Oscillatory Event-Related Brain Dynamics
Pantey, C., Elbert, T., & Luetkenhoener, B. (Editors).
New York, New York: Plenum Press (1994)

TUNING AND FILTERING IN ASSOCIATIVE LEARNING

Larry E. Roberts, Ronald J. Racine, Paula J. Durlach, and Sue Becker

Department of Psychology
McMaster University
Hamilton, Ontario
Canada L8S 4K1

INTRODUCTION

The individual organisms of most animal species occupy variable environments whose detailed features cannot be fully anticipated by a genetic code. The evolutionary response to this ecological constraint has been to build into that code mechanisms for abstracting the structure of individual environments and for generating behavior based on the representations thus formed. To adequately represent an environment, an organism must encode the physical features of its world, relations among those features and events that occur in that world including its own actions and their consequences, and the temporal flow of this information in time. Memory and perception can be considered to operate in the service of the representational problem, which is an associative one in the sense that these aspects of spatiotemporal information must be extracted and encoded if an ecological niche is to be adequately described.

The problem of how the brain carries out its associative and representational functions was addressed by early researchers in experimental psychology and neuroscience (Lashley, 1950; Hebb, 1949; Konorski, 1967; Deutsch, 1960). The accounts they suggested were structural in the sense that specific neural mechanisms were proposed, although computational modelling of those mechanisms was not attempted. In the two or three decades following these efforts, the study of learning shifted away from structural theories to laboratory investigations of Pavlovian and instrumental conditioning (Mackintosh, 1983) and to the development of quantitative theories of the acquisition process (Pearce & Hall, 1980; Rescorla & Wagner, 1972). Although this research has deeply enriched our understanding of associative mechanisms, it has also set into relief some problems regarding the nature of representations (Colwill, 1993; Miller & Barnet, 1993) and the organization

of action (Holland & Rescorla, 1982) that might be better solved within the framework of a structural approach (Brener, 1986). Interest in structurally-based accounts of learning has recently resurfaced, not only because the contribution of laboratory analyses of learning may be saturating, but also because developments in neuroscience and in neural network modelling have opened avenues of description and inquiry that were not available to early pioneers.

Recent structural models have typically taken as their starting point the task of capturing specific conditioned responses (e.g., Gluck et al., 1993) or selected phenomena of signalling such as blocking, overshadowing, and the learning of non-linear associative relations (e.g., Schmajuk & DiCarlo, 1992; Gluck & Meyers, 1993). The model described in this paper, on the other hand, attempts to give a more general account of an organism's adjustment to a learning situation. We begin by describing some background facts that invited the current approach. These facts concern the nature of Pavlovian and instrumental conditioning and how mechanisms that support these types of learning might be reflected in the organization of the brain. A neural network architecture is then described that performs learning functions by means of two processes, (a) a filtering process that identifies unexpected information arising over sensory pathways, and (b) a tuning process that selectively augments neural processing in sensory modalities that are conveying surprising information to the cortex. In the concluding sections we discuss implications of the model for simulation and experimental studies, and comment on how learning functions may be reflected in the electrical and magnetic field activity of the brain.

BACKGROUND

Any model of a learning system should try to capture basic findings from behavioral research which indicate how learning works. In this respect, the idea that organisms come to know or represent the spatiotemporal stucture of their worlds was long resisted by behavior theory, in favor of the idea that linkages between responses and controlling stimuli were sufficient to account for learned behavioral adaptation (see Mackintosh, 1983, for a review). However, there is now much evidence that environmental relations are encoded by animal (Rescorla, 1987, 1988) and human subjects (Dawson & Schell, 1987) during conditioning experiences. For example, in human subjects differentiation of behavior between signalling stimuli does not occur in the absence of the ability of the subjects to describe the Pavlovian or instrumental contingencies to which they have been exposed (Dawson & Bifemo, 1973; Hughes & Roberts, 1985). When verbal reports of reinforcement contingencies are dissociated from response differentiation in either conditioning arrangement (this occurring in 15% of subjects in the feedback experiments of Hughes & Roberts, 1985), it appears to be invariably the case that knowledge of the contingencies precedes the differentiation of overt behavior. Overall, the linkage between abstracting of environmental relations and adapting behavior to those relations is very close (Roberts, 1990).

Also important for structural theories of learning are numerous similiarities which exist between Pavlovian and instrumental conditioning at the associative level. These similarities pertain not only to the close linkage between associative

knowledge and response differentiation just mentioned, but also to basic principles of signalling in the two conditioning arrangements. For example, if multiple stimuli predict a reinforcer, subjects learn preferentially about the best predictor, in accordance with a principle of relative validity (Mackintosh, 1983; Wagner et al., 1968). Similarly, if reinforcement is allocated differentially among several alternative responses, the response that is scheduled to receive the highest rate of reinforcement is strengthened the most by the training arrangement (Herrnstein, 1970). Stimuli and responses also appear to be equivalent in their ability to signal a reinforcer. For example, bar pressing for food reward is impaired if responding produces an exteroceptive stimulus that is equally valid as a signal for reinforcement (Williams & Heyneman, 1982). These and other associative similarities summarized by Mackintosh (1983) are important, because they imply that the same associative mechanism is responsible for encoding of Pavlovian and instrumental contingencies (Weisman, 1977). The existence of a single associative mechanism, on the other hand, gives reason to try to describe and model this mechanism. In the model that we describe here, Pavlovian and instrumental contingencies are assumed to differ only with regard to whether stimuli that signal reward are conveyed principally by an exteroceptor (the Pavlovian case) or by kinesthetic pathways (the instrumental one).[1]

One might expect that the requirements of an encoding system that extracts relations among different types of events would be reflected in the general organization of the brain. In this respect it may be noteworthy that there is significant uniformity in the morphology and organization of cortical structures in the brain. It appears from histological evidence that up to 90% of the synapses in the cortex are excitatory or Type I synapses, and as many as 85% of these synapses may be provided by a single category of cell, the pyramidal neuron (Braitenberg & Schuz, 1991). There is less agreement about the uniformity of inhibitory neurons (Type II), some authors proposing half a dozen types while others only one basic type, in any case not too many (Douglas & Martin, 1990). Organization of these neuronal processing elements into laminae and columns is characteristic of most cortical regions, even though variations exist between regions that serve different functional roles. The capacity for plastic changes in the form of long term potentiation (LTP) has been documented at several levels of the brain including neocortical and paleocortical structures, magnocellular and intralaminar thalamic nuclei, basal forebrain regions, and peripheral ganglia (Gerren & Weinberger, 1983; Lynch & Granger, 1992; Racine et al., 1983). These findings suggest that learning depends not only on specialized neurons and local architectures that are adapted for specific purposes, but also on the way in which neurons are organized into cortical regions, and in the kind of information that is delivered to these regions over time.

[1]Although differences in the modality of the signalling stimulus do not appear to alter how association works. the properties of behavioral adaptations brought out by signalling (for example, whether the response is executed by striate or smooth muscles, the voluntary nature of the act, and its access to consciousness) are affected by modality differences (Roberts, 1990). How behavior is generated from environmental representations that are established by conditioning is a problem for modern behavior theory (Rescorla. 1988). The learning model described in this paper does not address this problem in a detailed way, but the hierarchically organized and distributed nature of encoding in the system is compatible with the form of analysis advanced by Brener (1986).

The model that we describe explores how these 'macroarchitectural' features of brain organization might support a learning system.

We have also been guided by some facts relating to brain electrical activity during performance on cognitive and perceptual tasks. Slow potentials of the brain shift toward negativity prior to the performance of motor responses (Gaillard, 1986) as well as prior to the delivery of informational stimuli when motor responding is absent (Chwilla & Brunia, 1991). P300-like waves are augmented by stimulus novelty (Johnson, 1986), and although dynamic changes in these waves have seldom been studied over extended periods, their dependence on novelty suggests they ought to subside as the eliciting event becomes predicted. Recent evidence suggests that P300 waves are generated by inhibitory mechanisms that are either widely distributed in cortical and subcortical structures (Halgren et al., 1986; Woodward et al., 1992) or are organized in a such way that their effects can be selectively manifested in these structures (Roberts et al., 1994; Rockstroh et al., 1992). The intrinsic repetitive firing behavior of many neurons is also relevant to biological and computational models of learning (Basar & Bullock, 1992). Coherent oscillatory activity has been recorded from ensembles of neurons in several brain structures and shows properties such as sensitivity to apparent motion (Singer et al., 1990) and deployment of attention (Murthy & Fetz, 1992) that are consistent with a role in perception and/or encoding. Slow waves, P300 events, and oscillatory rhythms are of interest to our agenda, not only because any model of learning must account for them, but also because these phenomena seem likely to express the dynamics of a learning system that is embodied in the general organization of the brain.

A more specific point of departure for our attempt to describe a learning system came from behavioral experiments which suggested that sensory pathways are selectively modulated or 'tuned' by conditioning arrangements. Roberts et al. (1991) monitored the target of visual fixation while human subjects learned to produce two patterns of action that were identified by auditory feedback signals alone. The purpose of the experiment was to determine whether the previously documented ability of subjects to accurately describe behavioral outcomes of training (Hughes & Roberts, 1985) might depend in part on whether their actions were self-monitored in vision. Contrary to expectation, the results showed that subjects typically did not watch what they were doing during feedback training, even though provision of feedback in the form of auditory signals meant that they were free do so. On the contrary, subjects tended to close their eyes on feedback trials as training progressed (see Figure 1A). Bramwell (1993) recently corroborated this observation in a different training arrangement (Figure 1B) and went on to show that if a visual detection task was superimposed on the auditory feedback problem so that vision was now required, spontaneous blinks and detection errors were increased compared to a condition in which the visual task was performed alone. These findings suggested that eye closures may have occurred during the auditory feedback problems of Figure 1, because processing in visual channels was relaxed or 'tuned out' by the presence of auditory feedback signals that also required processing. The progressive nature of the eye closure effect also suggested that repeated conjunctions between events in kinesthetic and auditory but not visual pathways may have favored the development of eye closures in this training environment.

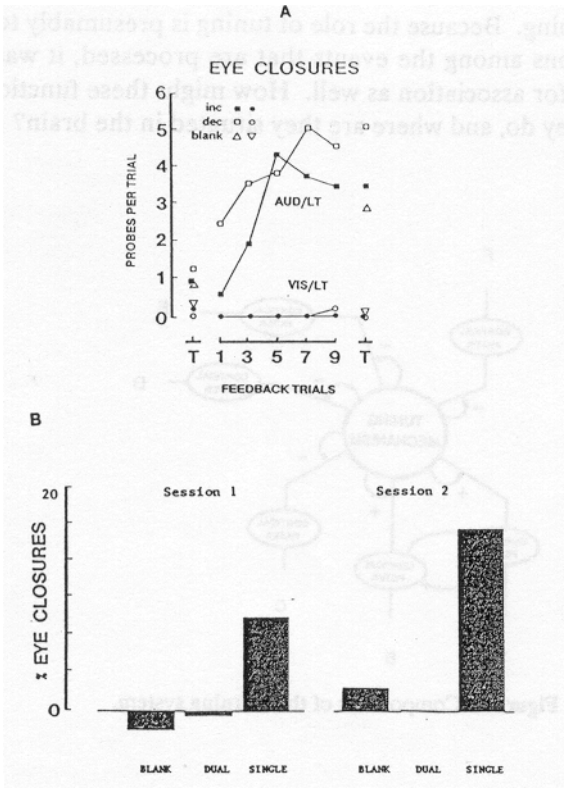

Figure 1. (A) Subjects received either auditory feedback (AUD/LT or visual feedback (VIS/LT) for generating two bidirectionally opposite patterns of unidentified behavior. Eye closures were observed during auditory feedback as training progressed. By the end of training, eye closures occupied approximately 65% of each auditory feedback trial relative to baseline (Roberts et al., 1991). The two behavior patterns that were trained consisted of striate muscular activities associated with increases and decreases in cardiac interbeat intervals. (B) Subjects solved an auditory feedback problem on single task trials and a visual detection task concurrently on dual task trials, in two training sessions separated by a brief rest interval. The onset of each trial was accompanied by a 'beep' issued from a computer. Measurement of eye closures showed that the subject's eyes were open on dual task trials and on blank trials which occurred during intertrial intervals. However, eye closures occurred and intensified over the course of training on single task trials where auditory feedback was processed (Bramwell, 1993).

Although tuning expressed as perceptual interference has not been widely studied in conditioning experiments, the fact that discriminative behavior often results from subjects learning to orient toward and approach objects or locations that are associated with reward, to the detriment of objects or locations that are not, suggests that the role of tuning in behavioral adaptation may be considerable (Jenkins & Sainsbury, 1970). Signalling principles such as blocking and overshadowing also appear to be amenable to analysis in terms of tuning (Mackintosh, 1983), and examples of lateral and surround inhibition which could support such phenomena are well documented in the nervous system (Kandel et al., 1991). We therefore began our project by attempting to describe an architecture that might tune sensory systems

during associative learning. Because the role of tuning is presumably to promote the formation of associations among the events that are processed, it was desirable to consider a mechanism for association as well. How might these functions be served, how much work can they do, and where are they situated in the brain?

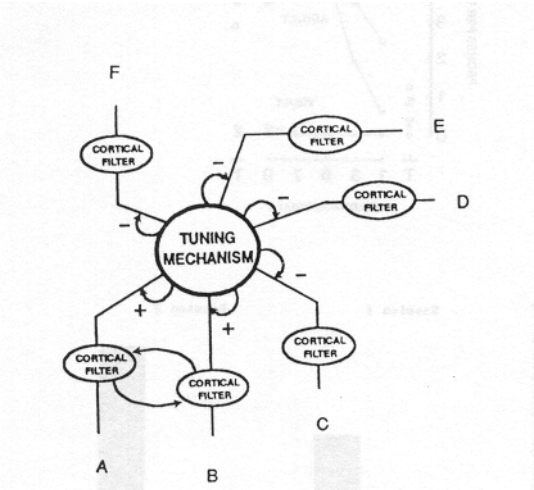

Figure 2. Components of the learning system.

TUNING AND FILTERING

If tuning among sensory modalities is to occur, information about activity in different modalities must converge in the nervous system. Thalamic nuclei offer an early opportunity for intermodal sensory processing and seem likely on this account to play a role in tuning functions. On the other hand, the extensive reciprocal connections which exist between thalamic nuclei and neocortical and paleocortical regions imply other levels of organization to and from which information is delivered, and are consistent with a principle of distributed associations (Lashley, 1950). These principles are not controversial, and we began by describing a learning system that included them.

Figure 2 describes the elements of the system in abstract terms, without explicit reference to neuronal mechanisms (Roberts et al., 1992). Pathways A-F represent sensory modalities that impinge on a tuning mechanism; in each modality there is also a cortical filter that processes sensory input. Briefly, we want tuning to work as follows. If an unexpected event occurs in one modality (A), processing in that modality is facilitated while processing in other modalities is dampened simultaneously (CDEF). However, if a biologically significant event has co-occurred in a second modality (B) because the two events are linked by an environmental contingency, that modality is spared from inhibition and is facilitated instead. Synaptic activity is therefore enhanced in neural networks that are driven by unpredicted and temporally coupled task events, and dampened elsewhere (receptor orienting acts may be part of this process). The purpose of tuning is to enhance the rate at which uncoded data enters the system, and to amplify the effect of these data

on synaptic plasticity. As a consequence of data-driven processing, links are forged between events at distributed levels of the system.

Associative functions are performed by the cortical filters which are shown in each modality in Figure 2. Although only two links are illustrated (the arrows connecting filters in modalities A and B), the pattern of connectivity among filters in different modalities is assumed to be complete. The cortical filters serve two functions. First, they gate sensory input to the tuning mechanism. Only unexpected events gain access. Second, the cortical filters are a major site of learned association. They receive input from many modalities, and they change their filtering characteristics as synaptic weights are modified by Hebbian rules. A further assumption of the model is that once two sensory events have become associated by filtering, those events lose their access to the tuning mechanism. In other words, association acts as a filter. We need the filtering function to capture phenomena such as automaticity, activational peaking over the course of training (Germana, 1968), and an expected diminution of P300 waves with experience (Johnson, 1986). At least one current model of Pavlovian conditioning (the SOP model of Wagner & Brandon, 1989) makes a similar assumption, although that assumption is couched in quite different terminology.

Before leaving Figure 2, we should comment on the concept of a modality (pathways A-F). This term is usually taken to refer to a sensory system such as vision, audition, kinesthesis (which signals reward in the instrumental case), and so on. But sensory systems are more finely differentiated than this term allows. For example, the visual system conveys different types of information over distinct pathways (the magnocellular and parvocellular, at the level of the lateral geniculate). Auditory cortex is tonotopically organized, and there is extensive columnar organization in visual and kinesthetic pathways. Although it is reasonable to suggest that tuning effects may extend to multiple levels including an entire sensory system (Figure 1), we prefer to link the sensory channels of our model with more detailed structural elements such as cortical columns. A tuning mechanism whose bandwidth is wide is likely to play a more important role in association than one whose bandwidth is narrow.

Figure 2 identifies two elements, tuning and filtering, that may be involved in learning, and although thalamic and cortical mechanisms are implicated in these functions, there are no neurons here. The next question was, how might these mechanisms work neuronally, and in real time?

NEURAL NETWORK IMPLEMENTATION

Given technological constraints, any modelling effort can only incorporate a small sample of the findings from neuroscience research. Although it is not an easy task to determine which findings should be included in a model, it is a task worth attempting because application of this information can have major effects on the functioning of a neural network (Lynch & Granger, 1992). In implementing the model of Figure 2 we have tried to accommodate some documented properties of real neural networks. Among these properties are (a) selected aspects of corticothalamic organization including reciprocal connectivity and a substantial bandwidth in the forward and backward paths; (b) an increase in the number units

available for encoding at higher levels of the system (this feature enabling sparse encoding); (c) a provision for top-down processing of sensory input (Peterhans & von der Heydt, 1989); and (d) modulatory mechanisms that appear to alter synaptic gain via basal forebrain and/or intralaminar thalamic pathways (Hasselmo & Bower, 1992). Our model is still in the developmental stage, with only pieces of it having been tested. In the following we describe the basic implementation and comment on how we are pursuing it.

An overview of the architecture is given in Figure 3 where three sensory channels (A, B, and C) are shown. Although each channel can be thought of as a cortical column, the nature of processing in the system is such that whole sensory modalities could be gated up or down by the tuning system. Thalamic relays convey activation to two central elements in each channel, a readout system and a comparator, both situated in the cortical column. It is important to note that the state of the readout system is determined by its thalamocortical input acting in concert with previous and concurrent activations, within and between sensory channels. The readout system integrates the information it receives and conveys a portion of its output as an inhibitory pattern to the next element in the system, the comparator.

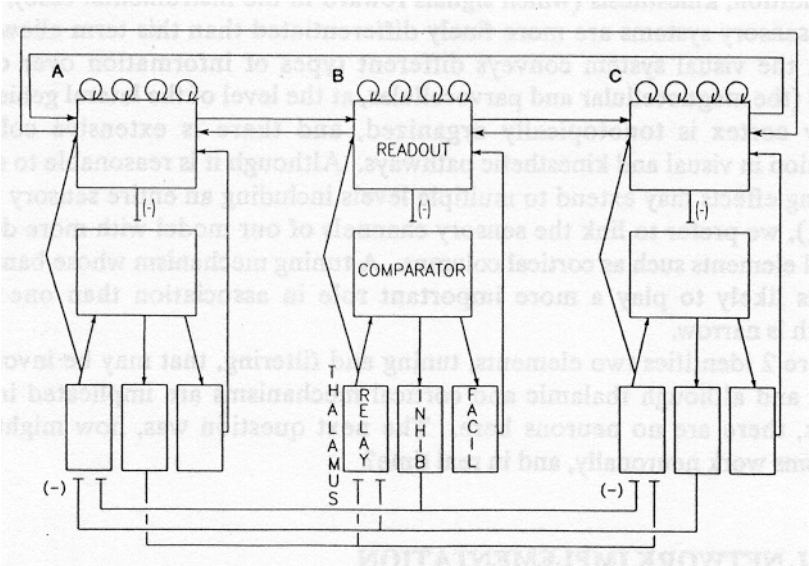

Figure 3. Overall organization of the model. Three channels or modalities are shown (A, B, C). In each modality thalamic relay neurons convey current input patterns to a readout system and a comparator situated in the cortex (both inputs are excitatory). The current state of the readout system depends on previous activations both with and between channels (cortical columns). The readout system sends part of its output as an inhibitory pattern to the comparator. The comparator compares the actual input with the expected input and generates a signal proportional to the mismatch. This signal is conveyed back down to the thalamus where it 1) activates an inhibition of channels which did not experience much of a mismatch or in which activation levels are currently quite low, and 2) drives a thalamic facilitator which increases the gains in the readout system proportional to the mismatch found in that system. These events are permissive for associations to form between channels.

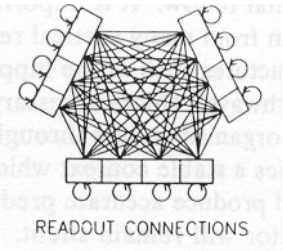

Figure 4. The readout level. The three layers represent three different cortical columns (modalities). The connectivity in this particular implementation is exhaustive and recurrent. This is where learning takes place when there is a mismatch of patterns arriving at the level of the comparator.

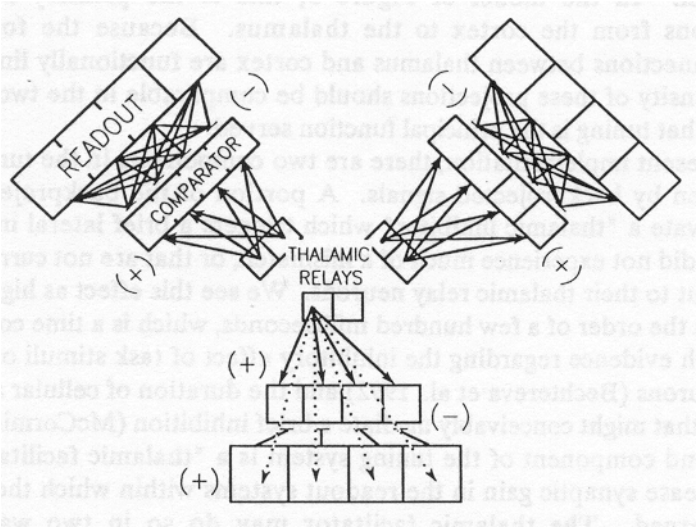

Figure 5. Removing the connections between modalities at the readout level, we can see the connections from the readout layers to the comparator. These connections are inhibitory, while connections ascending to the comparator from thalamic sensory nuclei are excitatory. When the actual pattern matches the expected pattern, there is no output from the comparator, but if a mismatch occurs, an output is generated proportionally which returns to the thalamic level as shown in Figure 3. The smaller number of cells at the thalamic layer is meant to depict increased capacity at the cortical level. Thalamic relays are also fully connected to the readout layer, but to simplify the figure this is illustrated only in the lower module.

It is the comparator acting in conjunction with the readout system that serves the filtering function. Specifically, the pattern of inhibition impinging on the comparator from the readout system constitutes a prediction of the state of excitation that is expected to be conveyed by thalamic relays at the next instant in time. If afferentation rising from thalamic relay neurons matches the readout pattern, the two patterns nullify one another and the comparator is silent. Psychologically this corresponds to the state of affairs in which there are no surprises in the environment, and

the moment by moment context that the organism finds itself in predicts, with considerable reliability, the events that follow. It is important that the readout system receive and integrate information from many cortical regions including not only intrinsic pathways arising from structures such as the hippocampus, but also from the exteroceptors and kinesthetic pathways. This is necessary if thalamic afferentation is to be predicted accurately as the organism moves through time and space.

Now, if the organism occupies a stable context which it has experienced extensively, the readout system should produce accurate predictions of the course of thalamic activity, and the comparator will remain silent. However, if a novel event occurs (such as the appearance of CS or US in a just-imposed conditioning arrangement, or an unpredicted stimulus in a cognitive task), inhibition conveyed to the comparator by the readout system will not cancel the excitation arriving from thalamic relay neurons. In this case, there is a mismatch, and the resulting output from the comparator is conveyed back down to the thalamus where it drives the tuning system. In the model of Figure 3, this is the primary role of the backprojections from the cortex to the thalamus. Because the forward and backward connections between thalamus and cortex are functionally linked by the model, the density of these projections should be comparable in the two pathways, to the extent that tuning is the principal function served.

In the present implementation, there are two components in the tuning system which is driven by backprojected signals. A portion of the backprojected signal serves to activate a "thalamic inhibitor" which triggers a brief lateral inhibition of channels that did not experience much of a mismatch, or that are not currently being driven by input to their thalamic relay neurons. We see this effect as highly specific and lasting on the order of a few hundred milliseconds, which is a time course that is consistent with evidence regarding the inhibitory effect of task stimuli on the firing of cortical neurons (Bechtereva et al., 1992) and the duration of cellular afterhyperpolarizations that might conceivably mediate a brief inhibition (McCormick, 1990).

The second component of the tuning system is a "thalamic facilitator" which serves to increase synaptic gain in the readout systems within which the mismatch patterns emerged. The thalamic facilitator may do so in two ways, (a) by depolarizing the apical dendrites of superficial pyramidal cells through intralaminar or modality-specific magnocellular structures, thus decreasing the threshold for activation of NMDA receptors; and (b) by gating down synaptic weights in association relative to afferent pathways through basal forebrain structures which appear to have this effect in at least some cortical regions (Hasselmo & Bower, 1992). The time constant of thalamic facilitation is thought to be longer than that of thalamic inhibition, perhaps lasting on the order of seconds, and the effect of facilitation is thought to be less specific, extending to nearby columns or an entire sensory modality. A combination of these two tuning events (lateral inhibition and thalamic facilitation) favors the formation of associations between temporally related patterns that are generated within the various activated channels, while affording protection of synaptic weights in channels whose current input is properly coded.[2]

[2] Encoding is selectively favored because (a) NMDA receptors mediating unexpected events in afferent pathways are facilitated, and (b) lateral inhibition ensures that only thalamic relay neurons that are driven reliably by their receptors discharge into the cortex during the period of inhibition. One encoding advantage conferred by inhibition lasting a few hundred milliseconds is that novel

In real biological systems, associations are formed at different levels. For example, LTP occurs in cortical synapses as well as in synapses found in the hippocampus, cerebellum, and magnocellular thalamic nuclei (Gerren & Weinberger, 1983; Lynch & Granger, 1992; Racine et al., 1983). It is therefore of interest to our agenda to explore the effect of adding plasticity at different levels. However, for the present we are restricting our attention to associations formed within the readout layer which forms the principal associative system. The architecture of this system is shown in greater detail in Figure 4. The three readout registers represent three different channels or cortical columns. In our initial implementation connectivity is complete and recurrent (the recurrent connections coding temporal information). This is where associative learning takes place when there is a mismatch from expected patterns at the level of the comparator.

In subsequent implementations we plan to explore the effect of adding recurrently connected hidden units to store all or some of the information regarding how spatiotemporal context changes over time. This might make it easier to uncouple patterns representing past versus current information.

Figure 5 shows relationships between the thalamic relay, readout, and comparator systems in more detail. As described above, connections from the readout system to the comparator are inhibitory, and those from thalamic relay neurons to the comparator are excitatory. In our current implementation, dual innervation of the comparator is enforced by the provision that synaptic weights must be either positive or negative and not change from one to the other. Eventually, we may include a bank of "interneurons" with empirically derived properties, and driven by excitatory afferents, to accomplish inhibition at the level of the comparator. One way to ensure the proper operation of the comparator is to employ anti-Hebbian learning on the connections between the readout and comparator levels, to force a match between learned and current patterns as new patterns are experienced. In contrast to the highly specific role of the comparator, the readout system plays a more general role of pattern storage by encoding associations between current and past inputs within and between modalities. An important feature of the arrangement shown in Figure 5 is that a smaller number of cells at the thalamic layer innervates a larger number of cells at the readout level. This means that there is an increase in capacity at the cortical level, an arrangement that favors sparse encoding and orthogonalization of input patterns (McNaughten & Nadel, 1990). Currently the readout system is modelled as a Hebbian pattern associator network. We are also exploring a form of competitive learning within the readout systems in order to enhance orthogonalization of stimuli and storage capacity in the encoding system.

events that follow unexpected stimuli within a brief time interval are likely to have been caused by those stimuli. Another advantage is that thalamic afferentation that arises from sources that are properly coded is less likely to be affected than would be the case were inhibition to be long lasting. It should be noted that because thalamic afferents feed into the readout or associative layer as well as the comparator (Figure 3, a property consistent with known columnar architectures), predicted events will continue to drive cortical processing even though alteration of synaptic weights is not supported by the tuning mechanism. It should also be noted that events which are not unexpected may still have some predictive value for new events. These events may enter into new associations but will do so more slowly because of the relative gating down of their channels.

PROBLEMS FOR SIMULATION AND EXPERIMENT

One reason for attempting to model a general learning system is that such a system almost certainly exists in the brain (Mackintosh, 1983; Weisman, 1973). An obvious risk, however, is that in attempting to do so one ends up modelling not just a learning system but brain function as a whole, which while ultimately desirable (Grossberg, 1987; Konorski, 1967) is probably not a realistic goal. One way that one might keep the task manageable is to approach the problem in steps, working on aspects that seem important and tractable. A stepwise approach may also make it easier to stay close to experimental findings and to frame experimental questions whose pursuit can change one's thinking about aspects of the model. In the preceding section we have already mentioned some of the variables we are exploring. Next, we comment on some additional questions pertaining to computer simulation and on some experimental findings that bear on the model.

Simulation Studies

There are a number of performance features, both general and specific, that we seek from this model, if it is to be a realistic model of associative encoding in the brain. For example, in general terms the model must obviously include a means by which very similar patterns can be separated (when appropriate) and dissimilar patterns can be grouped. As mentioned above, sparse encoding at neocortical levels is expected to assist orthogonalizing and grouping of stimulus inputs, particularly if a competitive learning rule is introduced at this level of the system. A further feature of the model is that the plasticity itself should be somewhat adaptable. By this we mean that the system should learn under several quite different situations: i) When there is a large departure from the expected pattern, ii) when there is a reliable departure from the expected pattern, even if the departure is not large; and iii) when there is an accompanying biologically significant event. The rate of learning should be proportional to these parameters.

Most of the physiological evidence presently available regarding mechanisms of synaptic plasticity in cortex has been gathered in allocortical and paleocortical structures (for example, the hippocampus) and points to simple forms of Hebbian learning in these structures (Brown et al., 1990). However, there is reason to suggest that more complicated forms of Hebbian learning may be needed to describe plasticity in the neocortex. Singer and his colleagues have recently reported that connection weights can show decrements as well as increments when both pre- and postsynaptic elements are active at a synaptic junction (Artola et al., 1990). The direction of change depends on separate thresholds for these effects, such that the postsynaptic unit must be activated beyond one threshold if a decrement in conductivity is to occur. There is, however, a second, higher, threshold. When this second threshold is reached, an increment in connection weights is observed. Increases in postsynaptic activation beyond this level increase the amount of the increment up to some asymptotic level. Although Singer and his colleagues found these effects in slices of visual cortex suspended in vitro (Artola et al., 1990), Racine and his coworkers have obtained evidence that this rule may apply in chronic preparations as well. Racine et al. (1994, I, II) found that coactivation of cholinergic systems with pilocarpine during LTP produced a long term depression of cortical

synapses. However, subsequent research which substituted externally applied DC currents for coactivation of cholinergic inputs enabled a long term potentiation of cortical neurons, perhaps because DC polarization of cortical neurons induced higher levels of postsynaptic activation than did coactivation of cholingeric systems.

In view of these findings in real biological systems, it is of interest to explore in a model that emphasizes cortical processing (as ours does) the effect of introducing a bidirectional encoding mechanism of the type reported by Artola et al. (1990). A similar algorithm was described earlier by Bienenstock, Cooper and Munro (1982) and is often called the BCM Rule. Such a rule offers several potential benefits. For example, simulations carried out by Hancock et al. (1991) suggest that it can facilitate orthogonalization of input patterns. Another possible function relates to how spurious correlations that may arise from the environment are dealt with by cortical processing. In the absence of modulatory inputs to push the postsynaptic activation beyond the second threshold, synapses that are driven by spurious correlations may be decremented and the effects of these correlations thus minimized.

Another question that we have considered is whether back propagation algorithms should be introduced as an alternative to Hebbian learning in the system. Although Hebbian rules are efficient in two-layer networks, back propagation is a powerful algorithm that can train hidden units in multi-layer architectures, which gives it an advantage in representing complex and nonlinear inputs to a system. However, one can question whether back propagation gives an accurate account of what real neural networks actually do (Zipser & Rumelhart, 1989). A feature common to back propagation models and the tuning model described here is that both compute errors in information processing. However, at this point the similarities appear to end. In back propagation models the error is detected at the output, whereas in our network (and probably also in real neural systems) errors are computed at the level of perception. Furthermore, in back propagation the error signal is conducted antidromically through the network. Because real neurons do not work this way, a neuronal implementation requires complementary hardware which can convey the error signal back to components of the forward path, so that synaptic weights are altered according to their contribution to the output (see Schmajuk & DiCarlo, 1992, and Zipser & Rumelhart, 1989, for examples of this approach). Although this solution may be biologically feasible, it is also biologically expensive. The tuning model described in this paper, on the other hand, uses error detection to direct the course of processing in the system. A major task of simulation studies will be to determine the extent to which this feature, in combination with anti-Hebbian and other biologically supported encoding rules, can support rapid and accurate learning comparable to that produced by back propagation.

To date, we have implemented the tuning function in a simple three-channel network, excluding the associative function (i.e., active connections between the readout registers of Figure 4). Although the properties of this simple system are not yet fully determined, it appears from our current understanding that this system will be able to generate the phenomena of blocking and overshadowing, provided that separate sensory channels are used to represent the component stimuli. One interesting implication of the qualitative architecture of Figure 2 is that blocking and overshadowing can be expected to occur as long as the tuning system is active during

learning. however, once this system has been disengaged owing to a well-formed association, salient features added to a predictive stimulus should again acquire signal value, to the extent that tuning is reinstated by them. At this time we are not aware of behavioral data that directly evaluate this possibility. Blocking may also occur at levels of the system which we have not modelled, such as those involved in generating overt responses (Lynch & Granger, 1992).

Experiments

One reason for attempting to formulate a structural model of learning is to explore experimental questions raised by it. We are using the model to guide the design and analysis of human psychophysiological experiments and animal studies.

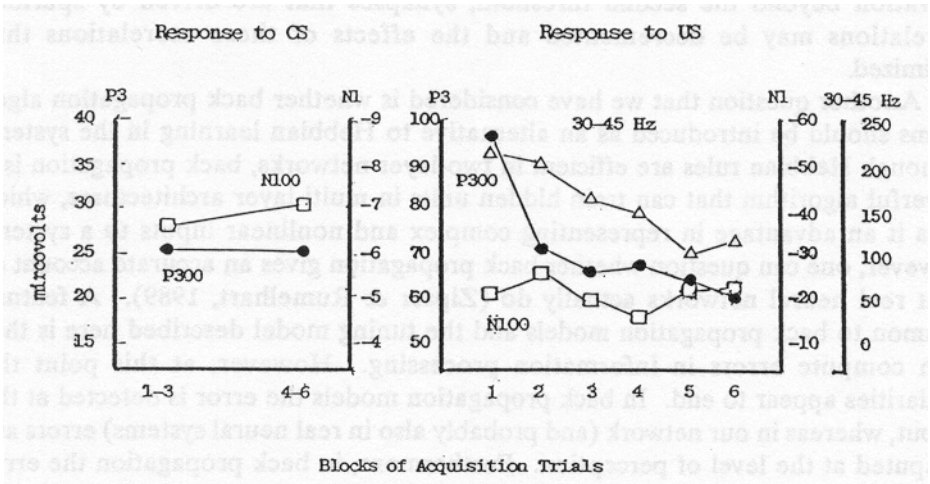

Figure 6. Response of N100 and P300 event-related potentials to CS+ (left panel) and US events (right panel) during discriminative aversive conditioning in human subjects. Spectral power between 30-45 Hz is also shown for a one-sec interval following delivery of the US (augmented power was not detected in this frequency range following CS+). lie metric for spectral power is the area of isoline plots containing baseline-corrected *t*-statistics exceeding p <.001, multiplied by the value of *t*. Results are taken from Flor et al. (1994) and are averaged over three procedurally isomorphic conditioning groups (Cz recording).

Results recently gathered by Flor et al. (1994) on aversive classical conditioning in human subjects give reason to suggest that the model may offer a first approximation of brain dynamics during learning. Flor et al. (1994) carried out discriminative conditioning using different human faces as conditioned stimuli (ass) and strong intracutaneous electric shock as the unconditioned stimulus (US). Three groups of subjects were trained which differed with respect to whether an angry, happy, or neutral face signalled the US event. In each group a slow negative wave was observed to develop in the EEG following CS+ but not CS-, and to extinguish when the US was discontinued. The time course of the slow wave effect is consistent with a role for such waves in preparing neural networks for alteration of synaptic weights

by US events. Because slow negativities appear to reflect depolarization of apical dendrites and show variable topography depending on the requirements of a task (Birbaumer et al., 1990, 1992), it is also reasonable to suggest a thalamic facilitator as their source. However, it appears likely that plasticity will have to be introduced into thalamic or basal forebrain structures in the model of Figure 3, because there is nothing in the model at present that can account for a differential response of slow waves to CS+ and CS- events. Flor et al. (1994) also recorded P300 waves that were elicited by CS and US events over the course of conditioning. P300 waves elicited by the US were found to diminish over acquisition trials, but P300s elicited by CS occurrence did not. These results, which are summarized in Figure 6, are consistent with the view that P300 waves are released by corticothalamic feedback consequent on surprise. CS events were not predictable from the cues that preceded them in Flor et al.'s training arrangement, but US events were perfectly predictable from the CS+.

Flor et al. (1994) subsequently applied power spectral analyses to their results, using a sliding window centered on successive 200 msec epochs within conditioning trials. Although the most prominent finding was augmented power between 0-5 Hz following CS+ and the US (this effect reflecting slow and P300-like waves), increased power was also detected between 30-45 Hz immediately following delivery of the US in each conditioning group. In two groups (these receiving either different neutral faces as CS+ and CS- stimuli, or an angry face as CS+ and a happy face as CS-) the 30-45 Hz effect disappeared by the third block of acquisition trials, whereas in the third group (this group receiving a happy face as CS+ and an angry face as CS) the effect diminished more slowly over the course of acquisition. Because the 30-45 Hz response appeared only following early US events, it appears to have been driven by US occurrence rather than by CS offset. Figure 6 includes the 30-45 Hz response so that its temporal course can be compared with the P300 elicited by the US. At this time we cannot say whether the 30-45 Hz response is a coherent oscillatory phenomenon, or whether it is composed instead of a complex of potentials brought out by strong somatosensory stimulation, some of which may relate motor responses elicited by the US rather to plastic changes induced by this event. The time course of the 30-45 Hz response to the US is compatible with a role in plasticity, but other functions could be served.

There is reason to inquire into a possible functional role for brain oscillatory responses in learning, even though the presence of such responses has not been well documented in learning experiments at the present time. Coherent oscillatory activity in the gamma band range (20-50 Hz) has been recorded from visual and somatosensory cortex of animal preparations during perceptual tasks (Singer et al., 1990; Murthy & Fetz, 1992) and from human subjects by means of neuromagnetic recordings taken during the delivery of brief auditory (Pantev et al., 1991) or somatosensory stimuli (Kaukoranta & Reinikainen 1985). It has been suggested that oscillatory phenomena of this type may segregate and bind sensory features that define objects or events in visual, auditory, or somatosensory fields (Pantev et al., 1991; Singer et al., 1990). Oscillatory activity that extends beyond the duration of a stimulus might also support the organization of limb or eye movements which are appropriate for a stimulus, insofar as priming of feature detectors for kinesthetic patterns by top-down processing is a mechanism by which the brain-defines and executes behavioral responses (Brener, 1986). In principle, oscillatory patterns

could serve multiple functional roles in adapting an organism to its circumstances. For example, if oscillating assemblies of neurons code sensory, motor, or other brain events that occur in the experience of an organism, such events can subsequently become assimilated into the prevailing context and into the stream of the organism's activities only by virtue of some effect of their oscillatory representations on synaptic conductivities. A role in learning is thus implied which is not incompatible with concurrent roles in perceptual identification and\or response selection.

Animal experiments may help to inform us further on the role of oscillatory activity and P300-like events in information processing. In particular, our model emphasizes the role of the corticothalamic backprojection in associative processes. One goal of these studies is therefore to see how neural activity in corticothalamic pathways is altered by the occurrence of unpredicted (surprising) events in the environment. We are also exploring whether corticothalamic activity produced by stimulus novelty is related to P300-like waves and oscillatory rhythms in different structures of the rat brain. Corticothamic feedback could be a source of oscillatory activity and promote encoding in highly selected cortical networks, if such feedback were to switch thalamic relay neurons in the forward path into a bursting mode. Were firing of relay cells to be sustained, perhaps by a combination of continued sensory input and corticothalamic feedback acting through specific reticular neurons (these neurons having inhibitory effects on relay cells), other thalamocortical circuits might be brought into the picture, to the extent that these circuits are similarly being driven by unpredicted events in a task situation. Repetitive firing has been documented in thalamic neurons (Llinas & Geijo-Barrientos, 1989; Steriade et al., 1991) and might relate to these processes, although the mechanism and functional role of thalamic oscillatory activity can only be speculated about at this time (Llinas, 1992).

Acknowledgement

Preparation of this paper was funded by a grant from the Natural Sciences and Engineering Research Council of Canada (OGP0000132).

REFERENCES

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

Basar, E., & Bullock, T.H. (1992). *Induced rhythms in the brain.* Boston MA: Birkhauser.

Bechtereva N.P., Abdullaev, Y.G., and Medvedev, S.V. (1992). Properties of neuronal activity in cortex and subcortical nuclei of the human brain during single-word processing. *Electroencephalography and Clinical Neurophysiology.* 82, 296-301.

Bienenstock, E.L., Cooper, L.N., & Munro, P.W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience,* 2, 32-48.

Birbaumer, N., Elbert, T., Canavan, A., and Rockstroh, B. (1990). Slow cortical potentials of the cerebral cortex and behavior. *Physiological Reviews,* 70, 1-41.

Birbaumer, N., Roberts, L.E., Lutzenberger, W., Rockstroh, B., & Elbert, T. (1992). Area-specific self-regulation of slow cortical potentials on the sagittal midline and its effects on-behavior. *Electroencephalography and Clinical Neurophysiology,* 84, 353-361.

Braitenberg, V., & Schuz, A. (1991). *Anatomy of the cortex: Statistics and geometry.* Berlin: Springer-Verlag.

Bramwell, L. (1993). *The role of visual task requirements in auditory feedback learning.* Honours B.A. thesis. McMaster University.

Brener, J. (1986). Operant reinforcement, feedback, and the efficiency of learned motor control. In M.G.H. Coles, E. Donchin, & S.W. Porges (Eds.), *Psychophysiology: Systems, processes, and applications* (pp. 309-327). New York: Guilford Press.

Brown, T.H., & Kariss, E.W., & Keenan, C.L. (1990). Hebbian synapses: Biophysical mechanisms and algorithnis. *Annual review of Neuroscience,* 13, 475-571.

Chwilla, D.J., & Brunia, C.H.M. (1991). Event-related potentials to different feedback stimuli. *Psychophysiolo,gy,* 28,123-132.

Colwill, R.M. (1993). An associative analysis of instrumental learning. *Current Directions in Psychological Science,* 2,111-116.

Dawson, M.E., & Diferno, M.A. (1973). Concurrent measurement of awareness and electrodermal classical conditioning. *Journal of Experimental Psychology,* 101, 82-86.

Dawson, M.E. & Schell, A.M. (1987). Human autonomic and skeletal classical conditioning: The role of conscious cognitive factors. In G. Davey (Ed.), *Cognitive processes and Pavlovian conditioning in humans* (pp. 27-55). New York: Wiley & Sons.

Deutsch, J.A. (1960). *The structural basis of behavior.* Cambridge: Cambridge University Press.

Douglas, RJ., & Martin, K.A.C. (1990). Neocortex. In G.M. Shepherd (Ed.), The *synaptic organization of the brain* (pp. 389-438). Oxford, UK: Oxford University Press.

Flor, H., Birbaumer, N., Roberts, L.E., Feige, E., Lutzenberger, W., Fuerst, M., & Hermann, C. *(1994). Electrocortical correlates of Pavlovian conditioning.* (submitted for publication).

Gaillard, A.W.K. (1986). The CNV as an index of response preparation. In W.C. McCallum, R. Zappoll, & F. Denoth (Eds.), Cerebral psychophysiology: Studies in event related potentials. *Electroencephalography and Clinical Nwrophystology, 38,* (Supplement), pp. 196-206.

Germana, J. (1968). Psychophysiological correlates of conditioned autonon-de response formation. *Psychological Bulletin,* 70,105-114.

Gerren, R., & Weinberger, N.M. (1983). Long term potentiation in the magnocellular medial geniculate nucleus of the anesthetized cat. *Brain Rejearch,* 265, 138-142.

Gluck, M.A., & Meyers, C.E. (1993). Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus* (in press).

Gluck, M.A., Goren, 0., Meyers, *C.,* & Thompson, R.F. (1993). A higher-order recurrent network model of the cerebellar structures of response timing in motor-reflex conditioning. *Journal of Cognitive Neuroscience* (in press).

Grossberg, S. (1987). *The adaptive brain.* Amsterdam: North Holland.

Halgren, E., Stapleton, J.M., Smith, M., and Altafullah, I. (1986). Generators of the human scalp P3(s). In R.Q. Cracco and I. Bodis-Wollner (Eds.), *Evoked Porentials,* pp. 269-284. New York NY: Alan R. Liss.

Hancock, P.J.B., Smith, L.S., & Phillips, W.A. (1991). A biologically supported error correcting learning rule. *Neural Computing,* 3, 201-212.

Hasselmo, M.E., & Bower, J.E. (1992). Cholinergic suppression specific to intrinsic not afferent fiber synapses in rat piriform (olfactory) cortex. *Journal of Neurophysiology,* 67, 1222-1229.

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

Hermstein, RJ. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior,* 13, 243-266.

Holland, P. C., & Rescorla, R.A. (1982). Behavioral studies of associative learning in animals. *Annual Review of Psychology,* 33, 265-308.

Hughes, D.E., & Roberts, L.E. (1985). Evidence of a role for response plans and self-monitoring in biofeedback. *Psychophysiology,* 22, 427-439.

Jenkins, H.M., & Sainsbury, R.S. (1970). Discrimination learning with the distinctive feature on positive or negative trials. In D. Mostofsky (Ed.), *Attention: Contemporary theory and analysis (pp.* 239-273). New York: Appleton Century Crofts.

Johnson, R. Jr. (1986). A triarchic model of P300 amplitude. *Psychophysiology*, 23, 367-384.

Kandel, E.R., Schwartz, J.H., & Jessell, T.M. (1991). *Principles of neural science* (third edition). Norwalk, CN: Appleton & Lange.

Kaukoranta, E., & Reinikainen, K. (1986). *Somatosensory evoked magnetic fields from SI: An interpretation of the spatiotemporal field pattern and effects of stimulus repetition rate.* Report Number TKK-F-A581, Helsinki University of Technology Low Temperature Laboratory. SF-02150, Espoo, Finland.

Konorski, J. (1967). *Integrative activity of the brain.* Chicago: University of Chicago Press.

Ushley, K.S. (1950). In search of the engram. *Society of Experimental Biology Symposium,* 4, 454-482.

Llinas, R.R. (1992). Oscillations in CNS neurons: A possible role for cortical interneurons in the generation of 40-Hz oscillations. In E. Basar & T.H. Bullock (Eds.), *Induced rhythms in the brain* (pp. 269-283). Boston MA: Birkhauser.

Llinas, R.R. & Geijo-Barrientos, E. (1989). In vitro studies of mammalian thalaimic and reticularic thalanile neurons. In M. Bentivoglio & R. Spreaafico (Eds.), *Cellular Thalamic Mechanisms.* Amsterdam NE: Elsevier.

Lynch, G., & Granger, R. (1992). Variations in synaptic plasticity and types of memory in cortico-hippocampal networks. *Journal of Cognitive Neuroscience,* 4.189-198.

Mackintosh, NJ. (1983). *Conditioning and associative learning.* Oxford, UK: Oxford University Press.

McCormick, D.A. (1990). Membrane properties and neurotransniitter actions. In G.M. Shepherd (Ed.), *The synaptic organization of the brain* (pp. 32-66.) Oxford, UK: Oxford University Press.

MeNaughten, B., & Nadel, L (1990). Hebb-Marr networks and the neurobiological representation of action in space. In M Gluck & D. Rumelhart (Eds). *Neuroscience and connectionist theory* (pp. 2-63). Hilisdale NJ: Lawrence Erlbaum.

Miller, R.R., & Barnet. R.C. (1993). The role of time in elementary associations. *Current Directions in Psychological Science.* 2, 106-111.

Murthy, V.N, & Fetz, E.E. (1992). Coherent 25- to 35-Elz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proceedings of the National Academy of Sciences USA,* 89, 5670-5674.

Pantey, C., Makeig, S., Hoke, M., Galambos, R., Hampson, S., & Gallen, C. (1991). Human auditory evoked gamma-band magnetic fields. *Proceedings of the National Academy of Science USA,* 88, 8996-9000.

Peterhans, E., & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience,* 9, 1749-1763.

Pearce, J.M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review,* 87, 532-552.

Racine, RJ., Milgram, N.W., & Hafner, 5. (1983). Long-term potentlation phenomena in the rat limbic forebrain. *Brain Research,* 260, 217-231.

Racine, RJ., Wilson, D., Teskey, O.C., & Milgram. N.W. (1994). Post-activation potentiation in the neocortex: I. Acute preparations. *Brain Research,* 637, 73-82.

Racine, RJ., Teskey, G.C., Wilson, D., & Seidlitz, E. (1994). Post-activation potentiation and depression in the neocortex of the rat: II. Chronic preparations. *Brain Research,* 627, 83-96.

Rescorla. R.A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist,* 43, 151-160.

Rescorla, R.A. (1987). A Pavlovian analysis of goal-directed behavior. *American Psychologist*, 42, 119-129.

Rescorla, R.A.. & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforeement. In A.H. Black & W.F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pages 64-99). New York: Appleton Century Crofts.

Roberts, L.E.. Rau, H., Lutzenberger, W., & Birbaumer, N. (1994). Mapping P300 Waves onto inhibition: Go/No-Go discrimination. *Electroencephalography and Clinical Neurophyliology,* 92, 44-55.

Roberts, L.E (1990). Evidence for a general associative process in Pavlovian and instrumental conditioning. In H. Lachnit (Chair). *Recent Advances in Pavlovian Conditioning*, Fifth International Congress of Psychophysiology, Budapest, Hungary.

Roberts, L.E., Preston, D., & Uttl, B (1991). *Tuning and capacity limitations in feedback (instrumental) learning.* Psychonomics, San Francisco, CA.

Roberts, L.E., Racine, R.J., & Durlach, P.J. (1992) *A macroarchitecture for associative learning*. Symposium on the Organization of Learning and Memory, Sixth International Congress of Psychophysiology, Berlin, Germany.

Rockstroh, B., Muller, M., Cohen, R., and Elbert, T. (1992). Probing the functional brain state during P300-evocation. *Journal of Psychophysiology,* 6, 175-184.

Schmajuk, N.A., & DiCarlo, J.J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. *Psychological Review,* 99, 268-305.

Singer, W., Gray, C., Engel, A., Konig, P., Artola, A., & Brocher. S. (1990). Formation of cortical cell *assemblies. Cold Spring Harbor Symposia on Quantitative Biology,* 55, 939-952.

Steriade, M., Curro Dossi, R., Pare,,D., & Oakson., G. (1991). Fast oscillations (20-40 Hz) in thalamocortical systems and their potentlation by mesopontine cholinergic nuclei in the cat. *Proceedings of the National Academy of Sciences USA,* 88,4396-4400.

Wagner, A.R., & Brandon, S.E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In S.B. Klein & R.R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 149-189). Hamale, NJ. Lawrence Erlbaum.

Wagner, A.R., Logan, F.A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology,* 76, 171-180.

Weisman, R.G. (1977). On the role of the reinforcer in associative learning. In H.Davis and H.M.B. Hurwitz (Eds.), *Operant-Pavlovian interactions* (pp. 1-22). Hillsdale. NJ: Lawrence Erlbaum.

Williams, B.A., & Heyneman, N. (1982). Multiple determinants of 'blocking' effects on operant behavior. *Animal learning and behavior,* 10, 72-76.

Woodward, S.H., Brown, W.S., Marsh, J.T., and Dawson, M.E. (1991). Probing the time-course of the auditory oddball P3 with secondary reaction lime. *Psychophysiology,* 28, 609-618.

Zipser, D., & Rumelhart, D.E. (1990). The neurobiological significance of the new learning models. In E.L Schwartz (Ed.), *Computational neuroscience* (pp. 192-200). Cambridge, MA: MIT Press.