REVIEW

P. Read Montague[1] and
Terrence J. Sejnowski[2]

[1]Division of Neuroscience
Baylor College of Medicine
Houston, Texas 77030
[2]Howard Hughes Medical Institute
The Salk Institute for Biological Studies
La Jolla, California 92037

The Predictive Brain: Temporal Coincidence and Temporal Order in Synaptic Learning Mechanisms

Abstract

Some forms of synaptic plasticity depend on the temporal coincidence of presynaptic activity and postsynaptic response. This requirement is consistent with the Hebbian, or correlational, type of learning rule used in many neural network models. Recent evidence suggests that synaptic plasticity may depend in part on the production of a membrane permeant-diffusible signal so that spatial volume may also be involved in correlational learning rules. This latter form of synaptic change has been called volume learning. In both Hebbian and volume learning rules, interaction among synaptic inputs depends on the degree of coincidence of the inputs and is otherwise insensitive to their exact temporal order. Conditioning experiments and psychophysical studies have shown, however, that most animals are highly sensitive to the temporal order of the sensory inputs. Although these experiments assay the behavior of the entire animal or perceptual system, they raise the possibility that nervous systems may be sensitive to temporally ordered events at many spatial and temporal scales. We suggest here the existence of a new class of learning rule, called a *predictive Hebbian learning rule*, that is sensitive to the temporal ordering of synaptic inputs. We show how this predictive learning rule could act at single synaptic connections and through diffuse neuromodulatory systems.

Introduction

Most biologically feasible theories of how experience-dependent changes take place in real neuronal networks use some variant of the notion that the efficacy or "strength" of a synaptic connection from one cell to another can be modified on the basis of its history. In this theoretical work it is generally assumed that modifications of synaptic efficacy, by acting over a large population of synapses, can account for interesting forms of learning and memory. This theoretical assumption prevails primarily because of its intuitive appeal, its accessibility to analysis, some provocative relations to biological data, and a lack of good alternatives.

Recent work demonstrates that simple abstract learning algorithms, if given appropriately coded input, can produce complicated mappings from input to output. These efforts include networks that learn to pronounce written text (Sejnowski and Rosenberg 1987), play master level backgammon (Tesauro 1994), and recognize handwritten characters (Le Cun et al. 1990). As pointed out by Crick (1989) and others, many of these efforts are not good models of the vertebrate brain; however, they can be quite valuable for identifying the informational requirements

involved in specific tasks. Moreover, they point out some of the computational constraints to which brains are subject. An awareness of the computational constraints involved in a particular problem can guide theories that explain how real brains are constructed (Churchland and Sejnowski 1992).

Although abstract networks have provided some insight into the top-down constraints that nervous systems face, these approaches are of limited use in gaining insight into how various problems have been solved by real brains. For example, the actual learning mechanisms that are used in biological systems also satisfy additional constraints that arise from the known properties of neurons and synapses. In this paper we focus on learning rules that are supported by biological data and consider the strengths and weaknesses of these rules by measuring them against both computational and biological constraints. Taking this dual approach, we show that computational concerns applicable to the behavior and survival of the animal can work hand in hand with biologically feasible synaptic mechanisms to explain and predict experimental data.

Correlational Learning Rules—Learning Driven by Temporal Coincidence

Theoretical accounts of how neural activity actually changes synaptic function typically rely on a local correlational learning rule to model synaptic plasticity. A correlational learning rule, often called a Hebbian learning rule, uses the correlation between presynaptic activity and postsynaptic response to drive changes in synaptic efficacy (Fig. 1) (Hertz et al. 1991; Churchland and Sejnowski 1992). One simple expression of a Hebbian learning rule is

$$\Delta w(t) = \eta x(t)y(t) \tag{1}$$

where, at time t, $w(t)$ is a connection strength (weight), $x(t)$ is a measure of presynaptic activity, $y(t)$ is a measure of postsynaptic activity (e.g., firing rate or probability of firing), and η is a fixed learning rate.

This kind of learning rule is called local because the signals sufficient for changing synaptic efficacy are assumed to be generated locally at each synaptic contact. One form of this learning rule was initially proposed by Donald Hebb in 1949 (Hebb 1949). Subsequent theoretical and computational efforts have exploited Hebb's idea and used correlational learning rules to account successfully for aspects of map formation and self-organization of visual and somatosensory cortex (von der Malsburg 1973; von der Malsburg and Willshaw 1977; Montague et al. 1991). For example, various computational schemes employing Hebbian learning rules have accounted for the formation of cortical receptive fields (Bienenstock et al. 1982; Linsker 1986, 1988), ocular dominance columns (Miller et al. 1989), orientation maps (von der Malsburg 1973; Obermayer et al. 1990; Miller 1994), directional selectivity (Sereno and Sereno 1991), and disparity tuning (Berns et al. 1993). Correlational learning rules also provide a reasonable theoretical framework for synaptic plasticity observed in the hippocampus (Kelso et al. 1986; Bliss and Lynch 1988), cerebellum (Ito 1986, 1989), and neocortex (Kirkwood et al. 1993).

Below, we review some of the biological evidence from the vertebrate nervous system that supports this simple learning rule as a descriptor of synaptic change during both activity-dependent development and synaptic modification in the adult. We subsequently suggest that changes
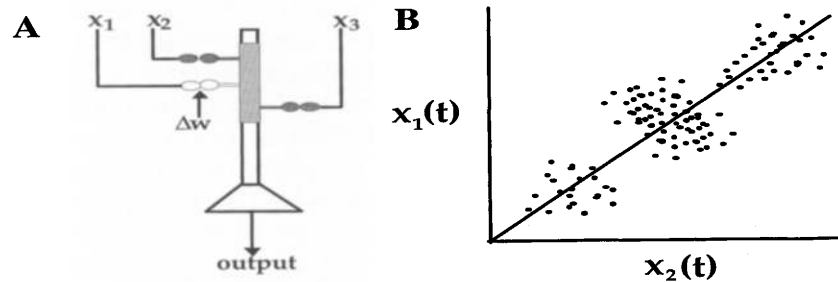
Figure 1: Hebbian Learning. (A) Inputs x_i provide excitatory drive to a neuron through connection strengths or "weights". Inputs x_2 and x_3 are sufficiently correlated to permit cooperation along a section of dendrite (shaded area) through voltage or second messengers. Through an expression like equation 1, the weights of these connections will be increased. Input x_1 is not active during this coincident activation of x_2 and x_3. The weight of x_1's connection could be decreased by a depression rule that depressed all synaptic contacts that were not sufficiently correlated with the postsynaptic response (shaded area). Without such a rule, weights can grow without bound. To prevent this, a homeostatic constraint that limits the total synaptic strength supported by the recipient neuron is typically used. This is just one possible way to normalize the weights. The issue of how and why normalization is biologically reasonable is critical. Normalization can give stability to the Hebb rule, but, depending on its implementation, it can cause the weight vector to converge to different values. In the presence of additional constraints (see text) for the learning rule, a Hebb rule will extract the principal component from the correlations in the input patterns that occur and the vector of weights will come to point in the direction of the first principle component of the "data" generated by the input activities (Oja 1982). The pattern of weights that develops can be analyzed in terms of the covariance matrix of the input activities (see text). (B) Graph of input activity along two inputs, x_1 and x_2. Each point is a pair of activity levels for the two inputs in (A). The inputs cluster along a straight line, indicating a strong correlation. The approximate direction of the principal component is along this line.

in this description are required by both experimental and theoretical work. The primary change is predicated on the need for brain mechanisms sensitive to temporally ordered input, a problem that has most likely been solved by brains across a range of spatiotemporal scales. We marshal arguments and review detailed evidence in support of this suggestion and point out those aspects of the proposed changes that are important for understanding learning and memory in the vertebrate brain.

Developmental Evidence for Hebbian Learning Rules

In the vertebrate nervous system, afferent axons find their appropriate target structures through interactions with local environmental cues and target-derived information (Bonhoeffer and Huff 1985; Dodd and Jessel 1988; Stuermer 1988; Harris 1989; Heffner et al. 1990; O'Leary et al. 1990; Placzek et al. 1990; Stretevan 1990). After reaching target structures, there is strong evidence that activity-dependent processes are critical in determining the development of mappings between peripheral sensory structures and their more centrally located target structures, including the optic tectum, thalamus, and cerebral cortex (Hubel and Wiesel 1965, 1970; Hubel et al. 1977; Meyer 1982; Stryker and Harris 1986; Stretevan et al. 1988). Specific mappings arise in these targets because temporal contiguity in axonal firing is somehow translated into

spatial contiguity of synaptic contacts. Hence, activity-dependent processes are involved at least with the initial *self-organization* of mappings in the tectum, thalamus, and cortex.

After normal developmental periods, activity-dependent processes are also involved in the *reorganization* of sensory mappings in the adult. For example, the adult cerebral cortex has been shown to be surprisingly plastic (for review, see Kaas 1991; Merzenich and Sameshima 1993) following changes to the environment such as retinal damage (Kaas et al. 1990; Gilbert and Wiesel 1992), changes in limb innervation (Merzenich et al. 1983, 1984; Wall et al. 1986; Clark et al. 1988), artificial scotomas (Pettet and Gilbert 1992), and other dramatic perturbations of sensory input (Clark et al. 1988; Garraghty et al. 1988; Garraghty and Kaas 1992). The plasticity observed during both activity-dependent development and map reorganization is consistent with the hypothesis that changes in synaptic efficacy are controlled by Hebbian learning rules (Reiter and Stryker 1988; Bear et al. 1990; Singer 1990; Rauschecker 1991; Merzenich and Sameshima 1993; Schlagger et al. 1993).

Collectively, these data suggest that (1) dynamic synaptic changes can occur throughout adulthood, (2) there is a strong relationship between the rules controlling synaptic change during development and in the adult, and (3) the learning rules appear to be Hebbian.

Cellular and Synaptic Evidence for Hebbian Learning Rules

In the vertebrate, the existence of Hebbian-like learning rules is supported further by detailed experimental evidence from work on excitatory glutamatergic synapses. Research over the last 20 years, has demonstrated that in the adult nervous system, long-term increases (Bliss and Lomo 1973; Kirkwood et al. 1993) and decreases (Ito 1986, 1989; Artola et al. 1990; Sejnowski et al. 1990; Dudek and Bear 1992; Mulkey and Malenka 1992) in synaptic efficacy can occur under appropriate conditions.

One form of long-term increase in synaptic efficacy is called long-term potentiation or LTP. The induction of LTP depends on correlated presynaptic activity and postsynaptic depolarization and requires an increase in postsynaptic calcium ion concentration (Dolphin et al. 1982; Collingridge and Bliss 1993; Davies et al. 1989; Malenka et al. 1989). Although Hebb (1949) never mentioned decreases in synaptic efficacy resulting from a lack of coincidence in presynaptic activity and postsynaptic response, such an extension to the original postulate was made by Stent (1973). This idea also finds support in the vertebrate nervous system.

One form of long-term decrease in synaptic efficacy is called long-term depression or LTD. In the hippocampus, the induction of homosynaptic LTD requires presynaptic activity without a coincident "response" from the dendrite (Dudek and Bear 1992; Mulkey and Malenka 1992), that is, the presynaptic terminal and the postsynaptic dendrite are not active at the same time. The response alluded to above is probably an increase in postsynaptic calcium levels subsequent to presynaptic release of neurotransmitter (Mulkey and Malenka 1992). Other forms of synaptic depression have been described in the neocortex (Artola et al. 1990; Artola and Singer 1993).

In a number of systems, the induction of LTP depends on activation of the *N*-methyl-D-aspartate (NMDA) glutamate receptor (for review, see Bliss and Lynch 1988; Nicoll et al. 1988). Because the NMDA receptor

provides a pathway for postsynaptic increases in calcium, its function influences both LTP and LTD. Interestingly, activation of the NMDA receptor is also one key event in the activity-dependent segregation of axonal terminals that occurs during the initial development of sensory mappings described above (Cline et al. 1987; Kleinschmidt et al. 1987; Scherer and Udin 1989; Bear et al. 1990; Cline and Constantine-Paton 1990; Simon et al. 1992). Thus, the NMDA receptor plays a critical role in both activity-dependent development and synaptic change in the adult.

NMDA Receptor as a Hebbian Coincidence Detector

Experiments have demonstrated that NMDA-dependent LTP can be viewed as Hebbian: Sufficient correlation in presynaptic activity and postsynaptic depolarization leads to an increase in the efficacy of the synapse (Wigstrom and Gustafsson 1985; Kelso et al. 1986; Malinow and Miller 1986). Current flow through the NMDA receptor can occur only in the presence of bound L-glutamate and a sufficiently depolarized membrane, that is, the receptor is ligand and voltage gated. This dual control of the NMDA receptor implicates the receptor itself as one substrate for detecting the correlation in presynaptic activity (L-glutamate release) and postsynaptic response (postsynaptic depolarization).

Taken together, the data on activity-dependent development of cortical maps and NMDA-dependent LTP suggest that a Hebbian learning rule is a reasonable theoretical starting point for modeling activity-dependent development and adult learning. Figure 1 shows a concrete example of a Hebbian learning rule.

Computational Properties of Hebbian Rules

Hebbian learning rules are *unsupervised learning rules* in that they lack a teaching signal that could supervise learning at each synaptic connection in a detailed manner, although, as discussed in a later section, they can be extended to include an additional reinforcement component. Although networks using Hebbian rules have inputs and produce outputs, there is no precise error information that instructs the network about whether an output was correct or not. Instead, Hebbian rules extract particular correlations or regularities among the presynaptic inputs through an averaging process, and the results are represented in the pattern of weights in the network (Hertz et al. 1991; Churchland and Sejnowski 1992). Through averaging, only those features or regularities that are redundant are extracted. The kinds of regularities to which Hebbian rules are sensitive depends on the network architecture. In this discussion, we let x or $x(t)$ represent the vector of input fibers $x = [x_1(t), x_2(t), \ldots, x_\eta(t)]$ at time t, and $E[x]$ represents the vector of expected values of the components of x (see Fig. 1).

Biologically feasible correlational models depend on the assumption that learning (weight changes) at synaptic connections is slow relative to the time required to present a number of input patterns sufficient to represent the statistics of the input. This assumption permits the particular learning rule in question to be analyzed in terms of a normalized correlation matrix of the input activities, called the covariance matrix, defined as $\langle Cov_{ij} \rangle \equiv [x - E(x)][x - E(x)]^T$ (Linsker 1986, 1988; Miller et al. 1989). As discussed below, this assumption can allow extensive analysis of the learning rule. In some models, learning (weight changes) takes place as each input pattern is presented (e.g., von der Malsburg 1973). This latter assumption means that if the network contains nonlinearities, the order of presentation of the input

patterns can critically influence the final weight patterns that result; that is, the network does not necessarily extract the redundancies in the input activity.

Figure 1 shows the simplest form of a Hebbian learning rule. In this diagram we consider the response of one linear neuron in a target region whose output at time t is

$$y = \sum_{j=1}^{\eta} w_j x_j = w^T x \qquad (2)$$

where $\langle\,\rangle$ represents a time average sufficiently large so that condition 1 above is met.

The expression $C_{ij} \equiv \langle x_i x_j \rangle$ defines the correlation matrix for the input vector x and is equivalent

Using the Hebbian learning rule in equation 1 and assuming that the weight changes are slow relative to the time over which the input patterns are presented (case 1 above), the change in a synaptic weight w_i can be represented as (the dependence on time is suppressed)

$$\langle \Delta w_i \rangle = \langle x_i y \rangle = \left\langle x_i \sum_{j=1}^{\eta} w_j x_j \right\rangle = \left\langle \sum_{j=1}^{\eta} x_j x_i w_j \right\rangle = \sum_{j=1}^{\eta} C_{ij} w_j \qquad (3)$$

$$C_{ij} \equiv \langle x_i x_j \rangle$$

where $\langle\,\rangle$ represents a time average sufficiently large so that condition 1 above is met.

The expression $C_{ij} \equiv \langle x_i x_j \rangle$ defines the correlation matrix for the input vector x and is equivalent to the covariance matrix when the individual components of x have zero means. The correlation matrix of the input activities is a symmetric matrix and therefore has real eigenvalues and a complete set of orthonormal eigenvectors. The possibility of analyzing the Hebbian learning rule given by equation 1 in terms of the eigenvectors of the covariance matrix has made this learning algorithm attractive. In the absence of other constraints, the eigenvectors of the covariance matrix represent synaptic weight patterns that can develop independently of one another and the eigenvector associated with the largest eigenvalue will determine the final weight pattern that develops. This way of viewing the development of weight patterns (activity-dependent development) makes contact with standard methods of analysis used in many fields such as statistics, physics, and applied mathematics.

There are a number of problems with the simple learning rule in equation 1. One problem is that this rule allows the synaptic weights to grow without bound. To prevent synaptic weights from growing without bound, a kind of homeostatic constraint that limits the total synaptic strength supported by the recipient neuron is typically used. This is an important issue because without this extra constraint, a simple Hebbian learning rule is not stable.

Below, we consider three proposals for limiting the growth of synaptic weights under a correlational learning rule. We measure each proposal

against the nature of the information required to limit the total synaptic strength. For clarity, we assume that a synaptic weight represents the strength of a single synaptic connection.

WEIGHT DECAY (LOCAL CONSTRAINT)

One proposal made by Oja (1982) suggested that a local decay of the synaptic weight is sufficient to give a correlational learning rule stability. He considered the following rule (again the dependence on time is suppressed)

$$\Delta w_i = \eta[x_i y - w_i y^2] \qquad (4)$$

At each update of the synaptic weight, the decay of the weight is proportional to the square of the postsynaptic response. This rule requires a form of heterosynaptic long-term depression, in that the second, negative term depends only on postsynaptic activity. Oja (1982) has shown that this learning rule will extract the first principal component of the input correlation matrix. In a general sense, the first principal component represents the dominant input pattern (Hertz et al. 1991). The important biological aspect of this rule is that weight decay is reasonably viewed as a local event that could take place at single synapses or groups of cooperating synapses. Other correlational schemes have been used to extract all the principal components of the covariance matrix (Sanger 1989), but they have not been offered as feasible biological models.

CLIPPING OR SATURATION (LOCAL CONSTRAINT)

Another local method to limit total synaptic weight involves simply clipping the weight at maximum and minimum values. This method has been used in a number of different correlational schemes (e.g., Von der Malsburg and Willshaw 1976; Bienenstock et al. 1982; Edelman and Reeke 1982; Linsker 1986; Gally et al. 1990; Montague et al. 1991)

$$\Delta w_i = \begin{cases} 0 & w < w_{min} \\ \eta(x_i - \theta_{pre})(y - \theta_{post}) & w_{min} < w < w_{max} \\ 0 & w_{max} < w \end{cases} \qquad (5)$$

where θ_{pre} and θ_{post} are respectively pre- and postsynaptic modification thresholds (Sejnowski 1977). Cases where both $(x_i - \theta_{pre})$ and $(y - \theta_{post})$ are negative are usually ignored in biological models to prevent this condition from causing an increase in synaptic strengths contrary to experimental findings. Variations on this rule assume nonlinear functions for the postsynaptic term and sliding thresholds that depend nonlinearly on the postsynaptic activity (Bienenstock et al. 1982; Artola and Singer 1993). These types of schemes are intuitively acceptable from a biological perspective; however, allowing synaptic weights to saturate influences the final state of the weight vector by introducing a sensitive dependency on initial conditions. MacKay and Miller (1990) have provided a good analysis of this situation and have shown that the eigenvector associated with the largest eigenvalue does not always determine the final pattern of weights under learning rules similar to equation 5. The value of this form of learning rule is that only local information is needed to prevent uncontrolled growth of synaptic weights.

WEIGHT NORMALIZATION (GLOBAL CONSTRAINT)

Weight normalization refers to a procedure whereby some measure of the total synaptic weight onto the recipient neuron is used to limit the growth of the synaptic weights. In multiplicative weight normalization (see von der Malsburg 1973), each synaptic weight onto a neuron is divided by the sum or sum of squares of every other weight onto this neuron. This division is carried out after each update of the weights. This form of normalization rescales the synaptic weights continuously when they are updated. After updating the weights with Δw_i computed according to equation 1

$$w_i(t) = w_i(t-1) + \Delta w_i(t) \qquad (6)$$

The weights are then rescaled according to

$$w_i(t) \leftarrow \frac{w_i(t)}{W} \qquad (7)$$

$$W = \sum_{j=1}^{\eta} w_i(t)$$

In subtractive weight normalization, a similar procedure is followed but the weights are rescaled according to

$$w_i(t) \leftarrow w_i(t) - z \qquad (8)$$

$$z = \frac{W - \sum_{j=1}^{\eta} w_i(t)}{\eta}$$

INFORMATIONAL CONCERNS

There are other ways to normalize the synaptic weights, but they all require that each synapse onto the neuron have access to the weight of every other synapse on the neuron. Current experimental evidence on synaptic plasticity suggests that local postsynaptic events are sufficient to change the efficacy of a synapse. If true, then how will an inactive synapse communicate its weight so that the total synaptic weight can be computed for the sake of a currently active synapse? Moreover, if one considers the kinds of intracellular signals that could actually communicate the total synaptic strength onto a neuron to every afferent synapse on the neuron, then the scheme runs into the problem of "time-stamping" each total so that delays in collecting the total synaptic weight do not interfere with appropriate normalization. There are ways around these problems; however, if we are to make the assumption that a weight is equivalent to the strength of an individual synaptic connection, then normalization on a cell by cell basis runs into difficulties.

Analysis and simulation of these computational constraints on Hebbian rules has provided insight into some of the virtues and limitations of learning rules that are driven by coincidence detection. Recent evidence concerning one biological mechanism of coincidence detection suggests that the above formulations of a correlational learning rule may have omitted a number of important properties.

Retrograde Communication and Volume Signals in Synaptic Transmission

There is evidence for changes in presynaptic terminals following the induction of long-term potentiation at Schaffer collaterals in area CA1 of the hippocampus (Dolphin et al. 1982; Bekkers and Stevens 1990; Malinow and Tsien 1990). Because the trigger for LTP induction is postsynaptic, there must be some mechanism that permits retrograde communication back to the overlying presynaptic terminal (Bliss and Collingridge 1993). One candidate mechanism for this retrograde communication is the production, diffusion, and action of nitric oxide in response to glutamatergic activity.

The membrane permeant gas nitric oxide (NO), produced in the vertebrate central nervous system subsequent to NMDA receptor stimulation, can influence both synaptic plasticity (Bohme et al. 1991; Haley et al. 1991; O'Dell et al. 1991; Schuman and Madison 1991) and transmission (O'Dell et al. 1991; Friedlander et al. 1992; Manzoni et al. 1992; Montague et al. 1992, 1994a; Kato et al. 1993). Recent experiments have suggested an analogous role for the membrane permeant gas carbon monoxide (CO) (Stevens and Wang 1993; Zhuo et al. 1993a,b).

In 1991, four groups demonstrated that inhibition of the synthetic enzyme for NO, nitric oxide synthase (NOS), blocks the induction of NMDA- dependent LTP in the mammalian hippocampus (Bohme et al. 1991; Haley et al. 1991; O'Dell et al. 1991; Schuman and Madison 1991). Although the meaning of these results is not undisputed (Williams et al. 1993), the findings suggest that NO production is one necessary step in LTP induction. This possibility is strengthened significantly by evidence demonstrating that the correlation of presynaptic activity and elevated levels of NO is sufficient to potentiate transmission only at recently active axonal terminals (Arancio et al. 1993; Zhuo et al. 1993a,b). In these experiments the inactive terminals are unaffected. The amount of potentiation that results from this pairing of activity and NO cannot be enhanced further by a tetanizing stimulus that is known to induce LTP (Zhuo et al. 1993a).

Taken together, the above results suggest that NO may play the role of *conjunction detector* for axonal terminals that are active coincident with high levels of the gas. NO is a nonpolar gas that moves readily through cell membranes. This physicochemical property suggests that NO would not be restricted to its site of production but could move rapidly throughout a surrounding local volume of neural tissue. The interior of any nearby synapse can therefore feel the effects of the local NO concentration whether or not its own presynaptic or postsynaptic elements have been active. This is a critical possibility because in the absence of specialized compartmentalization mechanisms for NO, this signal would act throughout a local volume of neural tissue (Fig. 2). Synaptic plasticity will operate in a local diffusion-defined domain because NO will simply accumulate in local regions containing multiple NO sources.

Volume Learning

A synaptic plasticity mechanism employing NO or some other diffusible substance for the modulation of synaptic function would permit plasticity to operate beyond the boundaries of conventional anatomically defined synapses to influence plasticity throughout a local volume of neural tissue. This novel form of plasticity has been termed *volume learning* because synaptic plasticity operates through a transient diffusion-defined
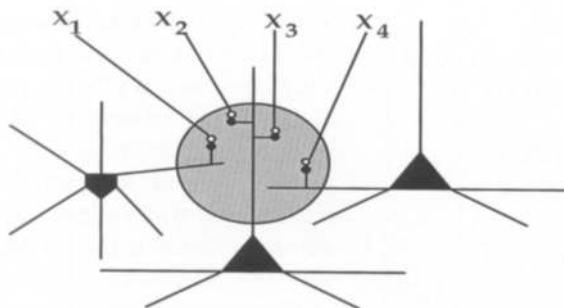
Figure 2: Volume learning. A short-lived membrane permeant substance is released at active synapses. The substance diffuses away from its sources to define a local domain (shaded zone). Temporal fluctuations in presynaptic firing patterns on inputs x_i are thereby transformed into spatiotemporal fluctuations in substance concentration. Specificity in learning is achieved by postulating that synapses that are active when the concentration of this compound is high are strengthened. Synapses that fire when the concentration of substance is low are weakened. The strength of a given synapse is influenced by the firing patterns and three-dimensional locations of neighboring synapses whether or not these synapses are made onto the same neuron. Associations among inputs x_i can develop in small volumes of tissue. An influence on synaptic weights in the local volume will occur if the substance transiently influences synaptic gain or is a direct and necessary signal for longer term synaptic plasticity. This mechanism would interact with other signals known to influence synaptic interactions on neurons through the postsynaptic compartment, e.g., changes in membrane voltage and second messenger production. The volume signal would need to reflect the local activity of glutamatergic synapses; however, as long as the production occurred subsequent to presynaptic activity, the substance could be produced locally from the *dendrites of other cells* or locally elicited from *other presynaptic terminals*. Note that this mechanism allows postsynaptic responses to be integrated in part in the extracellular space.

domain and permits associations among afferent inputs to form in small volumes of neural tissue (Gally et al. 1990; Montague et al. 1991, 1993; Montague et al. 1993b) (Fig. 4, below).

The case for NO as a membrane permeant substance involved in synaptic plasticity is strengthened by other experimental observations and computational work. For example, the synthetic enzyme NOS is calcium dependent (Bredt and Snyder 1992), and assays for NO production show that it is made in both the cortex and hippocampus subsequent to stimulation of NMDA receptors (Friedlander et al. 1992; Montague et al. 1994a). Computational work has shown that a model of synaptic plasticity that uses the covariance of presynaptic activity and a postsynaptically produced rapidly diffusible signal can account for the development of sensory mappings in the cerebral cortex and thalamus (Montague et al. 1991).

CO was mentioned earlier as another possible membrane permeant signaling molecule. CO is a very stable molecule and would enjoy a much longer biological half-life than NO; therefore, an associative mechanism utilizing CO would act over longer time scales than NO. Additionally, CO is not known to be made in response to NMDA stimulation and its use as a signaling molecule is energetically unfavorable (Edelman and Gally 1992). Hence, if CO is actually used in some kind of correlational learning rule in the brain, then the rule depends on events different from those known to influence

NMDA-dependent synaptic plasticity. A correlational mechanism using CO would extract more efficiently statistical regularities that exist at longer times scales than those regularities reported through NO production and diffusion.

Computational Consequences of Volume Learning

Independent of the identity of the rapid volume signal, there are a number of consequences that follow for any learning algorithms based on a volume learning mechanism. A simple version of a volume learning rule was proposed previously to account for the activity-dependent development of sensory mappings in the vertebrate thalamus and cortex and one form of synaptic plasticity in the adult state (Gally et al. 1990; Montague et al. 1991):

$$\Delta w(t) = \eta[x(t) - \theta_{\mathrm{pre}}][\mu(\mathbf{r},t) - T_{\mathrm{pre}}] \qquad (9)$$

where $\Delta w(t)$ is the change in the "weight" or synaptic efficacy of a connection, η is a constant controlling rate of change of synaptic efficacy, $x(t)$ is a measure of presynaptic activity, and θ_{pre} is a threshold that determines whether a terminal is active at time t. T_{pre} is a threshold, dependent on the activity of the presynaptic terminal, which determines the direction of synaptic change. The postsynaptic factor, formerly y in equation 1, is now dependent on the substance concentration $\mu(\mathbf{r},t)$ at time t located at position $\mathbf{r}$.

The substance concentration evolves in time and space according to

$$\frac{\partial \mu(\mathbf{r},t)}{\partial t} = \nabla^2 \mu(\mathbf{r},t) - \kappa\mu(\mathbf{r},t) + p(\mathbf{r},t) \qquad (10)$$

The first term on the right side of this equation governs the diffusion of the substance, whereas the second and third terms represent the sinks and sources of the substance. The constant κ controls the rate of exponential decay, whereas the production rate of substance $[p(\mathbf{r},t)]$ at location $\mathbf{r}$ and time t depends on the synaptic weights and patterns of input activity in the vicinity of $\mathbf{r}$. Large-scale computer simulations have demonstrated that a learning rule like equation 9, when acting against a background of axonal growth, can account for the self-organization of whisker barrels, the refinement of topographic mappings, formation of reciprocal connectivity between cortical regions sharing correlated input, and the formation of ocular dominance columns (Montague et al. 1991).

SPECIFICITY IN VOLUME LEARNING

Specificity in synaptic modification in a volume learning rule is maintained by postulating differential effects on active and inactive presynaptic terminals (Gally et al. 1990; Montague et al. 1991). This theoretical requirement is supported by experimental evidence showing that NO differentially affects active and inactive presynaptic terminals through activation of soluble guanylate cyclase (Zhuo et al. 1993a,b). Also, a form of "synaptic recruitment" has been observed in tissue slices from mammalian hippocampus and cerebral cortex, where potentiation of synaptic contacts spreads from the site of induction only to active presynaptic terminals throughout a local region (Kossel et al. 1990; Schuman and Madison 1991; also see Bonhoeffer et al. 1989). This latter physiological effect is consistent with the production of a diffusible

volume signal that potentiates only active synaptic contacts and does not potentiate or depress inactive contacts.

STABILITY IN VOLUME LEARNING

Using equation 9, weights can be prevented from growing without bound by clipping as described previously or through an appropriate balance of competing activity patterns. This kind of volume learning rule can be overwhelmed by overactive inputs because the threshold T_{pre} does not change with activity levels in the vicinity. Hence, this rule is reasonably thought of as a *set point model* where the thresholds for presynaptic activity and substance levels necessary for synaptic change do not adjust to the ambient conditions. It is likely that this aspect of this volume learning rule is biologically incorrect and a better way to express this learning rule is in its covariance version

$$\Delta w = \eta[x(t) - \theta_{pre}][\mu(\mathbf{r},t) - \overline{\mu}(\mathbf{r},t)] \qquad (11)$$

Here, $\overline{\mu}(\mathbf{r},t)$ represents a running average of $\mu(\mathbf{r},t)$ in the vicinity of the synapse. Equation 11 will adapt to ambient levels of the substance and thus adapt the learning mechanism to the average amount of activity in a region. This volume learning rule would produce qualitative changes in synaptic strengths according to Table 1.

A REPRESENTATIONAL ISSUE: HOW MAPPINGS FUNCTION UNDER RAPID VOLUME EFFECTS

If the interaction among presynaptic terminals is mediated in part by a transient diffusible signal produced by nearby active synapses, then lateral interactions can take place throughout a diffusion-defined domain. The size and shape of such a domain will depend on patterns of neural activity impinging on a region as well as physical factors such as the geometrical arrangement of synapses in three dimensions, the synthetic and catabolic rates for the substance, and the potential action of other unidentified barriers or sinks for the substance. The numerous possible implications of a variable distance over which synapses interact are not known. Although it has been demonstrated that a volume learning rule can direct the appropriate activity-dependent map formation (Montague et al. 1991, 1993a,b), the implications of a variable lateral interaction scale are crucial and unexplored in detail.

The robustness of a mechanism that permits associations between afferent inputs to develop in small volumes of tissue is not known. Rapid

Table 1: *Volume learning rule*

	$\mu(\mathbf{r},t) - \overline{\mu}(\mathbf{r},t) > T$	$\mu(\mathbf{r},t) - \overline{\mu}(\mathbf{r},t) < -T$
Active synaptic element	increase	decrease
Inactive synaptic element	decrease	no change

This contingency table shows the direction of change of synaptic weights under a volume learning rule. If nitric oxide is the volume signal, then experimental data suggest that the learning rule in the lower left portion of the table is incorrect. This rule decreases synaptic weights of inactive connections that experience suprathreshold levels of nitric oxide. Using these learning rules, an active synaptic contact must correlate with local activity or its weight will be decreased to 0. This latter condition arises because two depression rules are included in the table: one for depressing inactive terminals and one for depressing active terminals. These rules force cooperation throughout a local diffusion-defined domain because they do not allow a terminal to be appropriately silent. In this fashion, a local volume of tissue could only map information about one variable.

volume signals place a premium on the three-dimensional distribution of synaptic contacts throughout a local volume because these contacts represent potential sources of substance in response to neural activity. An emphasis is therefore implicitly placed on the dendritic morphology of the recipient cells in a region. This latter point gains in importance because NO can also influence local blood flow (Iadecola 1993) and neurotransmitter release (O'Dell et al. 1991; Zhuo et al. 1993a,b; Montague et al. 1994a). It has long been recognized that electrotonic and second messenger communication within branched dendritic structures is critically dependent on specific dendritic morphology. Rapid volume effects add an extra dimension to these interactions and may represent another means through which specific dendritic structure performs important and identifiable computational functions.

In Figure 3 we provide a simple example of how rapid volume signals can influence the manner in which information is represented in a cortical map in which volume signals operate. In this example we assume for simplicity that pyramidal cells are the only recipient cell type in the region. In Figure 3A, input 1 and input 2 are located one space constant from the soma. If we consider local signals confined to the dendrites, then these inputs will interact primarily through depolarization at the level of the soma. This assumes that local signals in the dendrites do not interact and that the voltage changes are integrated primarily at the soma. Changing the orientation of these dendrites, as shown in Figure 3B, does not change these statements; that is, for signals confined to the postsynaptic compartment, the notion of proximity is not changed by reorienting the dendrite. In the case where an active input (input 1) can elicit a rapid volume signal (shaded circular zone), the position of the dendrite is critical. Because of the capacity for synapses to interact directly through the tissue, proximity through the tissue space does not correspond to proximity along dendrites. Accordingly, the exact three-dimensional distribution of synapses throughout a region determines the nature of the feed-forward mapping into the region from the point of view of the volume signals. Hence, the distribution of feed-forward synapses can be sampled using different dendritic structures, so that a given mapping can simultaneously represent several different transformations depending on the dendritic structure of the recipient neurons.

From Temporal Coincidence to Temporal Order—The Goal of Prediction

Hebbian learning rules are correlational in the sense that the changes in synaptic strengths represent the associations between inputs. There are other relationships that are important to learn about events, such as the temporal order of the inputs. Hebbian learning rules are symmetric in time: They are sensitive only to the degree of temporal coincidence of inputs and not the temporal order (Fig. 4). For example, a Hebbian rule is not sensitive to whether input A follows input B; rather, it is sensitive only to the absolute separation in time of inputs A and B and is therefore symmetric in time. This lack of sensitivity to temporal order means that alone, a Hebbian rule would not permit the development of the predictive relationships that occur between stimuli during a classical or instrumental conditioning task (MacKintosh 1974, 1983). Conditioning experiments have shown that through learning, sensory stimuli can come to act as predictors of reward, punishment, and other salient stimuli (Dickinson 1980). One important constraint that has emerged from this
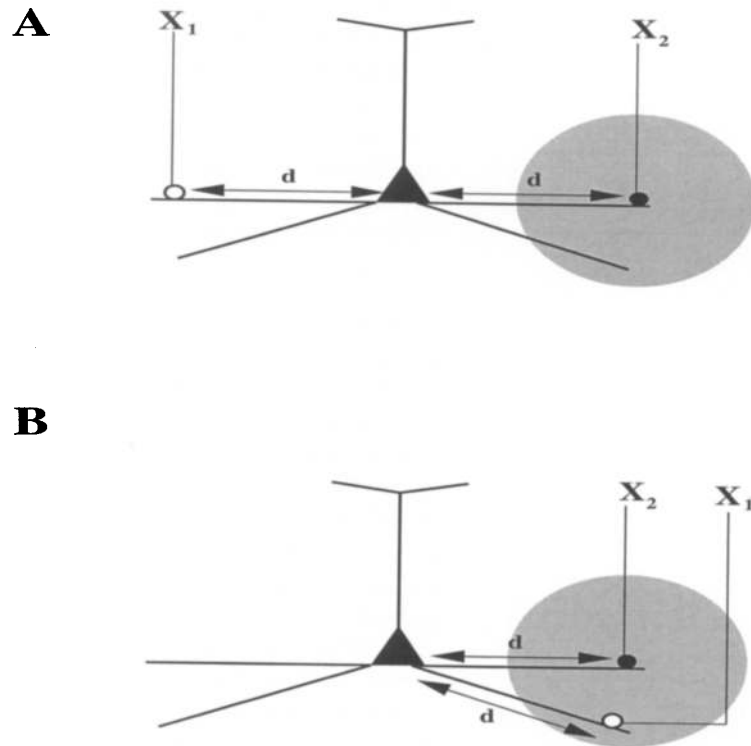
Figure 3: Sampling input activity with volume signals and postsynaptic signals. Simplified example illustrating some differences between signaling directly through a volume of tissue and confining signals to the postsynaptic compartment. (*A,B*) Fibers 1 and 2 represent inputs that participate in a feed-forward mapping into a region of cortex. Input 1 (○) is not firing, whereas input 2 (●) is firing and producing a modulatory substance. These two inputs are assumed to be the same electrotonic distance *d* from the soma. This illustration shows how a volume signal would influence the representation of information in this mapping. From the point of view of signals restricted to the dendritic compartment, the proximity of these synaptic inputs is not changed by simply moving the dendrite or repositioning one input at an equivalent electrotonic distance on another dendrite (*B*). Because rapid volume signals confer the capacity for synapses to interact directly through the tissue, proximity through the tissue space does not correspond to proximity along dendrites. (*B*) The repositioned input 1 is influenced by the volume signal elicited by input 2 although the capacity for interactions strictly through the dendritic compartment has not changed. Clearly, the fact that the signal can pass readily through membranes is important for understanding its function because the extracellular space may be very small in any given region of neuropil. A compound like NO (diffusion constant $= 2.60 \; 10^{-5} \; cm^2/sec$) would diffuse a root mean square distance of 10 μm in about 6.4 msec. This calculation for the root mean square distance NO diffusion assumes an unbiased random walk in three dimensions. In one dimension, the average squared distance traversed by a diffusing molecule with diffusion coefficient D is $<d^2> = 2Dt$, where d is distance, t is time, D is the diffusion constant. In three dimensions we have $<r^2> = <x^2 + y^2 + z^2> = <x^2> + <y^2> + <z^2> = 2Dt + 2Dt + 2Dt = 6Dt$, where r is the vector (x,y,z). In the mammalian cerebral cortex, synaptic densities have been estimated to be ~1 billion synapses/mm³, which yields >4000 synapses in a sphere of radius 10 μm. A volume signal using a low-molecular-weight membrane permeant molecule can allow a small number of synaptic contacts to tell many other synaptic contacts in a region about the statistics of their activity. This form of local "broadcast" does not require the use of axons.
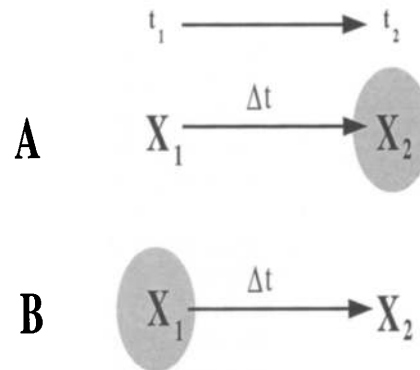
Figure 4: Detection of temporal order through a suppression of plasticity. x_1 and x_2 represent two inputs into a common region or onto a common cell (see Fig. 3). A correlational rule is sensitive only to the absolute separation in time ($\Delta t = t_2 - t_1$) of x_1 and x_2. (A) x_1 occurs before x_2; (B) their order is reversed. The two cases are equivalent for a correlational rule. Suppose that input x_2 elicits some signal that is necessary for learning but that also inactivates learning for some short period of time when it rises above a threshold (shaded area). The order of occurrence of x_1 and x_2 now becomes important. x_1 must be active prior to x_2 or else no changes in synaptic weights can occur. This simple scheme represents one of many possible ways to express an unsupervised local predictive learning rule (see text).

work is that this form of learning is asymmetric in time; that is, sensory events consistently preceding presentation of rewarding stimuli come to act as predictors of the reward, whereas sensory events following the presentation of reward do not come to act as predictors of the reward. Although these kinds of experiments assay the behavior of the entire animal, they highlight the importance of the causal structure of the world for the learning displayed by the animal.

Any system that uses predictions of its most likely next state and the most likely next state of the world has information to prepare itself for the future given its current inputs and plans for taking action. In this sense, prediction can be viewed as a *computational goal* of a system that must operate in an uncertain and variable environment. For example, a system that could predict how the sensory input would change as a consequence of making a movement would be of great benefit in planning actions (Jordan and Rumelhart 1992). A prediction can be compared with the actual changes following a movement and the error used to improve the prediction. A similar approach can be taken to predicting the locations of targets for eye movements. There are probably a variety of predictive systems in the brain. Given that prediction is an important goal, it is reasonable to expect the existence of predictive mechanisms in vertebrate nervous systems at many spatial and temporal scales.

Predictive Learning Rules—Learning Driven by Temporal Order

Previous theoretical and modeling work has focused on the need for predictive or anticipatory mechanisms for learning to explain animal behavior and reinforcement systems (Rescorla and Wagner 1972; Sutton and Barto 1981; Klopf 1982). These kinds of models have used neuron-like units and adaptive weights to reproduce various aspects of classically and instrumentally conditioned behaviors. Other work has carefully considered the neurobiological aspects of predictive models in the context of classical conditioning (Hawkins and Kandel 1984; Moore

et al. 1986; Gluck and Thompson 1987). In all of this work, some sensitivity to the temporal order of inputs is built into the structure of the network and the rules for modifying synaptic strengths.

Because predictive mechanisms are probably represented in the vertebrate brain at many spatial and temporal scales, it is reasonable to inquire about the nature of these mechanisms at both large and small scales. In the following sections we present arguments and modeling approaches that address predictive mechanisms in the vertebrate brain at the scale of single synapses and at the scale of global signals available to widespread recipient regions. We first review data from the behavior of glutamatergic synapses suggesting that a predictive learning rule may take place at single glutamatergic connections. In the succeeding section we review evidence for predictive mechanisms at global levels of processing.

A Local Predictive Learning Rule—Could Synaptic Change at Glutamatergic Synapses be Predictive?

We focus here on data related to glutamatergic transmission in the mammalian hippocampus. In a hippocampal slice preparation, Malenka and colleagues (Huang et al. 1992) have shown that weak synaptic activity (30 Hz, 0.25 sec) along a synaptic pathway will transiently block the subsequent capacity to induce LTP along that pathway. This result has been confirmed by Zorumski and colleagues (Izumi et al. 1992). The latter group have also extended these findings and have shown that this block of LTP induction is itself blocked by agents that inhibit NO production or chelate NO in the extracellular space. These results suggest an important theoretical possibility, provide important mechanistic clues, and construct links to the volume learning framework outlined in the preceding discussion. We first consider additional data pertinent to these findings and subsequently present the computational consequences of a mechanism that allows prior synaptic activity to block the subsequent capacity to modulate synaptic strength.

NO is known to be involved in a variety of feedback mechanisms at different spatiotemporal scales: (1) NO inhibits NOS activity directly (Klatt et al. 1992); (2) NO greatly diminishes calcium fluxes through the NMDA receptor (Izumi et al. 1992; Manzoni et al. 1992) through action at a site distinct from the glutamate-binding site (Lei et al. 1992; Lipton et al. 1993); and (3) by coupling glutamatergic activity to blood flow changes, NO production causes the influx of oxyhemoglobin into regions of neural tissue that were active previously and therefore have diminished oxygen tension. The oxyhemoglobin would tend to lose its oxygen and would be available to chelate free NO.

These forms of negative feedback that limit NO production in a region of neural tissue might have been expected simply on the grounds that because NO has so many potential influences, it must be tightly controlled. Whereas no definitive conclusions can be reached with these experimental results, they do suggest that there may be a substrate for negative feedback onto the NMDA receptor following events that are sufficient to cause calcium fluxes through the receptor. As explained below, this biological possibility has important computational consequences.

One important possibility is that this inhibition of the NMDA receptor may represent a blockade of the subsequent ability of synapses in the local volume to change their strengths, that is, inhibition of plasticity following events sufficient to engender long-term changes in synaptic strengths. Although the time scales are not quite appropriate, this

interpretation is supported by the above data from glutamatergic synapses in the hippocampus (Huang et al. 1992; Izumi et al. 1992). In situations where NO production participates in synaptic plasticity, the negative feedback to NOS has similar consequences.

Independent of the exact role played by NO production, these data collectively suggest that previously active synaptic contacts may be the only eligible candidates for potentiation, that is, only those pre- and postsynaptic elements with a decaying trace of activity (e.g., calcium) would be eligible for long-term modification. We extend this possibility to include both increases and decreases in synaptic strength. The simple invalidation of all local connections at the moment that an elicited postsynaptic signal is sufficiently high forces the system to "look backward" in time for some trace of activity at the synapses that were active previously. Hence, we have a substrate for a plasticity mechanism that is sensitive to the temporal order of inputs and may operate at single glutamatergic connections in an unsupervised manner (Fig. 4).

This mechanism may be viewed as a **local predictive learning rule**: *Those synaptic elements (pre- or post-) whose activity consistently precedes (predicts) epochs of sufficiently synchronous activity in a local volume of tissue will be potentiated. Those synaptic elements whose activity precedes (predicts) epochs of little or no synchronous activity in a locale are depressed.*

$$\Delta w(t) = \eta x(t - b)[\mu(\mathbf{r},t) - \overline{\mu}(\mathbf{r},t)] \tag{12}$$

This possibility provides a substrate for an unsupervised learning rule that can act predictively. As before, suprathreshold fluctuations in $\mu(\mathbf{r},t) - \overline{\mu}(\mathbf{r},t)$ are necessary for changes in synaptic strength; however, now only prior activity is relevant. In this formulation, b represents a fixed time interval so that equation 12 changes weights by comparing preceding presynaptic activity $x(t-b)$ with current postsynaptic responses $\mu(\mathbf{r},t) - \overline{\mu}(\mathbf{r},t)$. We have cast this rule as a volume learning rule primarily because of convenience: It allows a signal to be passed rapidly throughout a local domain. The signal would not have to be a rapidly diffusible signal but could represent fluctuations of any kind of elicited postsynaptic response made available to presynaptic terminals. *The weights that develop under such a learning rule act as predictions of future epochs of correlated activity in the local region.* In the succeeding section we show how a second predictive learning rule acting within a neural system can produce synaptic weights that represent predictions of future reinforcement.

One point to emphasize about this learning rule is that there must be some time delay (b) between the onset of activity in the presynaptic terminal and the time that the fluctuation in substance levels actually selects the direction and magnitude of the weight changes. A threshold for the size of fluctuations $\mu(\mathbf{r},t) - \overline{\mu}(\mathbf{r},t)$ is one way to enforce such a delay. This delay would then depend on how fast both μ and $\overline{\mu}$ can change. Note that allowing rapid changes in $\overline{\mu}$ could prevent learning altogether. We mention these possibilities to emphasize that the time constant associated $\overline{\mu}$ is a critical parameter and might actually be adaptable (Fig. 5).

There are other differences between this predictive learning rule and the covariance rule presented in Table 1.
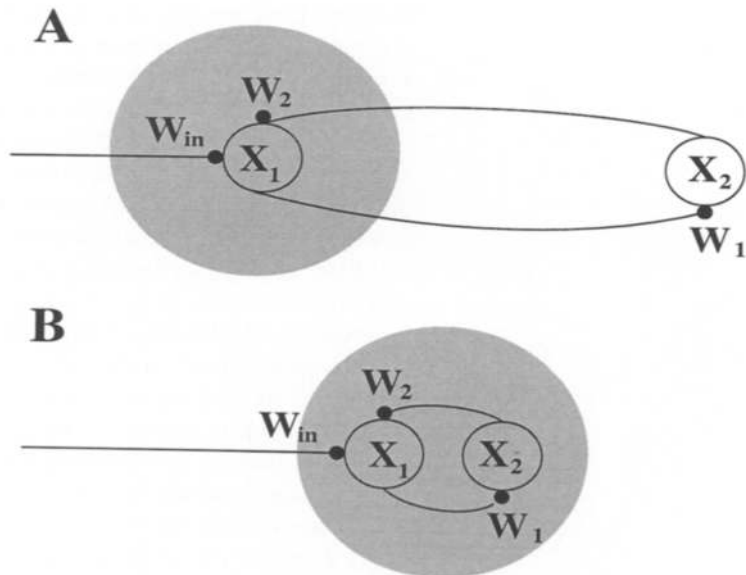
Figure 5: Two scenarios for a local predictive rule. Two neurons, 1 and 2, are connected to one another through connection weights W_1 and W_2, respectively. Additional input to neuron 1 is provided through connection weight W_{in}. When neuron 2 fires, it elicits a local signal (shaded area) that inactivates learning (long-term synaptic weight changes) for a short period. The nature of this local signal is left unspecified. Only those synaptic contacts that were active just prior to the production of this signal are eligible for weight changes. As discussed in the text, this scheme forces the system to "look back" in time for synaptic contacts eligible for long-term weight changes. The learning would be predictive because only those synaptic contacts whose activity consistently precedes activity in the terminal from neuron 2 can be potentiated. A synaptic contact whose activity is not followed by the signal elicited from 2 would be depressed (see Table 1). In this illustration, we assume that neuron 1 can elicit this signal because its weight is sufficiently large (for details, see text). (A) Neuron 2 is separated from neuron 1 by a distance sufficiently large that the signal (shaded area) only influences the input to 1. If this input precedes activity in 2, then W_{in} could be potentiated. (B) The input connections to 1 (W_{in}) as well as the output connection from 1 (W_1) are subject to the signal elicited by 2. In this case, both W_{in} and W_1 are subject to increase if activity in these connections precedes the signal elicited by W_2 (shaded area); otherwise, their activity will cause them to depress if this activity is not followed by activity in 2.

An appropriately silent terminal can remain in a region because potentiation and depression only occur for active terminals; that is, a terminal must be active to "assay" whether it correlates with activity in a local domain, thus allowing more than one variable to be mapped in a single local domain. Note also that this rule does not potentiate synapses that become active after epochs of correlated activity, possibly suggesting that modification of single glutamatergic connections is sensitive to temporally ordered events (Huang et al. 1992; Izumi et al. 1992).

INFLUENCE OF LOCAL PREDICTIVE LEARNING RULES ON DEVELOPING SYSTEMS

One reason that developmental experiments may appear to be consistent with Hebbian mechanisms of plasticity is that they represent long-term time averages. These could mask the fact that the terminals that segregate together do so because they are predicting the same local events through time. For example, consider the case of two axonal terminals, A and B, in

a common region of neural tissue. Suppose that (1) activity in A precedes the activity in B, (2) activity in B precedes the activity in A, or (3) both 1 and 2 hold. One would label any of these three conditions as supporting the statement that A and B were correlated in their activity. Under appropriate conditions, a synaptic learning rule sensitive to temporal order could strengthen or stabilize terminal A when its activity was followed by activity in B and similarly strengthen terminal B when its activity was followed by activity in A. In this fashion, although the learning rule is strictly predictive, the stabilization or strengthening of synaptic contacts over longer developmental epochs would thus appear to be consistent with a correlational rule acting at the individual synaptic contacts.

During activity-dependent phases of neural development, this local predictive rule also would tend to incorporate the causal structure of the inputs even at the level of single receptive fields: One set of inputs could predict future activity in another set of inputs onto a single neuron or group of neurons. For a visually responsive neuron, this would allow input in one region of the receptive field to predict activity in another portion of the receptive field. For a cell receiving inputs from multiple modalities, one modality could predict the future onset of inputs carrying information from another modality (Montague et al. 1993a; Pouget et al. 1993).

Predictive Learning through Diffuse Ascending Systems

Thus far we have considered only *unsupervised learning rules,* but there are numerous other signals important for adaptive behavior and learning that could be considered *supervised*, if arising from outside the brain, or *monitored*, if the reinforcement signal is internally generated (Churchland and Sejnowski 1992). For example, attentional and motivational states (Mountcastle et al. 1981; Cole and Robbins 1992) and rewards (Wise 1982) are all important components of learning and memory. Information about these kinds of influences is transmitted to target structures in part through the diffuse ascending systems of axons originating in small nuclei in the midbrain and basal forebrain (e.g., Cooper et al. 1970). The axons of these nuclei innervate large expanses of the cortical mantle and other structures and deliver to their targets various neurotransmitters including dopamine, norepinephrine, serotonin, and acetylcholine. Invertebrates have analogous neurons that have been shown to be involved in reinforcement and reward processing (Hammer 1991).

Behavioral and physiological work has shown that these diffuse systems can influence ongoing neural activity (Kaczmarek and Levitan 1987; Foote et al. 1991), memory (Damasio et al. 1985; Tranel and Damasio 1985; Goldman-Rakic et al. 1990; Decker and McGaugh 1991), and action choice (e.g., Bernheimer et al. 1973). Some of the same diffuse systems are also known to be required for the normal development of the response properties of cerebral cortical neurons. For example, removal of acetylcholine and norepinephrine disrupts normal ocular dominance plasticity (Bear and Singer 1986; Rauschecker 1991) and dramatically alters the rules for dendritic development in the somatotopic maps of rat cerebral cortex (Loeb et al. 1987). The question naturally arises as to how and why the same signals are used during development, learning, and behavioral control.

PERMISSIVE MODELS

There are a number of models of reinforcement learning that explicitly or implicitly appeal to one of the diffuse systems to deliver information about rewards to target structures. In a number of these models, the output of the diffuse system is used as a gating signal that defines epochs during which correlational learning can occur (Hawkins and Kandel 1984; Gluck and Thompson 1987; Rauschecker 1991). These models can be expressed as

$$\Delta w(t) = \eta x(t)y(t)r(t) \tag{13}$$

where $r(t)$ represents the output of a diffuse system that reports on reward to its target structures. This particular formulation of permissive gating would allow a system to backward-condition if r simply reported the *occurrence* and *magnitude* of a rewarding stimulus (Fig. 6). In general, backward conditioning does not occur; therefore, this rule is an incomplete description of how information about reinforcement should influence synaptic change.

Interestingly, experimental data support a coincidence rule for the influence that the diffuse system outputs exert on plasticity. If the temporal order effects are to be taken into account at the synaptic level, then the gating model in equation 13 can be sensibly modified by a more detailed consideration of the nature of the information that a diffuse system could be expected to deliver to its target structures. It is possible that these diffuse systems are not simply reporting the occurrence and
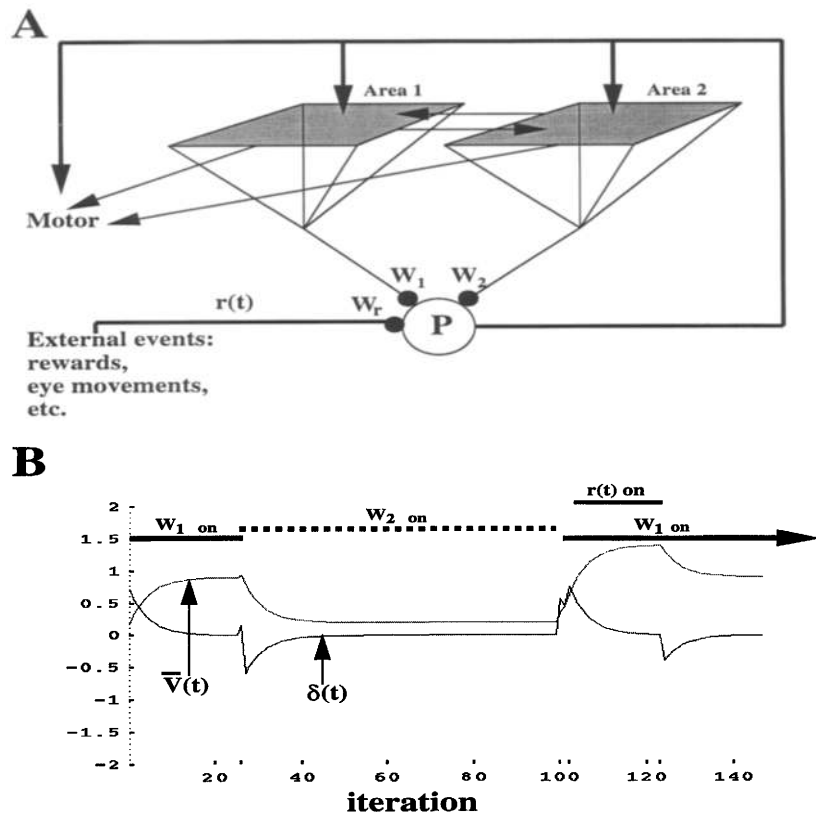

Figure 6: (*See facing page for legend.*)

magnitude of rewarding and salient events in the world but, rather, information about predicted rewards and events.

PREDICTIVE MODELS: MAKING PREDICTIONS IN THE REAL BRAIN

Recent research on computational models has suggested that the diffuse ascending systems could be reporting information about predictions of future stimuli simultaneously to widespread targets (Quartz et al. 1992;

Figure 6: Making and using scalar predictions in a real brain. (A) Making predictions with distributed representations. An activation pattern in one of the illustrated cortical areas influences the (scalar) output of P through an effective weight, e.g., W_1 from area 1. This weight is termed effective because the high convergence from area 1 onto P throws away topographically coded information forcing the input to P from area 1 to represent a scalar drive to P. In this sense, the anatomical convergence alone performs an important computational function converting a pattern of activity into a scalar drive to P. If an activation pattern in area 1 consistently preceded an eye movement reported along $r(t)$, then the output of P would reflect approximately the temporal difference between the drive to P through W_1 and the drive to P through the W_r. The more these two inputs to P differ, the larger the change in P's output. Similar values along these two inputs cause little or no change in the output of P. Because P responds to differences in input through time, its output represents a comparison *through time* of W_1 and W_r. This comparison permits the weight W_1 to act as a *prediction* of the effective drive that will be delivered along $r(t)$ in the future. P's output will reflect a crude *error* in that prediction, i.e., the degree to which there is a mismatch between input through W_1 and future input along $r(t)$. A simple means to modify the weights W_1 and W_2 so that they become better predictors of future external events is to change them according to a simple correlational rule (see text). Using predictions from distributed representations, B illustrates how $\delta(t)$ and $\overline{V}(t)$ change with changing patterns of activation in the two converging areas (area 1 and area 2). For this example, pattern 1 is restricted to area 1 and influences the output of neuron P through its effective weight $W_1 = 0.9$. Pattern 2 is similarly defined and influences P through an effective weight $W_2 = 0.2$. There is also a reward pathway with weight equal to 0.5. We assume that these weights have already been set by some learning process as described in the text. This example considers how switching between two distributed patterns influences the output of P. The times during which each pattern or the reward pathway is active are indicated: W_1 **on** (solid lines) ≡ pattern 1 active, W_2 **on** (broken line) ≡ pattern 2 active, and $r(t)$ **on** ≡ reward on. Each input pathway is off unless otherwise indicated. Each input takes the value 1 when it is active and 0 when it is inactive. At cycle 0, pattern 1 is active and remains so until cycle 26. Note that during this period that $\delta(t)$ decays to 0 while $\overline{V}(t)$ builds to 0.9 (the value of the weight W_1). At cycle 26, pattern 2 becomes active and causes a slight increase in $\Delta(t)$. At cycle 27, pattern 1 becomes inactive and pattern 2 remains active, thus completing the switch from pattern 1 to pattern 2. This switch is accompanied by a large negative deflection in $\Delta(t)$. In this example, we assume that decrease in the output of P delivers less neuromodulator to its targets. Under this assumption, switching from pattern 1 to pattern 2 makes it less likely that the output connections from neurons participating in pattern 2 will cause their target cells to fire. This output could be connections to motor output or to other cortical areas (see A, above). Hence, the weights W_1 and W_2 store information related to predictions made by the inputs to neuron P. These weights subsequently influence learning through their influence on the output of neuron P. As pattern 2 remains active, $\delta(t)$ decays back to 0 and $\overline{V}(t)$ decreases from W_1 (0.9) to its steady-state value W_2 (0.2). At cycle 100, pattern 2 goes off and pattern 1 comes on, resulting in a positive deflection in $\Delta(t)$. This switch would bias the output connections of neurons participating in pattern 2 and permit them to drive their target neurons. The first peak in $\Delta(t)$ is the result of the switch to pattern 1, and the second peak is the result of the onset of the reward pathway $r(t)$. The negative deflection in $\Delta(t)$ at cycle 123 is caused by the offset of the reward pathway. ($\lambda = 0.4$).

Montague et al. 1993a, 1994b; Pouget et al. 1993). These biological models are related to more abstract predictive models that have been developed and applied to engineering control problems and data from classical conditioning experiments (Sutton and Barto 1981, 1987, 1989; Moore et al. 1986; Sutton 1988).

The proposal for how predictions could be made in a real brain is outlined in Figure 6 (also see Montague et al. 1993a). In figure 6, a small subcortical nucleus, labeled P, receives highly convergent input from both cortical representations and inputs carrying information about rewarding or otherwise salient events in the world. The cells in this nucleus are assumed to respond transiently to their net excitatory input, that is, compute the change in activity, or an approximation to the temporal derivative. For this discussion we will represent this derivative as a difference between the ongoing activity and a running average of the activity

$$\delta(t) = V(t) - \overline{V}(t) \tag{14}$$

$$V(t) = \sum_{j=1}^{\eta} w_j x_j + r(t)$$

where $V(t)$ is the net input to the cell at time t, including the unconditional reward stimulus $r(t)$, and $\overline{V}(t)$ is a running average represented by:

$$\overline{V}(t) = \lambda V(t) + (1 - \lambda)\overline{V}(t - 1) \text{ for } 0 < \lambda < 1 \tag{15}$$

where λ is a constant that determines the distance into the past over which the activity is averaged. As λ approaches 0, the average reaches farther into the past. As λ approaches 1, the averaging interval becomes short, and the output of P closely approximates the net input $V(t)$ at time t. Combining equations 14 and 15, we have

$$\delta(t) = (1 - \lambda)[V(t) - \overline{V}(t - 1)] \tag{16}$$

The output of P *reflects a (scaled) temporal difference between the current net input and the previous running average of the net input.*
Plasticity of the weights from the cortex onto P and within the cortical layers is assumed to follow a simple correlational rule

$$\Delta w(t) = \eta x(t)\delta(t) \tag{17}$$

This rule thus retains a "gating" influence of the output of the diffuse system; however, the output $\delta(t)$ is not simply the magnitude of the rewarding stimulus but instead represents a particular comparison of the net input through time. The net weight converging onto P from the cortex, for example, W_1 in Figure 6A, can act as a prediction of the amount of reward $r(t)$.

When a rewarding stimulus is first encountered, it increases the output of P because at that moment, the output $\delta(t)$ will be proportional to $r(t) + \overline{V^*}(t-1)$, where $V^*(t)$ is the total net input to P not contributed by $r(t)$. As the actual delivery of information about the reward rises and falls back to baseline, the running average $\overline{V}(t)$ will follow slowly (Fig. 6B). During learning, the weights are changed according to equation 17 until the running average of the input $\overline{V}(t)$ from the cortex correctly predicts delivery of the reinforcement, so that $\delta(t) = 0$. The capacity for $\overline{V}(t)$ to do this clearly depends on a number of factors including the value of λ and the nature and time course of the rewarding events. Following learning, the output of P remains at zero (or at a constant level of firing if there is a spontaneous level of background activity); any change in the output of P is then a measure of the unexpected reward and represents a failure in the prediction of future reward. Alternatively, the output of P can be used to predict the likely reward value of a novel sensory stimulus based on previous learning with similar stimuli.

In the learning rule expressed by equation 17, the output of P chooses the direction of learning at its targets because it is a signed quantity. The sign is interpreted as an increase in weight if $V(t) - \overline{V}(t) - 1 > T$ (increased neuromodulator release at the target) and a decrease in weight if $V(t) - \overline{V}(t) - 1 > -T$ (decreased neuromodulator release at the target) for some threshold T. In this sense, the weights onto P act as *predictions of the (scalar) value of $r(t)$* and the output of P could be said to represent a prediction error. The simple correlational rule that acts at the targets of the diffuse axons thus permits these predictions errors to drive learning. There are a number of subtleties in this formulation that have been discussed in an engineering context (Sutton and Barto 1981; Sutton et al. 1987). Although the foregoing discussion omits the subtleties of how "on-line" predictions can be used in a learning rule, it does give the general character of the information that diffuse systems of axons can deliver to their targets.

A MECHANISTIC CONSTRAINT SUGGESTED BY ANIMAL CONDITIONING DATA

As formulated, the above scenario for making predictions about future reinforcement would permit the reinforcement pathway $r(t)$ to predict the onset of input from the cortical areas; this is, this scheme would permit backward conditioning under some circumstances. Under almost all normal circumstances, animals will not backward-condition. Consequently, we propose that the onset of the reinforcement pathway inhibits future plasticity for some period of time. This assumption forces the system to use previously active inputs as the predictions of reward in a manner analogous to our previous proposal for single glutamatergic connections. Under this assumption, backward conditioning will not occur because only those events that precede the input along $r(t)$ are eligible for synaptic change. For example, if $r(t)$ was driven by an eye movement, the only synapses eligible for long-term modification are those whose activity consistently precedes the eye movement. This requirement establishes the *eligibility* for synaptic modification and does not alter the construction and use of the predictions outlined above.

The cortical activity patterns that consistently predict the increases of $r(t)$ in the world will thus develop weights onto P that act to discount the influence that the subsequent increase in $r(t)$ has on the output of P. This framework has been applied to activity-dependent development and registration of mappings in the vertebrate nervous system (Quartz et al.

1992; Montague et al. 1993a; Pouget et al. 1993). In this paper information related to eye movements and eye position was used to influence the diffuse system output and thus directed the appropriate development of a sensorimotor mapping (Montague et al. 1993b) and the registration of this map with a head-centered auditory map (Pouget et al. 1993). A similar model has been used to account for decision behaviors in foraging animals (Montague et al. 1994b) and is consistent with human data from lesion patients who apparently lack descending inputs from the frontal cortex to the corresponding subcortical nuclei (Saver and Damasio 1991; A. Damasio and H. Damasio, pers. comm.). This latter result is consistent with physiological studies in behaving rats (Castro-Alamancos and Borrell 1992). This framework has also been proposed to account for the data of Ljunberg et al. (1992), where changes in firing rate observed in dopaminergic neurons in the ventral tegmental area during a conditioning task in awake behaving monkeys is consistent with these neurons reflecting a prediction error (Quartz et al. 1992; Montague et al. 1993a).

Summary

We have outlined some of the virtues and limitations of Hebbian, or correlational, learning rules as mechanisms crucial for appropriate development and learning in the vertebrate brain. We have also reviewed some of the consequences of a synaptic plasticity mechanism that employs a rapidly diffusible membrane permeant signal. The learning rule that results from the latter mechanism can be viewed as a *volume learning rule* because it permits a small volume of neural tissue to act as a single computational unit. Both Hebbian and volume learning rules depend primarily on the absolute separation in time of different inputs to drive changes in synaptic strengths and are otherwise insensitive to the exact temporal order of inputs.

The critical concept of prediction is left out of both Hebbian learning and volume learning. Through an appeal to biological evidence and computational expediency, we have proposed a local predictive learning rule that may act at glutamatergic synapses. This rule is supported by experimental data from the hippocampus and may represent only one of many local predictive mechanisms in the vertebrate brain. Finally, we have discussed how more global signals, like those broadcast through various diffuse ascending systems, can also be used to provide information about predictions to widespread target structures.

The benefits from being able to make rapid and accurate predictions may be sufficiently important that evolution may have favored brain systems that supported predictive learning. Learning mechanisms such as those explored in this paper may provide insights into the overall organization of brain systems.

Acknowledgments

References

Arancio, O., E.R. Kandel, and R.D. Hawkins. 1993. cGMP produces long-lasting presynaptic potentiation in hippocampal cell cultures. *Soc. Neurosci. Abstr.* **19:** 241.

Artola, A. and W. Singer. 1993. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends Neurosci.* **16:** 480–487.

Artola, A., S. Brocher, and W. Singer. 1990. Different voltage-dependent thresholds for inducing long-term depression and long-term potentiation in slices of rat visual cortex "see comments." *Nature* **347:** 69–72.

Bear, M.F. and W. Singer. 1986. Modulation of visual cortical plasticity by acetylcholine and noradrenaline. *Nature* **320:** 172–176.

Bear, M.F., A. Kleinschmidt, Q.A. Gu, and W. Singer. 1990. Disruption of experience-dependent synaptic modifications in striate cortex by infusion of an NMDA receptor antagonist. *J. Neurosci.* **10:**(3) 909–925.

Bekkers, J.M. and C.F. Stevens. 1990. Presynaptic mechanism for long-term potentiation in the hippocampus. *Nature* **346:** 724–729.

Bernheimer, H., W. Birkmayer, O. Hornykiewicz, K. Jellinger, and F. Seitelberger. 1973. Brain dopamine and the syndromes of Parkinson and Huntington, Clinical, morphological and neurochemical correlations. *J. Neurol. Sci.* **20:**(4) 415–455.

Berns, G.S., P. Dayan, and T.J. Sejnowski. 1993. A correlational model for the development of disparity in visual cortex that depends on prenatal and postnatal phases. *Proc. Natl. Acad. Sci.* (in press).

Bienenstock, E.L., L.N. Cooper, and P.W. Munro. 1982. Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *J. Neurosci.* **2:** 32–48.

Bliss, T.V. and G.L. Collingridge. 1993. A synaptic model of memory: Long-term potentiation in the hippocampus. *Nature* **361:** 31–39.

Bliss, T.V.P. and T. Lomo. 1973. Long-lasting potentiation of synaptic transmission in the dentate area of the anesthetized rabbit following stimulation of the perforant path. *J. Physiol.* **232:**(2) 331–356.

Bliss, T.V.P. and M. Lynch. 1988. Long potentiation of synaptic transmission in the hippocampus: Properties and mechanism. In *Long-term potentiation: From biophysics to behavior* (ed. P.W. Landfield and S.A. Deadwyler), pp. 3–72. Liss, New York.

Bohme, G.A., C. Bon, J.M. Stutzmann, A. Doble, and J.C. Blanchard. 1991. Possible involvement of nitric oxide in long-term potentiation. *Eur. J. Pharmacol.* **199:** 379–381.

Bonhoeffer, F. and J. Huff. 1985. Position-dependent properties of retinal axons and their growth cones. *Nature* **315:** 409–410.

Bonhoeffer, T., V. Staiger, and A. Aertsen. 1989. Synaptic plasticity in rat hippocampal slice cultures: Local "Hebbian" conjunction of pre- and postsynaptic stimulation leads to distributed synaptic enhancement. *Proc. Natl. Acad. Sci.* **86:** 8113–8117.

Bredt, D.S. and S.H. Snyder. 1992. Nitric oxide, a novel neuronal messenger. *Neuron* **8:**(1) 3–11.

Bredt, D.S., C.D. Ferris, and S.H. Snyder. 1992. Nitric oxide synthase regulatory sites. *J. Biol. Chem.* **267:** 10976–10981.

Castro-Alamancos, M.A. and J. Borrell. 1992. Facilitation and recovery of shuttle box avoidance behavior after frontal cortex lesions are induced by a contingent electrical stimulation in the ventral tegmental nucleus. *Behav. Brain Res.* **50:**(1–2) 69–76.

Churchland, P.S. and T.J. Sejnowski. 1992. *The computational brain.* MIT Press, Cambridge, MA.

Churchland, P.S., V.S. Ramachandran, and T.J. Sejnowski. 1993. A Critique of pure vision. In *Large-scale neuronal theories of the brain* (ed. C. Koch and J. Davis). MIT Press, Cambridge, MA.

Clark, S.A., T. Allard, W.M. Jenkins, and M.M. Merzenich. 1988. Receptive fields in the body-surface map in adult cortex defined by temporally correlated inputs. *Nature* **332:**(6163) 444–445.

Cline, H.T. and M. Constantine-Paton. 1990. NMDA antagonists disrupt the retinotectal topographic map. *Neuron* **3:** 413–426.

Cline, H.T., E.A. Debski, and M. Constantine-Paton. 1987. N-methyl-D-aspartate receptor antagonist desegregates eye-specific strips. *Proc. Natl. Acad. Sci.* **84:** 4342–4345.

Cole, B.J. and T.W. Robbins. 1992. Forebrain norepinephrine: Role in controlled information processing in the rat. *Neuropsychopharmacology* **7:** 129–142.

Cooper, J.R., F.E. Bloom, and R.H. Roth. 1970. *The biochemical basis of neuropharmacology*. Oxford University Press, Oxford, UK.

Crick, F.H.C. 1989. The recent excitement about neural networks. *Nature* **337:** 129–132.

Damasio, A.R., N.R. Graff-Radford, P.J. Eslinger, H. Damasio, and N. Kassell. 1985. Amnesia following basal forebrain lesions. *Arch. Neurology* **42:**(3) 263–271.

Davies, S.N., R.A.J. Lester, K.G. Reymann, and G.L. Collingridge. 1989. Temporally distinct pre- and post-synaptic mechanisms maintain long-term potentiation. *Nature* **338:** 500–503.

Dayan, P. and T.J. Sejnowski. 1993. TD(λ): Convergence with probability 1. *Machine learning* (in press).

Decker, M.W. and J.L. McGaugh. 1991. The role of interactions between the cholinergic system and other neuromodulatory systems in learning and memory. *Synapse* **7:** 151–168.

Dickinson, A. 1980. *Contemporary animal learning theory*. Cambridge University Press, Cambridge, UK.

Dodd, J. and T.M. Jessell. 1988. Axon guidance and the patterning of neuronal projections in vertebrates. *Science* **242:** 692–699.

Dolphin, A.C., M.L. Errington, and T.V. Bliss. 1982. Long-term potentiation of the perforant path in vivo is associated with increased glutamate release. *Nature* **297:** 496–498.

Dudek, S.M. and M. Bear. 1992. Homosynaptic long-term depression in area CA1 of hippocampus and effects of N-methyl-D-aspartate receptor blockade. *Proc. Natl. Acad. Sci.* **89:**(10) 4363–4367.

Edelman, G.M. and G.N. Reeke. 1982. Selective networks capable of representative transformations, limited generalizations, and associative memory. *Proc. Natl. Acad. Sci.* **79:** 2091–2095.

Edelman, G.M. and J.A. Gally. 1992. Nitric oxide: Linking space and time in the brain. *Proc. Natl. Acad. Sci.* **89:** 11651–11652.

Foote, S.L., C.W. Berridge, L.M. Adams, and J.A. Pineda. 1991. Electrophysiological evidence for the involvement of the locus coeruleus in alerting, orienting, and attending. *Prog. Brain Res.* **88:** 521–532.

Friedlander, M.J., P.R. Montague, F. Hester, and R.B. Marchase. 1992. Role of nitric oxide in NMDA receptor-mediated release of neurotransmitter in visual cortex. *Soc. Neurosci. Abstr.* **18:** 210.

Gally, J.A., P.R. Montague, G.N. Reeke, and G.M. Edelmen. 1990. The NO hypothesis: Possible effects of a short-lived, rapidly diffusible signal in the development and function of the nervous system. *Proc. Natl. Acad. Sci.* **87:** 3547-3551.

Garraghty, P.E. and J.H. Kaas. 1992. Dynamic features of sensory and motor maps. *Curr. Opin. Neurobiol.* **2:**(4) 522–527.

Garraghty, P.E., C.J. Shatz, and M.J. Sur. 1988. Prenatal disruption of binocular interactions creates novel lamination in the cat's lateral geniculate nucleus. *Visual Neurosci.* **1:** 93–102.

Gilbert, C.D. and T.N. Wiesel. 1992. Receptive field dynamics in adult primary visual cortex. *Nature* **356:** 150–152.

Gluck, M.A. and R.F. Thompson. 1987. Modeling the neural substrates of associative learning and memory: A computational approach. *Psychol. Rev.* **94:**(2) 176–191.

Goldman-Rakic, P.S., S. Funahashi, and C.J. Bruce. 1990. Neocortical memory circuits. *Cold Spring Harbor Symp. Quant. Biol.* **55:** 1025-1038.

Greenough, W.T. and C.H. Bailey. 1988. The anatomy of a memory: Convergence of results across a diversity of tests. *Trends Neurosci.* **11:** 142–147.

Haley, J.E., G.L. Wilcox, and P. Chapman. 1991. The role of nitric oxide in long-term potentiation. *Neuron* **8:** 211–216.

Hammer, M. 1991. "Analyse der funktionellen Rolle des Neurons VUMvx1 bei der Klassischen Konditionierung des Rüsselreflexes der Biene." Ph.D. dissertation, FU Berlin, Germany.

Harder, L.D. and L.A. Real. 1987. Why are bumble bees risk averse? *Ecology* **68:**(4) 1104–1108.

Harris, W.A. 1989. Local positional cues in the neuroepithelium guide retinal axons in embryonic Xenopus brain. *Nature* **339:** 218–221.

Hawkins, R.D. and E.R. Kandel. 1984. Is there a cell-biological alphabet for simple forms of learning? *Psychol. Rev.* **91:**(3) 375–391.

Hebb, D.O. 1949. *The organization of behavior.* Wiley, New York.

Heffner, C.D., A.G. Lumsden, and D.D. O'Leary. 1990. Target control of collateral extension and directional axon growth in the mammalian brain. *Science* **247:** 217–220.

Hertz, J., A. Krogh, and R.G. Palmer. 1991. *Introduction to the theory of neural computation.* Addison-Wesley, Redwood City, CA.

Huang, Y.Y., A. Colino, D.K. Selig, and R.C. Malenka. 1992. The influence of prior synaptic activity on the induction of long-term potentiation. *Science* **255:** 730–733.

Hubel, D.H. and T.N. Wiesel. 1965. Binocular interactions in striate cortex of kittens reared with artificial squint. *J. Neurophysiol.* **28:** 1041–1059.

———. 1970. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J. Physiol.* **206:** 419–436.

Hubel, D.H., T.N. Wiesel, and S. LeVay. 1977. Plasticity of ocular dominance columns in monkey striate cortex. *Phil. Trans. Roy. Soc. Lond. B.* **278:** 377–409.

Iadecola, C. 1993. Regulation of the cerebral microcirculation during neural activity: Is nitric oxide the missing link? *Trends Neurosci.* **16:**(6) 206–214.

Ito, M. 1986. Long-term depression as a memory process in the cerebellum. *Neurosci. Res.* **3:**(6) 531–539.

————. 1989. Long-term depression. *Annu. Rev. Neurosci.* **12:** 85–102.

Izumi, Y., D.B. Clifford, C.F. Zorumski. 1992. Inhibition of long-term potentiation by NMDA-mediated nitric oxide release. *Science* **257:** 1273–1276.

Jordan, M.I. and D.E. Rumelhart. 1992. Forward models—Supervised learning with a distal teacher. *Cognit. Sci.* **16:** 307–354.

Kaas, J.H. 1991. Plasticity of sensory and motor maps in adult mammals. *Annu. Rev. Neurosci.* **14:** 137–167.

Kaas, J.H., L.A. Krubitzer, Y.M. Chino, A.L. Langston, E.H. Polley, and N. Blair. 1990. Reorganization of retinotopic cortical maps in adult mammals after lesions of the retina. *Science* **248:** 229–231.

Kaczmarek, L.K. and I.B. Levitan. 1987. *Neuromodulation: The biochemical control of neuronal excitability.* Oxford University Press, Oxford, UK.

Kato, K., D.B. Clifford, and C.F. Zorumski. 1993. Nitric oxide inhibitors facilitate the induction of hippocampal long-term potentiation by modulating NMDA responses. *Soc. Neurosci. Abstr.* **19:** 905.

Kelso, S.R., A.H. Ganong, and T.H. Brown. 1986. Hebbian synapses in hippocampus. *Proc. Natl. Acad. Sci.* **83:**(14) 5326–5330.

Kirkwood, A., S.M. Dudek, J.T. Gold, C.D. Aizenman, and M.F. Bear. 1993. Common forms of synaptic plasticity in the hippocampus and neocortex *in vitro. Science* **260:** 1518–1521.

Klatt, P., K. Schmidt, and B. Mayer. 1992. Brain nitric oxide synthase is a haemoprotein. *Biochem. J.* **288:** 15–17.

Kleinschmidt, A., M.F. Bear, and W. Singer. 1987. Blockade of NMDA receptors disrupts experience-dependent plasticity of kitten striate cortex. *Science* **238:** 355–358.

Klopf, A.H. 1982. *The hedonistic neuron.* Taylor and Francis, New York.

Kossel, A., T. Bonhoeffer, and J. Bolz. 1990. Non-Hebbian synapses in rat visual cortex. *Neuroreport* **1:**(2) 115–118.

Le Cun, Y., B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel. 1990. Handwritten digit recognition with a back-propagation network. In *Advances in neural and information processing systems 2* (ed. D. Touretzky). Morgan Kaufman, Denver, CO.

Lei, S.Z., Z.H. Pan, S.K. Aggarwal, H.S. Chen, J. Hartman, N.J. Sucher, and S.A. Lipton. 1992. Effect of nitric oxide production on the redox modulatory site of the NMDA receptor-channel complex. *Neuron* **8:**(6) 1087–1099.

Linsker, R. 1986. From basic network principles to neural architecture: Parts I, II and II. *Proc. Natl. Acad. Sci.* **83:** 7508–7512,8390–8394,8779–8783.

————. 1988. Self-organization in a perceptual network. *Computer* March:105–117.

Lipton, S.A., Y.B. Choi, Z.H. Pan, S.Z. Lei, H.S. Chen, N.J. Sucher, J. Loscalzo, D.J. Singel, and J.S. Stamler. 1993. A redox-based mechanism for the neuroprotective and neurodestructive effects of nitric oxide and related nitroso-compounds. *Nature* **364:** 626–632.

Ljungberg, T., P. Apicella, and W. Schultz. 1992. Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* **67:**(1) 145–163.

Loeb, E.P., F. Chang, and W. Greenough. 1987. Effects of neonatal 6-hydroxydopamine

upon morphological organization of the posteromedial barrel subfield in mouse somatosensory cortex. *Brain Res.* **403:** 113–120.

MacKay, D.J.C. and K.D. Miller. 1990. Analysis of Linsker's simulations of Hebbian rules. *Neural Computation* **2:** 169–182.

MacKintosh, N.J. 1974. *The psychology of animal learning.* Oxford University Press, Oxford, UK.

———. 1983. *Conditioning and associative learning.* Oxford University Press, Oxford, UK.

Malenka, R.C., J.A. Kauer, D.J. Perkel, M.D. Mauk, P.T. Kelly, R.A. Nicoll, and M.N. Waxham. 1989. An essential role for postsynaptic calmodulin and protein kinase activity in long-term potentiation. *Nature* **340:** 554–557.

Malinow, R. and J.P. Miller. 1986. Postsynaptic hyperpolarization during conditioning reversibly blocks induction of long-term potentiation. *Nature* **320:** 529–530.

Malinow, R. and R.W. Tsien. 1990. Presynaptic enhancement shown by whole-cell recordings of long-term potentiation in hippocampal slices. *Nature* **346:** 177–180.

Manzoni, O., L. Prezeau, P. Marin, S. Deshager, J. Bockaert, and L. Fagni. 1992. Nitric oxide induced blockade of NMDA receptors. *Neuron* **8:**(4) 653–662.

Marr, D. 1982. *Vision.* Freeman, New York.

McGaugh, J.L. 1989. Involvement of hormonal and neuromodulatory systems in the regulation of memory storage. *Annu. Rev. Neurosci.* **12:** 255–287.

Merzenich, M.M. and K. Sameshima. 1993. Cortical plasticity and memory. *Curr. Opin. Neurobiol.* **3:** 187–196.

Merzenich, M.M., J.H. Kaas, J. Wall, R.J. Nelson, M. Sur, and D. Felleman. 1983. Topographic reorganization of somatosensory cortical areas 3b and 1 in adult monkeys following restricted deafferentation. *Neuroscience* **8:** 33–55.

Merzenich, M.M., R.J. Nelson, M.P. Stryker, M.S. Cynader, A. Schoppmann, and J.M. Zook. 1984. Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comp. Neurol.* **224:** 591–605.

Merzenich, M.M., G.H. Recanzone, W.J. Jenkins, and K.A. Grajski. 1990. Adaptive mechanisms in cortical networks underlying cortical contributions to learning and nondeclarative memory. *Cold Spring Harbor Symp. Quant. Biol.* **55:** 873–887.

Meyer, R.L. 1982. Tetrodotoxin blocks the formation of ocular dominance columns in goldfish. *Science* **218** 589–591.

Miller, K.D. 1994. A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. *J. Neurosci.* **14:** 409–441.

Miller, K.D., J.B. Keller, and M.P. Stryker. 1989. Ocular dominance column development: Analysis and simulation. *Science* **245:** 605–615.

Montague, P.R. 1992. The NO hypothesis. *Neuroscience year: Supplement 3 to the Encyclopedia of Neuroscience* (ed. B. Smith and G. Adelman), pp. 100–103. Birkhauser, Cambridge, MA.

Montague, P.R., J.A. Gally, and J.A. Edelman. 1991. Spatial signaling in the development and function of neural connections. *Cerebral Cortex* **1:** 1- 20.

Montague, P.R., C.D. Gancayco, R.B. Marchase, and M.J. Friedlander. 1992. Role of nitric

oxide in NMDA receptor-mediated neurotransmitter release in hippocampus. *Soc. Neurosci. Abstr.* **18:** 761.

Montague, P.R., P. Dayan, S.J. Nowlan, A. Pouget, and T.J. Sejnowski. 1993a. Using aperiodic reinforcement for directed self-organization. In *Advances in neural and information processing systems 5* (ed. C.L. Giles, S.J. Hanson, and J.D. Cowan). Morgan Kauffman Publishers, San Mateo, CA.

Montague, P.R., P. Dayan, and T.J. Sejnowski. 1993b. Volume learning: Signaling covariance through neural tissue. In *Neural systems: Analysis and modeling* (ed. F. Eeckman), Ch. 57. Kluwer Academic Publisher, Norwell, MA.

Montague, P.R., C.D. Gancayco, M. Winn, R.B. Marchase, and M.J. Friedlander. 1994a. Nitric oxide production links NMDA receptor activation to neurotransmitter release. *Science* **263:** 973–977.

Montague, P.R., P. Dayan, and T.J. Sejnowski. 1994b. Foraging in an uncertain environment using predictive Hebbian learning. In *Advances in neural information processing systems 6*. Morgan Kaufmann Publishers, San Mateo CA. (In press.)

Moore, J.W., J.E. Desmond, N.E. Bethier, D.E. Blazis, R.S. Sutton, and A.G. Barto. 1986. Simulation of the classically-conditioned nictitating membrane response by a neuron-like adaptive element: Response topography, neuronal firing, and interstimulus intervals. *Behav. Brain Res.* **12:** 143–154.

Mountcastle, V.B., R.A. Andersen, and B.C. Motter. 1981. The influence of attentive fixation upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *J. Neurosci.* **1:**(11) 1218–1225.

Mulkey, R.M. and R.C. Malenka. 1992. Mechanisms underlying induction of homosynaptic long-term depression in area CA1 of the hippocampus. *Neuron* **9:**(5) 967–975.

Nicoll, R.A., J.A. Kauer, and R.C. Malenka. 1988. The current excitement in long-term potentiation. *Neuron* **1:** 97–103.

Obermayer, K., H. Ritter, and K. Schulten. 1990. A principle for the formation of the spatial structure of cortical feature maps. *Proc. Natl. Acad. Sci.* **87:** 8345–8349.

O'Dell, T.J., R.D. Hawkins, E.R. Kandel, and O. Arancio. 1991. Tests of the roles of two diffusible substances in LTP: Evidence for nitric oxide as a possible early retrograde messenger. *Proc. Natl. Acad. Sci.* **88:** 11285–11289.

Oja, E. 1982. A simplified neuron model as a principal component analyzer. *J. Math. Biol.* **15:** 267–273.

O'Leary, D.D. and S.E. Koester. 1993. Development of projection neuron types, axon pathways, and patterned connections of the mammalian cortex. *Neuron* **10:**(6) 991–1006.

O'Leary, D.D., A.R. Bicknese, J.A. De Carlos, C.D. Heffner, S.E. Koester, L.J. Kutka, and T. Terashima. 1990. Target selection by cortical axons: Alternative mechanisms to establish axonal connections in the developing brain. *Cold Spring Harbor Symp. Quant. Biol.* **55:** 453–468.

Pettet, M.W. and C.D. Gilbert. 1992. Dynamic changes in receptive-field size in cat primary visual cortex. *Proc. Natl. Acad. Sci.* **89:** 8366–8370.

Placzek, M., M. Tessier-Lavigne, T. Jessell, and J. Dodd. 1990. Orientation of commissural axons in vitro in response to a floor plate-derived chemoattractant. *Development* **110:**(1) 19–30.

Pouget, A., P.R. Montague, P. Dayan, and T.J. Sejnowski. 1993. A developmental model of map registration in the superior colliculus using predictive Hebbian learning. *Soc. Neurosci. Abstr.* **19:** 858.

Quartz, S., P. Dayan, P.R. Montague, and T.J. Sejnowski. 1992. Expectation learning in the brain using diffuse ascending connections. *Soc. Neurosci. Abstr.* **18:** 1210.

Rauschecker, J.P. 1991. Mechanisms of visual plasticity: Hebb synapses, NMDA receptors, and beyond. *Physiol. Rev.* **71:**(2) 587–615.

Rauschecker, J.P. and W. Singer. 1979. Changes in the circuitry of the kitten's visual cortex are gated by postsynaptic activity. *Nature* **280:** 58–60.

Real, L.A. 1991. Animal choice behavior and the evolution of cognitive architecture. *Science* **253:** 980–986.

Redfern, P.A. 1970. Neuromuscular transmission in new-born rats. *J. Physiol.* **209:** 701–709.

Reiter, H.O. and M.P. Stryker. 1988. Neural plasticity without postsynaptic action potentials: Less-active inputs become dominant when kitten visual cortical cells are pharmacologically inhibited. *Proc. Natl. Acad. Sci.* **85:** 3623–3627.

Rescorla, R.A. and A.R. Wagner. 1972. A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. In *Classical conditioning II: Current research and theory* (ed. A.H. Black and W.F. Prokasy) pp. 64–69. Appleton-Century-Crofts, New York.

Sanger, T.D. 1989. Optimal unsupervised learning in a single-layer linear feed-forward neural network. *Neural Networks* **2:** 459–473.

Saver, J.L. and A.R. Damasio. 1991. Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. *Neuropsychologia* **29:** 1241–1249.

Scherer, W.J. and S.B. Udin. 1989. N-methyl-D-aspartate antagonists prevent interaction of binocular maps in Xenopus tectum. *J. Neurosci.* **9:** 3837–3843.

Schlaggar, B.L., K. Fox, and D.D. O'Leary. 1993. Postsynaptic control of plasticity in developing somatosensory cortex *Nature* **364:** 623–626.

Schultz, W. 1992. Activity of dopamine neurons in the behaving primate. *Sem. Neurosci.* **4:** 129–138.

Schuman, E.M. and D.V. Madison. 1991. A requirement for the intercellular messenger nitric oxide in long-term potentiation. *Science* **254:** 1503–1506.

Sejnowski, T.J. 1977. Storing covariance with nonlinearly interacting neurons, *J. Math. Biol.* **4:** 203–211.

Sejnowski, T.J. and C.R. Rosenberg. 1987. Parallel networks that learn to pronounce English text. *Complex Sys.* **1:** 145–168

Sejnowski, T.J., C. Koch, and P.S. Churchland. 1988. Computational neuroscience. *Science* **241:** 1299–1306.

Sejnowski, T.J., S. Chattarji, and P.K. Stanton. 1990. Homosynaptic long-term depression in hippocampus and neocortex. *Sem. Neurosci.* **2:** 355–363.

Selverston, A.I. 1985. *Model neural networks and behavior.* Plenum Press, New York.

Sereno, M.I. and M.E. Sereno. 1991. Learning to see rotation and dilation with a Hebb rule. In *Advances in neural and information processing systems 3* (ed. R.P. Lippman, J.E. Moody, and D.S. Touretzky), pp. 320–326. Morgan Kauffman, San Mateo, CA.

Shatz, C.J. and M.P. Stryker. 1988. Prenatal tetrodotoxin infusion blocks segregation of retinogeniculate afferents. *Science* **242:** 87–89.

Simon, D.K., G.T. Prusky, D.D. O'Leary, and M. Constantine-Paton. 1992. N-methyl-D-aspartate receptor antagonists disrupt the formation of a mammalian neural map. *Proc. Natl. Acad. Sci.* **89:** 10593–10597.

Singer, W. 1990. The formation of cooperative cell assemblies in the visual cortex. *J. Exp. Biol.* **153:** 177–197.

Sparks, D.L. and L.E. Mays. 1990. Signal transformations required for the generation of saccadic eye movements. *Annu. Rev. Neurosci.* **13:** 309–336.

Stent, G.S. 1973. A physiological mechanism for Hebb's postulate of learning. *Proc. Natl. Acad. Sci.* **70:**(4) 997–1001.

Stevens, C.F. and Y. Wang. 1993. Reversal of long-term potentiation by inhibitors of haem oxygenase. *Nature* **364:** 147–149.

Stretevan, D.W. 1990. Specific routing of retinal ganglion cell axons at the mammalian optic chiasm during embryonic development. *J. Neurosci.* **10:** 1995–2007.

Stretevan, D.W., C.J. Shatz, and M.P. Stryker. 1988. Modification of retinal ganglion cell axon morphology by prenatal infusion of tetrodotoxin. *Nature* **336:** 468–471.

Stryker, M.P. and W.A. Harris. 1986. Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *J. Neurosci.* **6:** 2117–2133.

Stuermer, C.A.O. 1988. Retinotopic organization of the developing retinotectal projection in the zebrafish embryo. *J. Neurosci.* **8:** 4513–4530.

Sutton, R.S. 1988. Learning to predict by the methods of temporal difference. *Machine Learn.* **3:** 9–44.

Sutton, R.S. and A.G. Barto. 1981. Toward a modern theory of adaptive networks: Expectation and prediction. *Psychol. Rev.* **88:**(2) 135–170.

———. 1987. A temporal-difference model of classical conditioning. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society.* Seattle, WA.

———. 1989. Time-derivative models of Pavlovian reinforcement. In *Learning and computational neuroscience* (ed. M. Gabriel and J. Moore). MIT Press, Cambridge, MA.

Swindale, N.V. 1980. A model for the formation of ocular dominance stripes. *Proc. R. Soc. Lond.: Series B.* **208:** 243–264.

Tesauro, G. 1994. TD-Gammon, a self-teaching backgammon program, achieves master-level play. *Neural Computation* **6:** 215–219.

Tranel, D. and A.R. Damasio. 1985. Knowledge without awareness: An autonomic index of facial recognition by prosopagnosics. *Science* **228:** 1453–1454.

von der Malsburg, C. 1973. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik* **14:** 85–100.

von der Malsburg, C. and D.J. Willshaw. 1977. How to label nerve cells so that they can interconnect in an ordered fashion. *Proc. Natl. Acad. Sci.* **74:** 5176–5178.

Wall, J.T., J.H. Kaas, M. Sur, R.J. Nelson, D.J. Felleman, and M.M. Merzenich. 1986. Functional reorganization in somatosensory cortical areas 3b and 1 of adult monkeys after median nerve repair: Possible relationships to sensory recovery in humans. *J. Neurosci.* **6:**(1) 218–233.

Wigstrom, H. and B. Gustafsson. 1985. On long-lasting potentiation in the hippocampus: A proposed mechanism for its dependence on coincident pre- and postsynaptic activity. *Acta Physiol. Scand.* **123:** 519–522.

Williams, J.H., M.L. Errington, Y.-G. Li, M.A. Lynch, and T.V.P. Bliss. 1993. The search for retrograde messengers in long-term potentiation. *Sem. Neurosci.* **5:** 149–158.

Willshaw, D.J. 1981. The establishment and the subsequent elimination of poly-neuronal innervation of developing muscle: Theoretical considerations. *Proc. R. Soc. Lond.: B.* **212:** 233–252.

Willshaw, D.J. and P. Dayan. 1991. Optimal plasticity in matrix memories: What goes up MUST come down. *Neural Computation* **2:** 85–93.

Willshaw, D.J. and C. von der Malsburg. 1976. How patterned neural connections can be set up by self-organization. *Proc. R. Soc. Lond.: B.* **194:** 431–445.

———. 1979. A marker induction mechanism for the establishment of ordered neural mappings: Its application to the retinotectal problem. *Philos. Trans. R. Soc. Lond.: B.* **287:** 203–243.

Wise, R.A. 1982. Neuroleptics and operant behavior: The anhedonia hypothesis. *Behav. Brain Sci.* **5:** 39.

Zhuo, M., S.A. Small, E.R. Kandel, and R.D. Hawkins. 1993a. Nitric oxide and carbon monoxide produce activity-dependent long-term synaptic enhancement in hippocampus. *Science* **260:** 1946–1950.

Zhuo, M., E.R. Kandel, and R.D. Hawkins. 1993b. Possible roles of carbon monoxide and cGMP in long-term potentiation in the CA1 region of hippocampus. *Soc. Neurosci. Abstr.* **19:** 241.

The predictive brain: temporal coincidence and temporal order in synaptic learning mechanisms.

P R Montague and T J Sejnowski

References	This article cites 136 articles, 47 of which can be accessed free at: http://learnmem.cshlp.org/content/1/1/1.full.html#ref-list-1
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**