The second stimulus was also presented for 60 min. For visualization purposes we only present 10 data items for each presentation, each of which corresponds to a 6 min step. In all the simulations presented below, the stimuli were moved from left to right relative to the animal model as in the experiments, and they had the same length 20 mm and height 5 mm. See the legend of Fig. 5 for the values of all the parameters used in the MP column model.

In the first set of simulations, two pairs of stimuli were studied. In Fig. 5A, stimulus **f**, the left-point triangle, was presented first. The response of the model was habituated after continuous presentation of the stimulus. After 60 min simulation time, stimulus **d**, the rectangle, was presented, and it triggered a new response. That is, **d** dishabituated the habituated response to **f**. The reverse order of presentation was studied in Fig. 5B, and apparently no dishabituation was demonstrated. In Fig. 5C and D, another pair of stimuli was simulated. From the results we can see that stimulus **b**, the right-pointing triangle, was able to dishabituate the habituated response to the rectangle, but stimulus **d** was unable to dishabituate the habituated response to **b**. Figure 5E shows the corresponding experimental results.

The simulated toad model is just one "average individual", whereas in the experimental results, a sum-

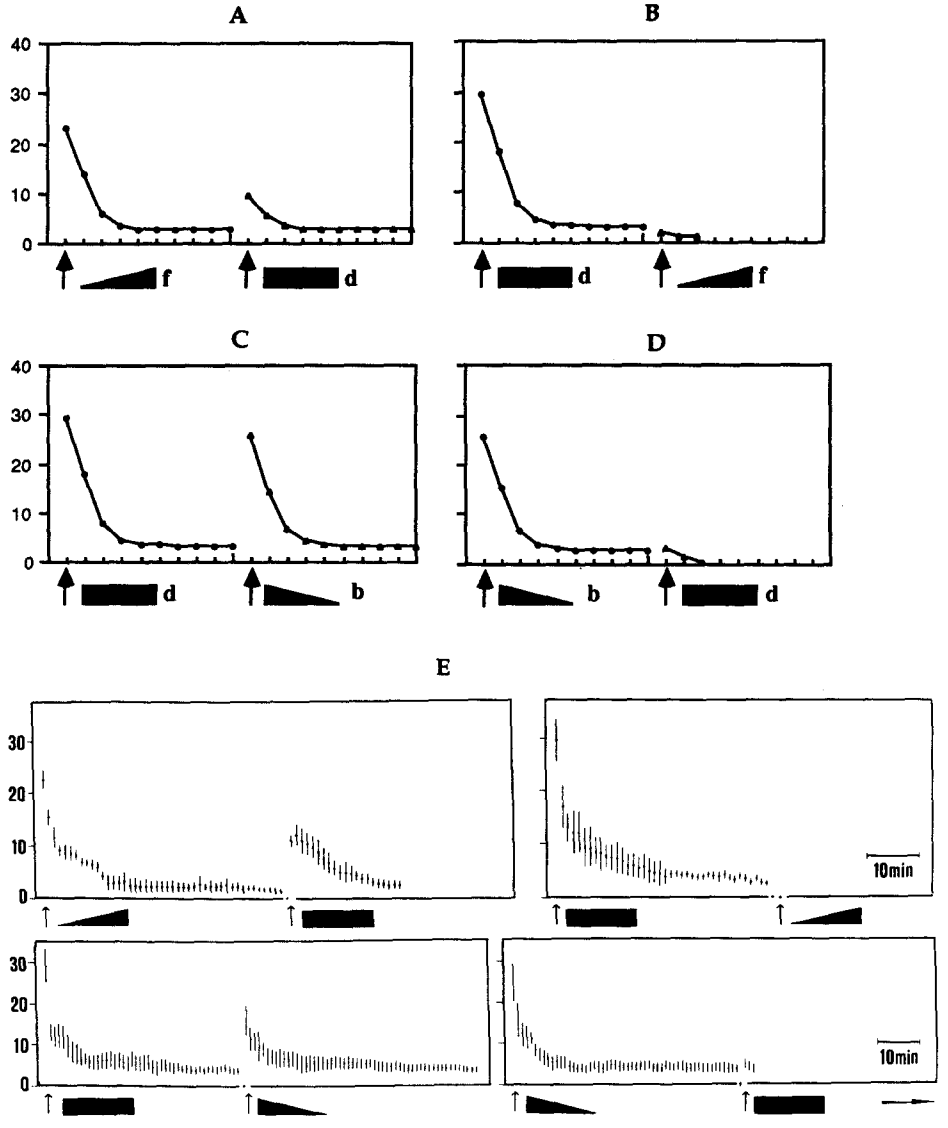

Fig. 5A–E. Computer simulation of habituation and dishabituation of prey orienting response in toads. The response was taken from the OUT cell of the MP column, and it was measured as the relative value of the initial response of each frame that is scaled to the same value measured experimentally. In a frame, the response to the first stimulus is indicated by little circles and that to the second one is indicated by little triangles. **A** Stimulus **f** was first presented and habituated, and then **d** was tested. **B** The reverse order of presentation. **C.** Stimulus **d** was first presented and habituated, and then **b** was tested. **D** The reverse order of presentation. **E** The experimental results obtained by Ewert and Kehl (1978) are shown (with permission) using the same combination of the stimuli. The parameter values are: $A_{p1} = 1.0$; $B_{p1} = 1.0$; $\varrho = 0.05$; $A_{mp2} = 0.1$; $B_{mp2} = 1.1$; $h_1 = 0.6$; $A_{mp3} = 1.0$; $B_{mp3} = 1.0$; $h_2 = 0.6$; $B_{mp1} = 1.0$; $A_{p2} = 1.0$; $B_{p2} = 0.1$; $A_{out} = 0.1$; $y_0 = 1.0$; $\tau = 200$; $\alpha = 3.2$; $\beta = 24$; $\gamma = 0.1$; and $z_i(0) = 0.99$

mary was taken from a group of individuals. From the comparison between Fig. 5A–D and Fig. 5E, it can be concluded that our model can clearly reproduce the experimental data with this set of stimulus pairs. The reduction of the initial activity of the second presentation as particularly seen in Fig. 5A is in the model due to overlapping in the population coding (see Fig. 3). The cells participated in coding stimulus **d** are partly involved in coding **f**. Thus, when **f** was habituated, the part of cells representing **d** was also habituated, and therefore the initial response to **d** was not as great as if without previous presentation of **f**. Note that the same phenomenon was exhibited in the experimental results (Fig. 5E).

In Fig. 6, we studied the pair of the top stimulus **a** and the bottom stimulus **h** in the dishabituation hierarchy. First, **h** was presented until habituation occurred, and then **a** was presented. A remarkable dishabituation was seen from the model response, as shown in Fig. 6A. However, after habituation to **a**, the model did not show any new response to presentation of **h**. Figure 6C shows the corresponding experimental results. A notable phenomenon with the **a**/**h** pair is that when **a** was presented after **h**, it elicited a larger initial activity than did **h**. This "overshooting" of dishabituation is exactly what occurred in the corresponding experiment (Fig. 6C). What yields this model phenomenon? This is because of the way of normalization of the cumulative shrinking mechanism. According to (5), after cumulative shrinking, the membrane potential of the MP2 representative cell of a stimulus is equal to $B_{mp2}N_{p1}(i, t)/(A_{mp2} + N_{p1}(i, t))$, which is a monotonically increasing function of $N_{p1}(i, t)$. Since $N_{p1}(i, t)$ is proportional to the AT input (1), The bigger is the AT response to a stimulus, the larger is the activity of the MP2 representative cell. The difference in activities of representative cells propagates through other layers of the column model until it reaches the OUT cell, where the initial overshooting caused by **a** was seen after habituation to **h** because the AT response to **a** is larger than to **h**.

In the last simulation we shall present, a separate process of dishabituation was investigated. The same procedure as in the experiment of Wang and Ewert (1992) was adopted in the simulation. Stimulus **f** was first presented for 60 min simulation time, until the response was habituated. Then stimulus **b** was presented for 0.5 min, and immediately afterwards **f** was brought back again. The simulation results are presented in Fig. 7A. Each little square in the plot represents the model response in a 4-min step, but the open circle represents half of the model response within that 0.5 min period, corresponding to the number of orienting turns in 0.5 min in the experiments. A strong dishabituation was shown when **b** was presented, just as in Fig. 7B which shows the corresponding experimental results. But habituation of **f** maintained when **b** was withdrawn and **f** was presented again. Presentation of **b** did not counteract the habituation effects caused by previous presentation of **f**, exactly as happened in the experiments.

In all simulation results presented so far, as demonstrated by Figs. 5–7, the model reproduces the experimental data remarkably well. Similarly good results were found in other simulations omitted here for space. Not only is the dishabituation hierarchy fully preserved as indicated by average firing rate of the AT model presented in Wang and Arbib (1991a), but also habituation properties, such as habituation curves and overshooting of dishabituation, are demonstrated by the model as seen in quantitative experimental data.

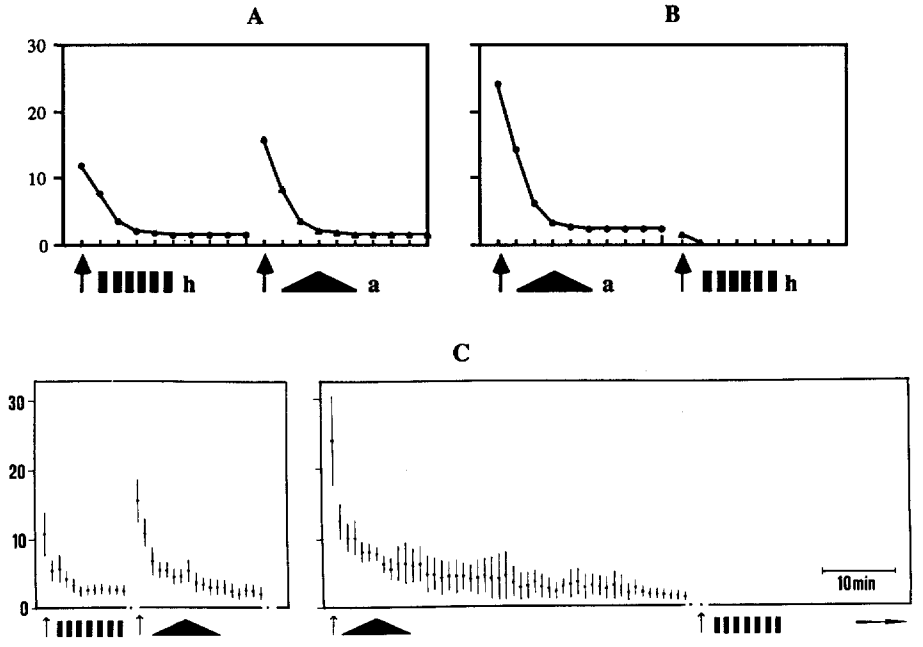

Fig. 6A, B. Computer simulation of habituation and dishabituation of prey orienting response in toads. **A** Stimulus **h** was first presented and habituated, and then **a** was tested. **B** The reverse order of presentation. **C** The corresponding experimental results obtained by Ewert and Kehl (1978, with permission). See the legend of Fig. 5 for parameter values and other explanations

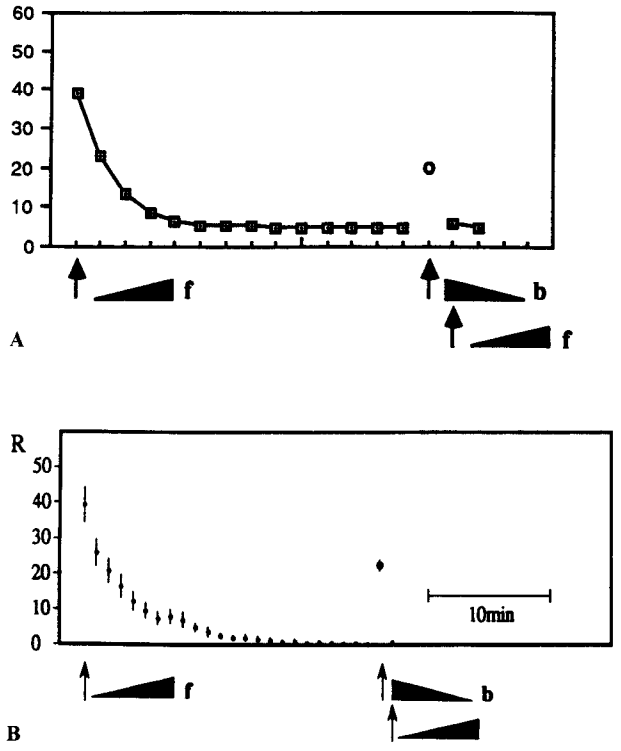

Fig. 7A. Computer simulation of separate process of dishabituation of prey orienting response in toads. Stimulus **f** was first presented for 60 min of simulation time, and then **b** was tested for 30 s. Immediately afterwards, **f** was tested again. See the legent of Fig. 5 for parameter values and other explanations. **B** The corresponding experimental results obtained by Wang and Ewert (1992)

4 Predictions

This model of the medial pallium extends the *Rana computatrix* model of visual information processing in frog and toad (Arbib 1991) one step further than the anterior thalamus, which is based on retinal and tectal processing. The theory we have developed in the previous sections yields a number of neurobiological predictions, as presented below.

Dishabituation mechanisms

Why can presentation of a certain stimulus dishabituate habituation to some stimuli, but not to others? That is, what causes dishabituation? The model allows us to predict that dishabituation is nothing but the release of a normal prey-catching behavior. The reason that presentation of a second stimulus fails to release a new response, however, is because of inhibition caused by habituation to the first stimulus. Thus, dishabituation is not a result of facilitation, as the name may suggest. The failure of dishabituation is caused by cross-talk between the first and the second stimulus. In light of the prediction, dishabituation naturally has a separate neuronal process, as indicated by the experimental results shown in Fig. 7B.

Long-term memory effects

In habituation training, one session of an experiment usually takes tens of minutes. An interesting phenomenon occurs if we observe the data curves carefully (see Figs. 5–7 for experimental data curves): toads did not reduce their orienting activity to zero, but rather saturated at some level. In experiments, training with one stimulus was often stopped when the toad reacted at some low level of activity stably for a period of time (Ewert, personal communication, 1990). The same phenomenon was also observed when the experiments reported by Wang and Ewert (1992) were conducted. Our model offers an explanation for why prey-catching orienting activity saturates at a non-zero level. According to the model of synaptic plasticity (10 and 11), we know that at equilibrium the weight of a plastic synapse will settle down at $y_i(t) = \alpha y_0 z_i(t)/(\alpha z_i(t) + \beta(N_{mp2}(i, t) - h_1))$, which is a non-zero value with a constant input from the MP2 layer. This is the reason why in all simulation results shown above, the model also saturated at a non-zero activity level. It is straightforward that if $z_i(t)$ becomes smaller, the saturation point will be lower. According to the analysis in Sect. 2, $z_i(t)$ can be significantly reduced only following long-term habituation training, which can be obtained with a series of training trials. We thus predict that, due to the buildup of long-term memory traces, toads will saturate at lower and lower levels of activity as a series of training proceeds, and eventually reach zero level. Theoretically speaking, the same long-term effect should occur if sufficiently long training is applied even with one session. It is interesting to note that Cervantes-Pérez et al. (1991) once tested an animal for 220 min, in order to reach the zero level.

Resting habituation

We know from the model that if a higher intensity stimulus H is presented first to the MP model and later habituated, a lower intensity stimulus L cannot elicit a new response. Would habituation still occur toward repeated presentation of L when L cannot elicit a response? Our model predicts that a habituation process still takes place in this situation. This is because in the model habituation is due to the reduction of the efficacies of the synapses that MP2 cells make on P2 and MP3 cells and it solely depends on presynaptic inputs from MP2 cells. The reason that L cannot elicit a new response is due to unilateral inhibition resulting from habituation to H that exerts on P2 cells. Although unable to release a prey-catching response, stimulation of L elicits normal MP2 activities, thus causing habituation process to happen. We refer to this kind of habituation as resting habituation.

Influence of habituation on dishabituation

Our model provides a framework which allows us to investigate phenomena beyond those recognized so far by experimentalists. Particularly, we consider in a stimulus pair that the relative amplitude of the initial

response of the second presentation is not a coincidence, but tells a rich story of how the first presentation affects the second one. To summarize what is studied in the last section, we offer the following explanations: (a) The overlapping of cell participation in representing different patterns leads to a smaller amplitude of initial activity elicited by a dishabituating stimulus, as exemplified in Fig. 5A and its corresponding experimental data; (B) Overshooting of dishabituation is a result of higher AT input to MP from the second presentation of H in comparison to the first presentation of L.

5 Discussion

In the MP column, we have introduced five types of neuron, each of which plays a different role in visual information processing. The existence of three of them, MP1, MP2, and MP3 is supported by physiological recordings. In the model, all these three types show spontaneous activities. MP2 cells are not affected by habituation (Fig. 2), whereas after habituation training MP3 cells decrease and MP1 cells increase their spontaneous firing activities. These properties of model cell types conform with the original classification of these cell types (Finkenstädt 1989a). The fact that MP shows a significant increase in 2DG uptake after habituation (Finkenstädt and Ewert 1988a) may be a result of net outcome of a relatively greater increase in MP1 activities and a decrease in MP2 activities. Introduction of P1 and P2 types is based on computational considerations.

Spontaneous activities as shown in MP1, MP2 and MP3 cells have an important computational role in the model. A cell of this type can either increase or decrease its stable activity level, dependent on different situations, whereas a cell without spontaneous firing can only increase its activity level after stimulation. It is the reduction of spontaneous activity of MP3 cells after habituation that gives rise to inhibition which can be exerted upon a later presentation of a different stimulus. The aftereffects of habituation are manifested through decrease and increase in spontaneous activity of MP3 and MP1 cells respectively, and finally through unilateral inhibition stemming from MP1 cells. From another perspective, our modeling suggests important computational values of having spontaneous activities in the central nervous system.

In the model, different visual stimuli are represented by activity distributions of different cell groups (see Fig. 3). Presentation of a new stimulus triggers a new response because it activates a different group of cells. Hierarchical stimulus specificity exhibited in amphibians can be explained by unilateral inhibition within a functional unit, as depicted in Fig. 2. The dishabituation mechanisms we expound in this paper are different from those revealed in the study of *Aplysia* (Kandel 1976), where dishabituation caused by sensitization is due to facilitation from an external channel. On the other hand, our model structure is consistent with neuronal substrates proposed by the dual-process theory (Groves and Thompson 1970) which claims that sensitization (dishabituation) develops independently. This medial pallium model provides a computational basis for the comparator model (Sokolov 1960) of stimulus specific habituation in vertebrates. If we leave out the unilateral inhibition from MP1 cells, the same structure can readily exhibit bilateral dishabituation implicated by the comparator model.

In the model, short-term and long-term memory are modeled by a single S-shaped curve. Long-term memory develops as training accumulates. Conciseness is one functional advantage of modeling two phases of learning by a single curve. The model does not imply that STM and LTM share the same neural mechanisms. What it says is that at the electrophysiological level, the two forms of memory could be represented by a single dynamic process. This combination might yield some insights into the relation between STM and LTM biological processes. This view is also consistent with the fact that in *Aplysia* the two forms of habituation share the same locus – the presynaptic terminals of sensory neurons (Castellucci et al. 1978). In our model, $y_i(t)$ (see 10) can be modified by presynaptic stimulation quite rapidly during a short time period while $z_i(t)$ (see 11) stays basically the same, and in terms of the synaptic mechanisms involved the change in $y_i(t)$ would correspond to the change in neurotransmitter release. A change to $z_i(t)$ needs much more profound training in the model than a change to $y_i(t)$, and this change would correspond to structural changes in presynaptic synapses (Bailey and Chen 1983). Available toad data on long-term memory do not permit fine tuning on the model of LTM. Perhaps the only data item is a recent observation by Cervantes-Pérez et al. (1991) that stimulus-specific habituation may last hundreds of days. Nonetheless, consideration of long-term memory has provided valuable information for modeling the MP structure. A preliminary version of the MP model (Wang and Arbib 1991b) can account for short-term habituation effects, but fails to explain long-term storage. With the present model, habituation can be maintained for a very long time depending on parameters chosen for formula (11), since presentation of a different stimulus does not disrupt previous memory, as it did in the previous version.

MP is the structure where different learning processes are integrated. The anatomy of our present model is a neural network with minimum complexity that can explain various experimental data. We hypothesize that in anurans, stimulus-specific conditioning shares the neural circuitry for stimulus-specific habituation. Future medling will address the role of MP in conditioning and learning related with predator stimuli, and in olfactory processing.

References

Arbib MA (1991) Neural mechanisms of visuomotor coordination: the evolution of rana computatrix. In: Arbib MA, Ewert J-P (eds) Visual structures and integrated functions. Research notes in neural computing. Springer, Berlin Heidelberg New York, pp 3–30

Bailey CH, Chen MC (1983) Morphological basis of long-term habituation and sensitization in *Aplysia*. Science 220:91–93

Castellucci VF, Carew TJ, Kandel ER (1978) Cellular analysis of long-term habituation of the gill-withdrawal reflex of *Aplysia California*. Science 202:1306–1308

Cervantes-Pérez F, Guevara-Pozas AD, Herrera-Becerra AA (1991) Modulation of prey-catching behavior in toads: Data and modeling. In: Arbib MA, Ewert J-P (eds) Visual structures and integrated functions. Research notes in neural computing. Springer, Berlin Heidelberg New York, pp 397–416

Eikmanns KH (1955) Verhaltensphysiologische Untersuchungen über den Beutefang und das Bewegungssehen der Erdkröte (*Bufo bufo* L.). Z Tierpsychol 12:229–253

Ewert J-P (1984) Tectal mechanisms that underlie prey-catching and avoidance behaviors in toads. In: Vanegas H (ed) Comparative neurology of the optic tectum. Plenum, New York, pp 246–416

Ewert J-P (1987) Neuroethology: toward a functional analysis of stimulus-response mediating and modulating neural circuitries. In: Ellen P, Thinus-Blonc C (eds) Cognitive processes and spatial orientation in animal and man, pt. 1. Martinus Nijhoff, Dordrecht, pp 177–200

Ewert J-P, Ingle D (1971) Excitatory effects following habituation of prey-catching activity in frogs and toads. J Comp Physiol Psychol 77:369–374

Ewert J-P, Kehl W (1978) Configurational prey-selection by individual experience in the toad *Bufo bufo*. J Comp Physiol A 126:105–114

Finkenstädt T (1989a) Stimulus-specific habituation in toads: 2DG studies and lesion experiments. In: Ewert J-P, Arbib MA (eds) Visuomotor coordination: amphibians, comparisons, models, and robots. Plenum, New York, pp 767–797

Finkenstädt T (1989b) Visual associative learning: searching for behaviorally relevant brain structures in toads. In: Ewert J-P, Arbib MA (eds) Visuomotor coordination: amphibians, comparisons, models, and robots. Plenum, New York, pp 799–832

Finkenstädt T, Ewert J-P (1988a) Stimulus-specific long-term habituation of visually guided orienting behavior toward prey in toads: a ^{14}C-2DG study. J Comp Physiol A 163: 1–11

Finkenstädt T, Ewert J-P (1988b) Effects of visual associative conditioning on behavior and cerebral metabolic activity in toads. Naturwissenschaften 75:85–87

Gluck MA, Thompson RF (1987) Modeling the neural substrates of associative learning and memory: A computational approach. Phychol Rev 94:1–16

Grossberg S (1976) Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. Biol Cybern 23:121–134

Grossberg S, Kuperstein M (1986) Neural dynamics of adaptive sensory-motor control. North-Holland, Amsterdam

Groves PM, Thompson RF (1970) Habituation: a dual-process theory. Psychol Rev 77:419–450

Herrick CJ (1933) The amphibian forebrain. VIII. Cerebral hemispheres and pallial primordia. J Comp Neurol 58:737–759

Hoffman HH (1963) The olfactory bulb, accessory olfactory bulb and hemisphere of some anurans. J Comp Neurol 120:317–368

Kandel ER (1976) Cellular basis of behavior: An introduction to behavioral neurobiology. Freeman, New York

Kicliter E, Ebbesson SOE (1976) Organization of the "nonolfactory" telencephalon. In Llinás R, Precht W (eds) Frog neurobiology. Springer, Berlin Heidelberg New York, pp 946–972

Lara R, Arbib MA (1985) A model of the neural mechanisms responsible for pattern recognition and stimulus specific habituation in toads. Biol Cybern 51:223–237

Sokolov EN (1960) Neuronal models and the orienting reflex. In: Brazier MAB (ed) The central nervous system and behavior: III. Macy Foundation, New York, pp 187–276

Thompson RF (1986) The neurobiology of learning and memory. Science 233:941–947

Wang DL, Arbib MA (1991a) How does the toad's visual system discriminate different worm-like stimuli? Biol Cybern 64:251–261

Wang DL, Arbib MA (1991b) Hierarchical dishabituation of visual discrimination in toads. In: Meyer J-A, Wilson S (eds) Simulation of adaptive behavior: From animals to animats. MIT Press, Cambridge, pp 77–88

Wang DL, Ewert J-P (1992) Configurational pattern discrimination responsible for dishabituation in common toads *Bufo bufo* (L.): Behavioral tests of the predictions of a neural model. J Comp Physiol A 170:317–325

Wang DL, Hsu CC (1988) A neuron model for computer simulation of neural networks. Acta Automatica Sinica 14:424–430

Wang DL, Hsu CC (1990) SLONN: A simulation language for modeling of neural networks. Simulation 55:69–83

Wang DL, King IK (1988) Three neural models which process temporal information. In Proc 1st Annu Conf Intern Neur Netw Soc. Boston, MA, pp 227