A NEURAL NETWORK MODEL OF
ADAPTIVELY TIMED REINFORCEMENT LEARNING
AND HIPPOCAMPAL DYNAMICS

Stephen Grossberg and John W. L. Merrill

January, 1992


Permission to copy without fee all or part of this material is granted provided that: 1. the copies are not made
or distributed for direct commercial advantage, 2. the report title, author, document number, and release
date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER
FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy
otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1991

Boston University Center for Adaptive Systems and
Department of Cognitive and Neural Systems
111 Cummington Street
Boston, MA 02215
A NEURAL NETWORK MODEL OF ADAPTIVELY TIMED REINFORCEMENT LEARNING AND HIPPOCAMPAL DYNAMICS

Stephen Grossberg† and John W.L. Merrill‡
Center for Adaptive Systems and Department of Cognitive and Neural Systems
Boston University, 111 Cummingston Street, Boston, MA 02215

ABSTRACT

A neural model is described of how adaptively timed reinforcement learning occurs. The adaptive timing circuit is suggested to exist in the hippocampus, and to involve convergence of dentate granule cells on CA3 pyramidal cells, and NMDA receptors. This circuit forms part of a model neural system for the coordinated control of recognition learning, reinforcement learning, and motor learning, whose properties clarify how an animal can learn to acquire a delayed reward. Behavioral and neural data are summarized in support of each processing stage of the system. The relevant anatomical sites are in thalamus, neocortex, hippocampus, hypothalamus, amygdala, and cerebellum. Cerebellar influences on motor learning are distinguished from hippocampal influences on adaptive timing of reinforcement learning. The model simulates how damage to the hippocampal formation disrupts adaptive timing, eliminates attentional blocking, and causes symptoms of medial temporal amnesia. It suggests how normal acquisition of subcortical emotional conditioning can occur after cortical ablation, even though extinction of emotional conditioning is retarded by cortical ablation. The model simulates how increasing the duration of an unconditioned stimulus increases the amplitude of emotional conditioning, but does not change adaptive timing; and how an increase in the intensity of a conditioned stimulus "speeds up the clock", but an increase in the intensity of an unconditioned stimulus does not. Computer simulations of the model fit parametric conditioning data, including a Weber law property and an inverted U property. Both primary and secondary adaptively timed conditioning are simulated, as are data concerning conditioning using multiple interstimulus intervals (ISIs), gradually or abruptly changing ISIs, partial reinforcement, and multiple stimuli that lead to time-averaging of responses. Neurobiologically testable predictions are made to facilitate further tests of the model.

1. Introduction

This article contributes to the development of a behavioral and neurobiological theory of learning and memory. The theory describes how processes of learning, recognition, reinforcement, and recall interact to focus attention upon motivationally desired goals, to

† Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175), DARPA (AFOSR 90-0083), the National Science Foundation (NSF IRI-87-16960), and the Office of Naval Research (ONR N00014-91-J-4100).
‡ Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0128).

Acknowledgements: Thanks to Cynthia Bradford and Diana Meyers for their valuable assistance in the preparation of the manuscript and illustrations.
generate appropriate actions towards these goals, and to regulate selective forgetting of environmental contingencies that no longer predict behavioral success. Recent contributions to the theory are found in Grossberg (1987a), Grossberg and Levine (1987), Grossberg and Schmajuk (1987, 1989), Levine and Prueitt (1989, 1991) and Schmajuk and Di Carlo (1991). Although derived from postulates aimed at explaining vertebrate behavior, the theory has also been applied to explain neurobiological data concerning classical conditioning of the invertebrate Aplysia (Buonomano, Baxter, and Byrne, 1990). Other empirically supported predictions and as-yet-untested predictions of the theory are reviewed in Carpenter and Grossberg (1991a, 1991b) and Grossberg (1991). These models have also been incorporated into the control architecture of freely moving adaptive robots for use in technology (Baloch and Waxman, 1991a, 1991b).

The present article further develops a part of the theory, introduced in Grossberg and Schmajuk (1989), which analyses how recognition events control motivated behaviors that are adaptively timed. Several different types of brain processes organize the temporal unfolding of serial order in behavior. The present model instantiates one type of timing control, called spectral timing, and shows how it can modulate the course of recognition learning, reinforcement learning, and the timed onset of a goal-oriented action. The model’s formal processing stages are also compared with anatomical, neurophysiological, and biochemical data about several brain regions, notably the hippocampal formation.

1.2 Timing the Balance between Exploration for Novel Rewards and Consumption of Expected Rewards

The spectral timing model clarifies the following type of behavioral competence. Many goal objects may be delayed subsequent to the actions that elicit them, or the environmental events that signal their subsequent arrival. Humans and many animal species can learn to wait for the anticipated arrival of a delayed goal object, even though its time of occurrence can vary from situation to situation. Such behavioral timing is important in the lives of animals which can explore their environments for novel sources of gratification. On the one hand, if an animal does not inhibit its exploratory behavior, then it will starve to death by restlessly moving from place to place, unable to remain in one place long enough to obtain food there. On the other hand, if an animal inhibits its exploratory behavior for too long while waiting for an expected source of food to materialize, then it will starve to death if food is not, after all forthcoming.

Thus the animal’s task is to accurately time the expected delay of a goal object based upon its previous experiences in a given situation. It needs to balance between its exploratory behavior aimed at searching for novel sources of reward, and its consummatory behavior aimed at acquiring expected sources of reward. To effectively control this balance, the animal needs to be able to suppress its exploratory behavior and focus its attention upon an expected source of reward at around the time that the expected delay transpires for acquiring the reward.

1.3 Distinguishing Expected Nonoccurrences from Unexpected Nonoccurrences: Inhibiting the Negative Consequences of Expected Nonoccurrences

This type of timing calibrates the delay of a single behavioral act, rather than the organization of a correctly timed and speed-controlled sequence of acts. Suppose, for example, that an animal typically receives food from a food magazine two seconds after pushing a lever, and that the animal orients to the food magazine right after pushing the lever. When the animal inspects the food magazine, it perceives the nonoccurrence of food during the subsequent two seconds. These nonoccurrences disconfirm the animal’s sensory expectation that food will appear in the magazine. Because the perceptual processing cycle that processes this sensory information occurs at a much faster rate than two seconds, it can compute this sensory disconfirmation many times before the two second delay has elapsed.
The central issue is: What spares the animal from erroneously reacting to these expected nonoccurrences of food during the first two seconds as predictive failures? Why does the animal not immediately become so frustrated by the nonoccurrence of food that it shifts its attentional focus and releases exploratory behavior aimed at finding food somewhere else? Alternatively, if the animal does wait, but food does not appear after the two seconds have elapsed, why does the animal then react to the unexpected nonoccurrence of food by becoming frustrated, shifting its attention, and releasing exploratory behavior?

Grossberg and Schmajuk (1989) argued that a primary role of the timing mechanism is to inhibit, or gate, the process whereby a disconfirmed expectation would otherwise negatively reinforce previous consummatory behavior, shift attention, and release exploratory behavior. The process of registering sensory mismatches or matches is not itself inhibited; if the food happened to appear earlier than expected, the animal could still perceive it and eat. Instead, the effects of these sensory mismatches upon reinforcement, attention, and exploration are inhibited.

One of the main tasks of the present work is to show how processes such as adaptive timing, reinforcement learning, attention, and motor learning differ, yet are linked in the control of behavior. Thus the exposition needs to describe several different types of circuits that form part of a larger neural system. These results were announced in Grossberg and Merrill (1991). Part I, Section 2, summarizes data concerning timed learning of the rabbit nictitating membrane response and the pigeon FI scallop. These data are used in Section 3 to suggest how the model solves a problem called the Timing Paradox. Section 4 describes the new spectral timing model, and illustrates its processes with computer simulations. Sections 5–9 describe how the model explains some difficult parametric conditioning data, notably data about secondary conditioning and the effects of changing stimulus intensity or duration. Sections 10–15 interpret the adaptive timing mechanism in terms of interactions between dentate granule cells and CA3 pyramidal cells in the hippocampus, notably at NMDA receptors. Neurobiological data in support of this hypothesis are summarized and new predictions made. Sections 16–22 summarize computer simulations that show how the model can replicate quantitative properties of data from several types of conditioning experiments. All of these computer simulations use a single set of parameters. Robustness of the model's properties is also demonstrated using different sets of parameters.

Part II, beginning with Section 23, shows how the spectral timing model may be embedded into a larger neural system for the control of recognition learning, reinforcement learning, and motor learning. These sections also summarize behavioral and neural data in support of each processing stage of this model system. The relevant anatomical sites are in thalamus, neocortex, hippocampus, hypothalamus, amygdala, and cerebellum. The behavioral data include an explanation of blocking in normal animals, elimination of blocking in hippocampectomized animals, impairing timing in hippocampectomized animals, medial temporal amnesia in hippocampectomized animals, subcortical fear conditioning and abnormal fear extinction in animals with cortical lesions, and disruption of motor learning by cerebellar lesions. Various of these data were reported after the corresponding model stages were published. Such data illustrate the predictive power of the theory. No claim is made that all neural processes are modeled in this system. Rather, the system is a lumped model that attempts to provide a minimal representation of the processes that are rate-limiting in explaining the targeted data bases. The present work may be viewed as a step in the progressive unlumping of the model to analyse ever finer neural processing stages.

PART I
SPECTRAL TIMING

2. Examples of Spectral Timing: Conditioning the Nictitating Membrane Response and the FI scallop
A well-studied example of spectral timing is the conditioning of the rabbit nictitating membrane response. Rabbits, like many mammals, have a translucent sheet of tissue called a "nictitating membrane", that acts as a third eyelid. The nictitating membrane response, which extends this sheet across the eye, can be classically conditioned. For example, a conditioned stimulus, or CS, can be paired with noxious unconditioned stimulus, or US, such as a periorbital shock or airpuff, that elicits membrane extension. Smith (1968) studied the effect of manipulating the time lag between CS onset and US onset. This lag is called the interstimulus interval, or ISI. The CS was a 50 msec tone and the US was a 50 msec electric shock. The ISI values were 125, 250, 500, and 1000 msec. The fact that conditioning occurred at ISI's much larger than the CS duration implies that an internal trace of the CS is stored in short term memory (STM) subsequent to CS offset. Because an internal CS trace is needed to bridge the ISI gap between CS offset and US onset, such a paradigm is called trace conditioning, to distinguish it from the delay conditioning paradigm wherein the CS and US overlap in time.

Figure 1

Smith (1968) found that the conditioned response, measured as percentage of responses and response amplitude, was determined by both ISI and US intensity, whereas response onset rate and peak time were determined by the ISI essentially independently of US intensity. An increase in the mean of the peak response time correlated with an increase in the variance of the response curve, for each ISI (Figure 1). Grossberg and Schmajuk (1989) called this the Weber law property.

Figure 2

Figure 2 summarizes a computer simulation by Grossberg and Schmajuk (1989) of the Weber law property. The data and computer simulation in Figures 1 and 2 clarify why this type of timing is called spectral timing: A temporally sensitive "spectrum" of activations exists, that is densely distributed across all finite ISIs up to some maximum, and that can be tuned by learning to enhance those spectral components that cluster around the experienced set of ISIs. Another example of spectral timing arises in appetitive instrumental conditioning with a fixed delay to reward in both rats (Roberts, 1981; Meck and Church, 1987) and pigeons (Roberts, Cheng, and Cohen, 1989). In these experiments, animals were rewarded for the first lever press (in rats) or key peck (in pigeons) that occurred a fixed time after a signal was presented (Figure 3). A characteristic pattern of responding evolved, referred to as the FI scallop (Mackintosh, 1974). Examples of such scallops are shown in Figure 4.

3. The Timing Paradox and a Solution

The Weber law property of a spectral timing model provides a way for an animal to distinguish between the expected and unexpected nonoccurrences that were discussed in Section 1, without losing the capacity to time its conditioned responses. The Timing Paradox described in this section clarifies why this is a nontrivial problem. The Timing Paradox comprises the following, apparently contradictory, pair of constraints. On the one hand, in response to any fixed choice of conditionable ISI, the learned optimal response delay approximates the ISI. Thus a model of adaptive timing must accurately discriminate between individual temporal delays. On the other hand, expected nonoccurrences throughout the ISI should not be treated as predictive failures. Thus the inhibitory signal that prevents this from happening must be distributed throughout the ISI. How can a timing model both be sharply enough tuned to precisely learn the ISI, yet be broadly enough tuned to inhibit orienting responses throughout the entire ISI interval?

A spectral timing model reconciles the two requirements of accurate optimal temporal delay and temporally distributed activation via the Weber law property (Figure 2). According to this property, the breadth of the model's temporal response scales with its peak time.
Consequently the onset of the CS causes the immediate initiation of an output signal which is sustained throughout the entire ISI, but the peak output of the signal is accurately located at the expected arrival time of the US. A solution of the Timing Paradox is achieved by assuming that the output of the spectral timing model obeys the Weber law property, and that this output signal is used to inhibit the expression of a predictive failure. This inhibition occurs throughout the ISI interval, yet timed output peaks at the ISI. Different pathways carry the inhibitory signal that blocks expression of predictive failure, and the excitatory signal that energizes timed responding. These pathways will be characterized with increasing precision in Section 4 and Part II.

**4. START: A Unified Model of Adaptive Timing and Conditioned Reinforcer Learning.**

The new adaptive timing model will now be defined. It combines Spectral Timing mechanisms with mechanisms from Adaptive Resonance Theory (see Part II). Hence it is called the START model. The START model builds upon a previous model of reinforcement learning whose processing stages are compared with behavioral and neural data below. Here we provide just enough exposition to define the model and to compare its emergent properties with these data.

As illustrated in Section 2, the model uses reinforcement learning experiments, notably classical conditioning experiments, to test its mechanisms. Each sensory event is therefore called a conditioned stimulus, or CS. The i<sup>th</sup> sensory event is denoted by $CS_i$. Event $CS_i$ activates a population of cells that is called the i<sup>th</sup> sensory representation $S_i$ (Figure 3). Another population of cells, called a drive representation $D$, receives a combination of sensory, reinforcement, and homeostatic (or drive) stimuli. Reinforcement learning, emotional reactions, and motivational decisions are controlled by $D$ (Grossberg, 1971). In particular, a reinforcing event, such as an unconditioned stimulus, or US, is capable of activating $D$.

Various authors have invoked representations analogous to drive representations. Bower and his colleagues have called them emotion nodes (Bower, 1981; Bower, Gilligan, and Monteiro, 1981) and Barto, Sutton, and Anderson (1983) have called them adaptive critic elements. During conditioning, presentation of a CS, before a US causes activation of $S_i$ followed by activation of $D$. Such pairing causes strengthening of the adaptive weight, or long term memory trace, in the modifiable synapses from $S_i$ to $D$. This learning event converts $CS_i$ into a conditioned reinforcer. Conditioned reinforcers hereby acquire the power to activate $D$ via the conditioning process.

Figure 5

In the START model, reinforcement learning in $S_i \rightarrow D$ pathways is supplemented by a parallel learning process that is concerned with adaptive timing. As shown in Figure 3, both of these learning processes output to $D$, which in turn inhibits a population of cells called the orienting subsystem. The orienting subsystem is denoted by $A$ because it is a source of nonspecific arousal signals that are capable of initiating frustrative emotional reactions, attention shifts, and orienting responses (see Part II). The inhibitory pathway from $D$ to $A$ is the gate that prevents these events from occurring in response to expected disconfirmations (Section 1).

**A. Limited Capacity Short Term Memory**

The sensory representations $S_i$ compete for a limited capacity, or finite total amount, of activation. Winning populations are said to be stored in short term memory, or STM. The competition is carried out by an on-center off-surround interaction among the populations $S_i$. The property of STM storage is achieved by using recurrent, or feedback, pathways among the populations. A tendency to select winning populations is achieved by using
membrane equations, or shunting interactions, to define each population’s activation, and a proper choice of feedback signals between populations (Grossberg, 1973, 1982a). Expressed mathematically, each CS$_i$ activates an STM representation $S_i$ whose activity $S_i$ obeys the shunting on-center off-surround competitive feedback equation:

$$\frac{d}{dt}S_i = -\alpha_A S_i + \beta_A (1 - S_i)(I_i(t) + f_S(S_i)) - \gamma_A S_i \sum_{k \neq i} f_S(S_k).$$

In (1), $I_i(t)$ is the input that is turned on by presentation of CS$_i$. Term $-\alpha_A S_i$ describes passive decay of activity $S_i$. Term $\beta_A (1 - S_i)(I_i(t) + f_S(S_i))$ describes the excitatory effect on $S_i$ of the input $I_i(t)$ and the feedback signal $f_S(S_i)$ from population $S_i$ to itself. Activity $S_i$ can continue to grow until it reaches the excitatory saturation point, which is scaled to equal 1 in (1). Term $-\gamma_A S_i \sum_{k \neq i} f_S(S_k)$ describes inhibition of $S_i$ by competitive signals $f_S(S_k)$ from the off-surround of populations $k \neq i$. Figure 6 summarizes a computer simulation of how a brief CS$_i$ gives rise to a sustained STM activation $S_i$, which is partially inhibited by competition from $S_i$’s activation in response to a US. The signal function $f_S$ may be chosen to have any of the forms depicted in Figure 7 without qualitatively altering model properties. In this article, the simple rectification function

$$f(w) = [w - \mu]^+ \equiv \max(w - \mu, 0)$$

of Figure 7a is used, except in equation (8) below, which uses a sigmoid signal function as in Figure 7b.

Figure 6

Figure 7

B. Drive Representation

The computer simulations reported herein use only a single drive representation $D$. Explanations of data arising from competing drive representations are discussed in Grossberg (1984, 1987a). The activity $D$ of the drive representation $D$ obeys the equation

$$\frac{d}{dt}D = -\alpha_D D + \beta_D \sum_i f_D(S_i)C_i + \gamma_D R.$$  

In (3), term $-\alpha_D D$ describes the passive decay of activity $D$. Term $\beta_D \sum_i f_D(S_i)C_i$ describes the total excitatory effect of all the sensory representations $S_i$ on $D$. In this term, the signal function $f_D$ is chosen as in (2), and $C_i$ is the adaptive weight, or long term memory (LTM) trace, in the pathway from the sensory representation $S_i$ of $CS_i$ to the drive representation $D$. This LTM trace is denoted by $C_i$ because its size measures how well $S_i$ can activate $D$, and thus how $CS_i$ has become a conditioned reinforcer through learning. Because $C_i$ multiplies $f_D(S_i)$, a large activation of $S_i$ will have a negligible effect on $D$ if $C_i$ is small, and a large effect on $D$ if $C_i$ is large. Term $\gamma_D R$ describes the total output of the spectral timing circuit to $D$. Output $R$ is defined in (11).

Figure 8

Figure 8c summarizes a computer simulation in which the activity $D$ responds to CS and US signals after 50 conditioning trials. Figures 8a and 8b summarize the corresponding STM traces $S_1$ of the CS and $S_0$ of the US, respectively.
C. Conditioned Reinforcement

The adaptive weight $C_i$ that calibrates conditioned reinforcement obeys a gated learning law (Grossberg, 1969b):

$$\frac{d}{dt} C_i = \alpha_C S_i (-C_i + \beta_C (1 - C_i) f_C(D)). \tag{4}$$

Learning by $C_i$ is turned on and off by the signal $S_i$ from $S_i$, which thus acts like a learning gate, or modulator. Once turned on, $C_i$ performs a time-average of activity at the drive representation $D$ via the signal $f_C(D)$, which is chosen as in (2). Activity $C_i$ cannot exceed the finite value 1, due to the shunting term $1 - C_i$. The value of $C_i$ can both increase and decrease during the course of learning.

The remaining equations of the model describe the adaptive timing process.

D. Now Print Signal

The Now Print signal $N$ is turned on by sufficiently large and rapid increments in the activity $D$ of $D$. Signal $N$ turns on the learning process that encodes adaptive timing information. The transient signal $N$ is derived from the sustained activity $D$ by the action of a slow inhibitory interneuron (Figure 9). The transformation from sustained activity $D$ to transient activity $N$ can be realized mathematically by the function

$$N = [f_C(D) - E - \epsilon]^+. \tag{5}$$

Figure 9

In (5), $E$ is the activity of an inhibitory interneuron that time-averages $f_C(D)$, as in equation

$$\frac{d}{dt} E = \alpha_E (-E + f_C(D)), \tag{6}$$

before inhibiting the direct excitatory signal $f_C(D)$. Equation (5) means that $N = 0$ if $f_C(D) - E \leq \epsilon$, and $N = f_C(D) - E - \epsilon$ if $f_C(D) - E > \epsilon$. Figures 6d and 7c illustrates how $N$ responds to increments in $D$. An important property of $N$ is that it increases in amplitude, but not significantly in duration, in response to larger inputs $f_C(D)$.

E. Activation Spectrum

The START model is said to control “spectral” timing because each drive representation $D$ is associated with a population of cell sites whose members react at a spectrum of rates $r_j$. Neural populations whose elements are distributed along a temporal or spatial parameter are familiar throughout the nervous system. Two examples are populations of spinal cord cells that obey the size principle (Henneman, 1957, 1985), and the spatial frequency-tuned cells of the visual cortex (Jones and Keck, 1978; Musselwhite and Jeffreys, 1985; Parker and Salzen, 1977a, 1977b; Parker et al., 1982a, 1982b; Plant et al., 1983; Skrandies, 1984; Vassilev and Strashimirov, 1979; Vassilev et al., 1983; Williamson et al., 1978).

The spectral activities $x_{ij}$ that are associated with drive representation $D$ and activated by sensory representation $S_i$ obey the equation

$$\frac{d}{dt} x_{ij} = r_j (-x_{ij} + (1 - x_{ij}) f_x(S_i)), \tag{7}$$
where \( f \) satisfies (2). By (1) and (7), presentation of \( CS_i \) to \( S_i \) via an input \( I_i \) generates an output signal \( f(x(S_i)) \) that activates the local potentials \( x_{ij} \) of all cell sites in the target population. The potentials \( x_{ij} \) respond at rates proportional to \( r_j, j = 1, 2, \ldots, n \). These potentials activate the next processing stage via signals

\[
f(x_{ij}) = \frac{x_{ij}^8}{\delta x_{ij}^8 + x_{ij}^8}.
\]  

Signal \( f(x_{ij}) \) is a sigmoid function of activity \( x_{ij} \). Figure 10a shows the activation spectrum \( f(x_{ij}(t)) \) that arises from presentation of \( CS_i \) to \( S_i \) via input \( I_i \) in (1), using a choice of rate parameters \( r_j \) in (20) which range from 10 (fast) to 0.0025 (slow). The method by which the simulations were performed is described in the Appendix.

**Figure 10**

**F. Habituation Transmitter Spectrum**

Each spectral activation signal \( f(x_{ij}) \) interacts with a habituative chemical transmitter \( y_{ij} \) via the equation

\[
\frac{d}{dt} y_{ij} = \alpha_y (1 - y_{ij}) - \beta_y f(x_{ij}) y_{ij}.
\]  

According to equation (9), the amount of neurotransmitter \( y_{ij} \) accumulates to a constant target level 1, via term \( \alpha_y (1 - y_{ij}) \), and is inactivated, or habituates, due to a mass action interaction with signal \( f(x_{ij}) \), via term \( -\beta_y f(x_{ij}) y_{ij} \). The different rates \( r_j \) at which each \( x_{ij} \) is activated causes the corresponding \( y_{ij} \) to become habituated at different rates. A habituation spectrum is thereby generated. The signal functions \( f(x_{ij}(t)) \) in Figure 10a generate the habituation spectrum of \( y_{ij}(t) \) curves in Figure 10b.

**G. Gated Signal Spectrum**

Each signal \( f(x_{ij}) \) interacts with \( y_{ij} \) via mass action. This process is also called gating of \( f(x_{ij}) \) by \( y_{ij} \) to yield a net signal \( g_{ij} \) that is equal to \( f(x_{ij}) y_{ij} \). Each gated signal \( g_{ij}(t) = f(x_{ij}(t)) y_{ij}(t) \) has a different rate of growth and decay, thereby generating the gated signal spectrum shown in Figure 10c. In these curves, each function \( g_{ij}(t) \) is a unimodal function of time, where function \( g_{ij}(t) \) achieves its maximum value \( M_{ij} \) at time \( T_{ij} \). \( T_{ij} \) is an increasing function of \( i \), and \( M_{ij} \) is a decreasing function of \( j \).

These laws for the dynamics of a chemical transmitter were described in Grossberg (1968, 1969a). They capture the simplest first-order properties of a number of known transmitter regulating steps (Cooper, Bloom, and Roth, 1974), such as transmitter production (term \( \alpha_y \)), feedback inhibition by an intermediate or final stage of production on a former stage (term \( -\alpha_y y_{ij} \)), and mass action transmitter inactivation (term \( -\beta_y f(x_{ij}) y_{ij} \)). Alternatively, they can be described as the voltage drop across an RC circuit, or the current flow through an appropriately constructed transistor circuit. These properties are sufficient to explain the article's targeted data, so finer transmitter processes, such as transmitter mobilization effects, are not considered herein.

**H. Spectral Learning Law**

Learning of spectral timing obeys a gated steepest descent equation
\[
\frac{d}{dt}z_{ij} = \alpha z f(x_{ij})y_{ij}(-z_{ij} + N),
\]  

where \( N \) is the Now Print signal of (5). Each long term memory (LTM) trace \( z_{ij} \) in (10) is activated by its own sampling signal \( g_{ij} = f(x_{ij})y_{ij} \). The sampling signal \( g_{ij} \) turns on, or gates, the learning process, and causes \( z_{ij} \) to approach \( N \) during the sampling interval at a rate proportional to \( g_{ij} \). The attraction of \( z_{ij} \) to \( N \) is called steepest descent. Thus (10) is an example of learning by gated steepest descent. Each \( z_{ij} \) changes by an amount that reflects the degree to which the curves \( g_{ij}(t) \) and \( N(t) \) have simultaneously large values through time. If \( g_{ij} \) is large when \( N \) is large, then \( z_{ij} \) increases in size. If \( g_{ij} \) is large when \( N \) is small, then \( z_{ij} \) decreases in size. As in equation (4), \( z_{ij} \) can either increase or decrease as a result of learning.

Associative learning by gated steepest descent was incorporated into neural network models in Grossberg (1969b) and is the learning law that was used to introduce adaptive resonance theory (Grossberg, 1976a, 1976b). An associative learning law of this form was subsequently used by Levy and his colleagues to model their data on hippocampal LTP (Levy, Brassel, and Moore, 1983; Levy and Desmond, 1985). Singer (1983) has also used such a law to model his experiments on adaptive tuning of visual cortical cells during the visual critical period. These experiments support the adaptive resonance theory predictions (Grossberg, 1976a, 1976b) that both hippocampal LTP and feature detector tuning in visual cortex should obey such a learning law.

As noted above, the time interval between CS onset and US onset is called the interstimulus interval, or ISI. Using the spectral learning law (10), the individual LTM traces differ in their ability to learn at different values of the ISI. This is the basis of the network’s timing properties. Figure 11 illustrates how six different LTM traces \( z_i, i = 1, \ldots, 6 \), learn during this simulated learning experiment. The CS and US are paired during 4 learning trials, after which the CS is presented alone on a single performance trial. In this computer simulation, the CS input \( I_{CS}(t) \) remained on for a duration of 0.05 time units on each learning trial. The US input \( I_{US}(t) \) was presented after an ISI of 0.5 time units and remained on for 0.05 time units. The upper panel in each part of the figure depicts the gated signal function \( g_{ij}(t) \) with \( r_j \) chosen at progressively slower rates. The middle panel plots the corresponding LTM trace \( z_{ij}(t) \).

**Figure 11**

I. Doubly Gated Signal Spectrum

The lower panel plots the twice-gated signal \( h_{ij}(t) = f(x_{ij}(t))y_{ij}(t)z_{ij}(t) \). Each twice-gated signal function \( h_{ij}(t) \) registers how well the timing of CS and US is learned and read-out by the \( i^{th} \) processing channel. In Figure 11d, where the once-gated signal \( g_{ij}(t) \) peaks at approximately the ISI of 0.5 time units, the LTM trace \( z_{ij}(t) \) shows the maximum learning. The twice-gated signal \( h_{ij}(t) \) also shows a maximal enhancement due to learning, and exhibits a peak of activation at approximately 0.5 time units after onset of the CS on each trial. This behavior is also generated on the fifth trial, during which only the CS is presented.

J. Output Signal

The output of the network is the sum of the twice-gated signals \( h_{ij}(t) \) from all the spectral components corresponding to all the CS. Thus
\[ R = \sum_{i,j} f(x_{ij})y_{ij}z_{ij}. \]

The output signal computes the cumulative learned reaction of the whole population to the input pattern. Figure 10c shows the function \( R \) derived from the \( h_{ij} \) shown in Figure 10d. A comparison of Figures 10c-e illustrate how the output \( R(t) \) generates an accurately timed response from the cumulative partial learning of all the cell sites in the population spectrum. The once-gated signals \( g_{ij}(t) \) in Figure 10c are biased towards early times. The twice-gated signals \( h_{ij}(t) \) in Figure 10d are biased towards the ISI, but many signals peak at other times. The output \( R(t) \) combines these partial views into a cumulative response that peaks at the ISI.

5. The Problem of Self-Printing during Adaptively Timed Secondary Conditioning

The START model overcomes four types of problems whose solution is needed to explain behavioral and neural data about adaptively timed conditioning. These are the problems of (1) self-printing during adaptively timed secondary conditioning, (2) asymmetric effects of increasing CS or US intensity on timed responding, (3) different effects of US duration on timing than on reinforcement, and (4) combinatorial explosion of network pathways. These problems and their solution by the START model are described below, along with supportive data. Problems (1), (3), and (4) were not solved by the Grossberg and Schmajuk (1989) model.

A major problem for any model of adaptive timing is to explain how adaptively timed secondary conditioning can occur. In primary conditioning, a conditioned stimulus \( CS_1 \) is paired with an unconditioned stimulus US until \( CS_1 \) becomes a conditioned reinforcer. In secondary conditioning, another conditioned stimulus \( CS_2 \) is paired with \( CS_1 \) until it, too, gains reinforcing properties. Various experiments have shown that the conditioned response to \( CS_2 \) can be adaptively timed (Gormezano and Kehoe, 1984; Kehoe, Marshall-Goodell, and Gormezano, 1987). Indeed, Gormezano and Kehoe (1984) claimed that, in their experimental paradigm, “first- and second-order conditioning follow the same laws” (p. 314), although they also acknowledged that some variables may differentially effect first-order and second-order conditioning in other paradigms.

Adaptively timed secondary conditioning could easily erase the effects of adaptively timed primary conditioning in the following way. In order for \( CS_1 \) to act as a conditioned reinforcer, \( CS_1 \) must gain control of the pathway along which the US activates its reinforcing properties. Suppose that \( CS_1 \) activated its sensory representation \( s_1 \) via an input \( (I_{CS_1}) \) pathway and that US expressed its reinforcing properties via an input \( (I_{US}) \) pathway. Also suppose that conditioned reinforcer learning enabled \( CS_1 \) to activate \( I_{US} \). Thereafter, presentation of \( CS_1 \) would simultaneously activate both the \( I_{CS_1} \) pathway and the \( I_{US} \) pathway. This coactivation would create new learning trials for \( CS_1 \) with a zero ISI. In other words \( CS_1 \) could self-print a spectrum with zero ISI due to \( CS_1-CS_1 \) pairing via the \( I_{CS} \) and conditioned \( I_{US} \) pathway. Thus, as \( CS_1 \) became a conditioned reinforcer, it could undermine the timing that it learned through \( CS_1-US \) pairing during primary conditioning. Such self-printing could, for example, occur on secondary conditioning trials when a \( CS_2 \) is followed by a conditioned reinforcer \( CS_1 \).

6. Simulations of Secondary Conditioning

The START model overcomes the self-printing problem with its use of a transient Now Print signal \( N \), as in (5). During primary conditioning, onset of the US causes a brief output burst from \( N \). During secondary conditioning, onset of the conditioned reinforcer \( CS_1 \) also
causes a brief output burst from $N$. However, the spectrum activated by $CS_1$ takes awhile to build up, so essentially all of its activities $x_{ij}$ and sampling signals $f(x_{ij})y_{ij}$ are very small during the brief interval when $N$ is large (Figures 10a and 10c). By the spectral learning law (10), negligible self-printing occurs. The main effect of the self-printing that does occur is to reduce every spectral LTM trace $z_{1j}$ in (10) by a fixed proportion of its value, thus scaling down the size of $R(t)$ without changing the timing of its peak.

Figure 12a depicts the model output $R(t)$ when the Now Print threshold $\epsilon$ in (5) is set to a high enough level to guarantee that no self-printing or secondary conditioning occur. Here $CS_1$ never activates a Now Print signal. Figure 12b shows the output when $\epsilon$ is set lower, thus allowing secondary conditioning and some self-printing to occur. Correct timing still obtains.

Figure 12

Figure 13 shows how the model behaves during secondary conditioning. The left hand half of each panel shows the output of the model in response to the primary conditioned stimulus $CS_1$, and the right hand half of each panel shows the model output in response to the secondary conditioned stimulus $CS_2$. The peak time arising from the presentation of $CS_2$ occurs near the expected time of arrival of $CS_1$, rather than the expected time of the US. This property is consistent with the environment that a model or animal experiences, since the subject never sees $CS_2$ paired with the primal US, but rather sees it paired as a predictor of $CS_1$, which serves as a CR in this context.

Figure 13

7. The Asymmetry between CS and US Processing in Timing Control

Although $CS_1$ can attain properties of a conditioned reinforcer through $CS_1-US$ pairing, this does not imply that all the functional properties of a conditioned reinforcer and an unconditioned stimulus are interchangeable. In fact, increasing the intensity of a conditioned reinforcer $CS_1$ can “speed up the clock” (Maricq, Roberts, and Church, 1981; Meck and Church, 1987; Wilke, 1987), whereas increasing the intensity of a primary US can increase the amplitude of conditioned response, but does not change its timing (Smith, 1968).

The fact that parametric changes of CS and US cause different effects on adaptive timing places a strong constraint on possible mechanisms of how adaptive timing is learned during secondary conditioning. Although the CS acquires reinforcing properties of a US when it becomes a conditioned reinforcer, it does not acquire all of its timing properties. Our proposed solution of the self-printing problem implies an explanation of the different responses caused by an increase in CS intensity or US intensity. This explanation holds even if the $CS_1$ and $US$ sensory representations $S_1$ and $S_0$, respectively, each send signals along the same types of pathways to the drive representation and the adaptive timing circuit. The explanation is summarized below.

An increase in $CS_1$ intensity causes an increase in the amplitude of input $I_1(t)$ in (1). The larger input causes a larger peak amplitude of activity $S_1$ in (1), and a larger signal $f_x(S_1)$ in (7). By (7), the rate with which a spectral activation $x_{1j}$ reacts to signal $f_x(S_1)$ equals $r_j(1 + f_x(S_1))$. Thus an increase in $CS_1$ intensity speeds up the processing of all spectral activations $x_{1j}$. Because $CS_1$ is a conditioned reinforcer, some of its LTM traces $z_{1j}$ are nonzero. Thus the total output $R$ in (11) peaks at an earlier time, and causes the total output $D$ from $D$ in (3) to also peak at an earlier time.

In contrast, a primary reinforcer such as a US does not generate a significant output $R(t)$ from the spectral timing circuit, even if it is allowed to generate a large signal $f_x(S_0)$ to the adaptive timing circuit in (7). This is true because a large US generates a signal
$f_x(S_0)$ to the spectral activations in (7) at the same time that it generates a large signal $f_D(S_0)$ to $\mathcal{D}$ in (3) and a large Now Print signal $N$ in (5). Thus a $US$ creates the conditions of a “zero ISI experiment” for purposes of spectral learning. All the LTM traces $z_{0j}$ in (10) therefore remain very small in response to any number of $US$ representations. An increase in $US$ amplitude thus cannot cause speed-up of the output $R(t)$ in (11), because this output remains approximately zero in response to any $US$ intensity. In summary, the same mechanism that explains how the self-printing problem is avoided also explains why an increase in $CS$ intensity, but not $US$ intensity, speeds up the conditioned response.

The primary effect of an increase in $US$ intensity is to increase the amplitude of the signal $f_D(S_0)$ in (3) to the drive representation $\mathcal{D}$. This causes an increase in the amplitude of $D$ and thus an increase in the amplitude of the conditioned response that is modulated by $D$. This explanation of how a $US$ increases the amplitude of the conditioned response also holds if the $US$ sends no signal $f_x(S_0)$ directly to the adaptive timing circuit. See Grossberg and Schmajuk (1989) for a further discussion of this issue.

Figure 14

8. Different Effects of US Duration on Emotional Conditioning and Adaptive Timing: Sustained and Transient Responses

The existence of a transient Now Print signal $N$ plays a central role in our explanations of how to avoid self-printing during secondary conditioning, and of different effects of $CS$ and $US$ intensity on learned timing. Another type of data lends support to the hypothesis that the activity $D$ and the Now Print signal $N$ both exist but respond to the $US$ in different ways. These data show that an increase in $US$ duration can significantly increase the strength of emotional conditioning (Ashton, Bitgood, and Moore, 1969; Boe, 1966; Borozci, Storms, and Broen, 1964; Church, Raymond, and Beauchamp, 1967; Keehn, 1963; Strouthes, 1965). How can a brief Now Print signal $N$ whose duration does not increase significantly with $US$ duration coexist with emotional conditioning properties that do increase significantly with $US$ duration?

An answer can be given using properties of drive representations $\mathcal{D}$. The activation $D$ of a drive representation by a $US$ does persist longer when the $US$ duration is increased, and does thereby increase the strength of emotional conditioning at the $S \rightarrow \mathcal{D}$ synapses that are modelled by equations (3) and (4); see Grossberg (1972b, Section 4) and Grossberg (1982a) for further discussion of this property. This sustained activation $D$ of a drive representation gives rise to a transient Now Print signal $N$ at a different processing stage - a transient detector - that is downstream from $\mathcal{D}$ itself, as displayed in Figures 5 and 9. Thus $D$ and $N$ represent responses of “sustained cells” and “transient cells” - a distinction familiar from visual perception - which here instantiate different functional properties of emotional conditioning and conditioning of adaptive timing, respectively. The parametric data properties summarized in Sections 6-8 illustrate that the processes of emotional conditioning and adaptive timing, although linked, are not the same. They also support the START model’s proposal of how these processes interact.


According to any spectral timing theory, each $CS_i$ activates a sensory representation $S_i$ that broadcasts signals along many parallel pathways. This can lead to a combinatorial explosion of cell bodies if the spectra are incorrectly instantiated. For example, suppose that each pathway activated a different cell, and that each cell’s activity computed a different $x_{ij}, j = 1, 2, \ldots, n$. Then there would exist as many copies of the spectral timing model as there are sensory representations in the brain. In addition, each spectrum contains 80 activities $x_{ij}$ in our computer simulations. Such a model would require a huge number of
cells to represent a different spectrum for every possible sensory representation. This is, in fact, the type of circuit used in the Grossberg-Schmajuk model.

In the START model, each drive representation, not every sensory representation, has its own spectral cells. Thus the pathways from all sensory representations that correspond to any given drive representation share the same neurons. This modification greatly reduces the number of cells that are needed to achieve spectral timing of arbitrary conditionable CS–US combinations, since there are many fewer drive representations (e.g., for hunger, thirst, sex, etc.) than there are sensory representations. As in Figure 3, each spectrum is computed in parallel with its drive representation. Since the present simulations only consider one type of reinforcer, only one drive representation is depicted. In general, each CS sends an adaptive pathway to every drive representation to which it can be conditioned, as well as adaptive pathways sufficient to sample the corresponding spectral representation. The "coordinates" of each drive representation and its spectrum encode reinforcement and homeostatic variables. In contrast, the CS-activated pathways to these circuits carry signals that reflect the sensory features of the CSs. Thus the fact that different perceptual stimuli may elicit characteristic responses at the cells which represent adaptive timing does not, in itself, imply that these perceptual stimuli are "encoded" at those cells.

It is suggested below how hippocampal cells can form an adaptive timing circuit, and how dendrites of hippocampal pyramidal cells can represent a drive-based spectrum that avoids the combinatorial explosion. A brief historical discussion will first be given to clarify the larger neural context in which this proposal needs to be evaluated.

10. Conditioning and the Hippocampus

Learning within the $S \rightarrow D$ pathways of Figure 3 was predicted in Grossberg (1971, 1975) to have the hippocampal formation as a final common path. It was also predicted that this type of learning is a variant of conditioned reinforcer learning. The distinction between the different learning processes that govern emotional conditioning and adaptive timing was not, however, made in these early articles. In experiments on conditioning the rabbit nictitating membrane response, Berger and Thompson (1978) reported that hippocampal learning does occur, thereby providing partial support for the prediction. At first, these investigators interpreted their results as the discovery of a general neural "engram". Subsequent experiments studied the effects of selective ablations on learning in both hippocampus and cerebellum (McCormick, Clark, Lavond, and Thompson, 1982). These experiments led to the conclusion that hippocampal learning appears to be a variant of the predicted conditioned reinforcer learning, whereas the cerebellum carries out a type of motor learning. Thompson et al. (1984, p. 82) distinguished these two types of learning as "conditioned fear" and "learning of the discrete adaptive response", respectively, a distinction that had also been predicted, and that is elaborated in Part II.

It should also be emphasized that this interpretation of hippocampal function does not contradict other data which implicate the hippocampal formation in the learning of spatial and attentional tasks (Isaacson and Pribram, 1986; O'Keefe and Nadel, 1978). Such a hybrid functional role for hippocampus in conditioned reinforcement, spatial approach and avoidance, and attentional blocking was mechanistically outlined and predicted by the theory's earliest model circuits (Grossberg, 1971, 1975; reviewed in 1987a). In support of such a hybrid function, Eichenbaum and Cohen (1988) have summarized recent data showing that the same hippocampal cells which have place fields in a radial-arm maze can also show conditioned responses in classical conditioning tasks. The present article focuses on clarifying how the emotional conditioning and adaptive timing processes are designed and related to each other. In particular, as indicated in Figure 5, reinforcement learning and adaptive timing are suggested to take place in different neural circuits, but circuits that interact with and modulate each other during normal behaviors. As reviewed below, aspects of emotional
conditioning may be spared even if adaptive timing is deranged, just as aspects of motor performance may be spared even if adaptive timing is deranged.

11. Adaptively Timed Conditioning of Hippocampal Pyramidal Cells

A large number of experiments have by now documented a role for hippocampal cells in adaptively timed conditioning. As Berger, Berry, and Thompson (1986, p. 204) have noted, “One of the striking features of these conditioning-induced changes in hippocampal activity is that a close parallel develops between the pattern of CS-evoked hippocampal pyramidal cell activity and the shape of the conditioned response – both during NM (nictitating membrane) aversive (Berger, Laham, and Thompson, 1980) and CJM (jaw movement) appetitive (Berry and Oliver, 1982) paradigms”. In addition, when animals were conditioned using different CS – US ISI intervals, the poststimulus histograms of pyramidal cell firing paralleled the shape of the NM response at ISIs of 150 msec and 250 msec. A 50 msec ISI did not lead to a conditioned NM response, and no enhancement of hippocampal unit activity occurred in either the CS – US interval or the US interval (Hoehler and Thompson, 1980). In addition, in a signal detection task in which a white noise CS was varied from suprathreshold to threshold intensity, hippocampal firing to the CS completely predicted the occurrence of a behavioral response (Kettner and Thompson, 1982).

Such data led Berger, Thompson, and their colleagues to characterize the response pattern of hippocampal pyramidal cells as a “temporal model” of the conditioned response, a proposal that was also espoused by Solomon (1979, 1980). The START model suggests how this “temporal model” develops and how it is integrated into a larger neural system for reinforcement learning, recognition learning, and motor learning.

12. Comparison of Conditioned Properties of Hippocampal Pyramids, NMDA Receptors at Dentate Cells, and Hippocampal Afferents

Berger, Berry, and Thompson (1986) reported data from dentate granule cells showing “increased firing rate beginning in the CS period and continuing through the US period... For any given cell, the latency of increased firing was constant and was time-locked to the CS” (p. 213). This difference between the “time-locked” responses of dentate granule cells and the adaptively timed responses of hippocampal pyramidal cells suggests that pyramidal cells and dentate cells process hippocampal afferents in different ways. Berger et al. (1986) also reviewed data indicating that the high correlation between firing of hippocampal pyramidal cells and conditioned responses cannot be explained solely by conditioned changes in afferents to the hippocampus. In particular, Berger and his colleagues (Berger et al. 1986; Berger and Weiss, 1987) reviewed data about firing patterns in two major hippocampal afferents, the medial septum and the entorhinal cortex, during conditioning of the nictitating membrane response. In the medial septum, each new stimulus generates a short, transient burst of activity followed by rapid habituation to the baseline response pattern. The firing rate pattern in the entorhinal cortex is more like that seen in areas CA3 and CA1 of the hippocampus. However, the hippocampal behavioral trace is much stronger than the entorhinal trace and evolves more slowly. Whereas the entorhinal trace takes only 10-20 trials to form and to reach its asymptotic level, the hippocampal trace starts forming when responses start being generated and continues to grow stronger through the first 100-150 trials (Berger et al., 1986). These data are consistent with the hypothesis that at least part of the hippocampal trace is endogenously generated. It needs also to be noted that, although two of the most important projections to the hippocampus arise from the medial septum and the entorhinal cortex, other areas do send projections there too, among them the supramamillary region and the dorsal diagonal gyrus, which project to region CA3 of the hippocampus via the fornix, and the anterior and medial dorsal thalamic nuclei, which project to region CA1 of the hippocampus via the cingulum. Collingridge and Davies (1989, p. 130) discussed additional evidence that “an increase in transmitter release maintains (hippocampal) LTP. This
evidence is strongest for the perforant path input to the dentate gyrus, where the increase in transmitter release is dependent on the activation of NMDA receptors”.

13. Conditioning at Dendritic Spines of Dentate Granule Cells

The formal processing stages of the START model have a natural hippocampal interpretation that mimics the observed differences between dentate granule cells and hippocampal pyramidal cells during conditioning, and uses a learning mechanism at the model analog of granule cells that is interpreted below in terms of a learned control of transmitter release, with associated alterations in protein synthesis.

The combinatorial explosion of cells that was described in Section 9 is avoided by assuming that the spectral activations $x_{ij}$ are local potentials at the dendritic spines of hippocampal dentate granule cells. Thus the $x_{ij}$ do not correspond to separate cells, but rather to dendritic spines of a single cell that accumulates signals from many sensory representations. The many pathways from different sensory representations to the dendrites still need to exist, but their targets are a much smaller population of cells and their dendrites. The microscopic biophysical details of this interpretation will be developed elsewhere. Here we show how the formal linkage of spectral learning properties to hippocampal circuitry leads to new explanations and predictions about hippocampal anatomy and neurophysiology.

In this interpretation, there exists a subset of dentate granule cells that reacts at a single spectral averaging rate $r_j$ in (7), and different subsets of granule cells react at different rates $r_j$. Each such cell possesses a large number of dendrites that are densely encrusted with dendritic spines. Each spine is assumed to structurally realize a private channel at which individualized activations $x_{ij}$ can be processed at the rate $r_j$. Learning is activated by a Now Print signal $N$ that globally activates the entire cell. Figure 14b-e indicates that the twice-gated operation $f(x_{ij})y_{ij}z_{ij}$ in (11) may be realized in several different ways.

Figure 14

14. Convergence of Dentate Granule Cells at CA3 Pyramidal Cells

This interpretation of the START model suggests that (1) conditioning occurs at dentate granule cells, (2) the latency of conditioned firing is constant at individual granule cells, and (3) the hippocampal pyramidal cells to which dentate cells project form a “temporal model” of adaptively timed behavioral responses. These data are consistent with the model hypothesis, formalized in equation (11), that the individual terms $f(x_{ij})y_{ij}z_{ij}$, which correspond to each fixed and different rate $r_j$, summate to generate an adaptively timed model $R$ of the behavioral response. We interpret the cells corresponding to different values of $r_j$ as different (subsets of) dentate granule cells, and the cells corresponding to the output $R$ as CA3 pyramidal cells (Figure 15). It is also assumed that different subsets of CA3 pyramidal cells correspond to different drive representations (Olds, 1977).

Figure 15

This interpretation of (11) suggests that many dentate granule cells converge on individual CA3 pyramidal cells. This property is consistent with the fact that, in the rat, there are approximately 1,000,000 dentate granule cells but only 160,000 CA3 pyramidal cells (Boss, Peterson, and Cowan, 1983; Boss, Turlejski, Stanfield, and Cowan, 1987; Squire, Shimamura, and Amaral, 1989). In addition, a CA3 cell receives approximately 80 mossy-fiber inputs from dentate granule cells (Squire, Shimamura, and Amaral, 1989). It may thus not be a coincidence that the Grossberg and Schmajuk (1989) computer simulations and our own found that 80 values of $r_j$ provide an excellent fit to behavioral data on the conditioned
NM response. On the other hand, in unreported simulations we have shown that qualitative model properties are robust when the number of populations is increased or decreased by a factor of four. In any case, the order of magnitude between anatomical and model convergence is acceptable. This anatomical interpretation can be unequivocally tested in terms of the following

**Prediction:** Sets of dentate cells, or perhaps a combination of dentate cells and mossy fibers, exhibit dynamics capable of giving rise to a full spectrum of activation rates $r_j$.

Gray (1982, pp. 97-100) has surveyed experiments that are consistent with this neurophysiological interpretation of START learning. These data show habituation in dentate granule cells to stimulation of the perforant path, analogous to $y_{ij}$ habituation in response to inputs $S_i$ in equations (8) and (9); potentiation at the dentate synapse in response to perforant path stimulation, analogous to $z_{ij}$ conditioning in response to inputs $S_i$ in equations (8) and (10); swelling in the dendritic spines of dentate granule cells after tetanizing stimulation of the perforant pathway, analogous to the anatomical interpretation of the $f(x_{ij})y_{ij}z_{ij}$ process at dentate spines; and facilitation of dentate response to a perforant path pulse by a prior conditioning pulse to the septum, analogous to the action of the Now Print signal on conditioning in equation (10).

The START model is also consistent with more recent data concerning the effects of manipulations of the dentate gyrus upon the behavior of animals. Diaz-Granados et al. (1991) showed that selective dysgenesis of the dentate gyrus in rats due to neonatal X-irradiation impaired performances in a differential-reinforcement-of-low-rate-of-responding (DRL) task: After X-irradiation, which selectively prevents the formation of granule cells, animals were unable to run slowly down a hallway to receive reward. In a similar paradigm, Robinson (1991) impaired acquisition of the conditioned nictitating membrane response, and Thompson and Disterhoft (1991) showed that NMDA agonists and antagonists have opposite effects upon long-interval trace eye blink conditioning. These results are consistent with the proposed interpretation of the START model: Animals without functional dentate granule cells, or with impaired NMDA receptors, should be unable to adaptively time their conditioned responses.

This interpretation is also consistent with the lack of effect of granule cell dysgenesis (Armstrong et al., 1991) or NMDA antagonist treatment (Sutherland et al., 1991) upon place learning. The START model suggests that the reinforcing value of an event is less affected by these manipulations than is the ability to adaptively time reinforced behavior. Given this interpretation, the close temporal relationship between being in a place and getting rewarded there may be spared, but distant temporal relationships may be unbridgeable. This possibility may be further testable in the context of fear conditioning, where NMDA antagonists impair fear conditioning over an ISI of 4 seconds (Davis et al., 1991). If this failure is substantially due to a failure of adaptive timing, then near-normal fear conditioning may be found over sufficiently short interstimulus intervals if NMDA antagonists are selectively applied to dentate granule cells.

The occurrence of associative learning on dendritic spines also helps to explain how the read-out (or performance) of old associative memories can be decoupled from the read-in (or learning) of new associative memories. Such a dissociation is needed to solve the self-printing problem (Section 5). By (11) read-out of old associative memories is accomplished by the twice-gated signals $f(x_{ij})y_{ij}z_{ij}$; also see Figure 15. These signals need to be separated from the influence of twice-gated signals activated by other conditioned stimuli and other spectral averaging rates. Dendritic spines can provide this functional separation during read-out, while also being responsive to more global events, such as the Now Print signal $N$, during LTM read-in via equation (10).

The hypothesis of the START model that hippocampal LTP occurs at dendritic spines in order to functionally dissociate the read-out of old associative memories from the read-in
of new associative memories was discussed in Grossberg (1975, see Figure 25). This type of process has recently excited a great deal of further work based upon new experimental approaches to hippocampal LTP and the discovery of the NMDA receptor (Brown et al., 1988; Brown and Zador, 1990; Eccles, 1989). The START model is consistent with data showing that conditioning takes place at NMDA receptors in the perforant-to-dentate pathway. As Collingridge and Davies (1989, p. 130) have noted: “Most neurochemical evidence suggests that an increase in transmitter release maintains LTP. This evidence is strongest for the perforant path input to the dentate gyrus, where the increase in transmitter release is dependent on the activation of NMDA receptors.”

15. NMDA Receptors and Adaptive Timing

The recent experiments suggesting that an increase in presynaptic transmitter release may help to control LTP at dentate granule cells and includes activation of NMDA receptors (Collingridge and Davies, 1989; Errington, Lynch, and Bliss, 1987) are consistent with another early prediction (Grossberg, 1968, 1969a) about associative learning. This prediction suggested that associative learning is achieved by “joint control of presynaptic excitatory transmitter production by presynaptic and postsynaptic levels of membrane potential. This control is presumed to be effected by the interaction of the pairs (Na+, K+) and (Ca++, Mg++) of antagonistic ions whose binding properties to intracellular sites and enzymes set cellular production levels” (Grossberg, 1969a, p. 325). In particular, a synergistic interaction of a voltage-dependent, postsynaptically generated, inward Ca++ current with inward Na+ and outward K+ currents was predicted, as well as a competitive interaction between Ca++ and Mg++. Recent studies of LTP at NMDA receptors have reported and greatly elaborated contemporary understanding of this sort of interaction, including the competition between Ca++ and Mg++ (Collingridge and Davies, 1989).

Related predictions may now be testable at the perforant path – CA3 pyramidal cell synapse. One prediction suggests that certain “nerve cells are capable of learning as ‘chemical dipoles’ ” (Grossberg, 1969a, p. 325) control the availability of the proper relative amounts of Ca++, Mg++, Na+, and K+, among other chemicals, at the cell sites where they are needed. Such control is suggested to coordinate potentiation of presynaptic transmitter production with levels of postsynaptic protein synthesis aimed at enabling the postsynaptic cell to cope with time-varying loads of presynaptic input. Akers et al. (1986) have shown that “protein kinase activation leading to phosphorylation of neural proteins appears to occupy a pivotal role in the development and expression of synaptic plasticity” in response to perforant path stimulation (p. 587). Further experiments are needed to test possible correlations between presynaptic and postsynaptic effects. This proposal also suggested that the shape of neurons realizes a type of structural dipole that helps to support the dynamics of the chemical dipole. The two poles of the structural dipole, at the dendritic apparatus/cell body and the synaptic knobs, respectively, were suggested to help maintain chemical gradients along the axons between these poles. Interactions between the cell nucleus, the cell membrane, microtubules, and tight junctions between presynaptic terminals and postsynaptic cells were proposed to maintain these gradients. Further details concerning these predicted chemical dipole properties, may be found in Grossberg (1969a).

Sections 16–22 summarize computer simulations that demonstrate other model properties that are consistent with conditioning data.

16. Stability of Learning Over Many Trials

Some learning models become unstable when they experience a large number of learning trials. Figure 16 shows the output of the model after 4, 50 and 100 learning trials, illustrating that this output pattern persists over many trials, even after asymptote is reached.

Figure 16
17. Robustness of the Model

Model properties are robust under physically plausible perturbations of its structure or parameters. For example, Figure 17a shows that the asymptotic behavior of the model is qualitatively preserved under large changes in the learning rate $\alpha_e$ in (10). Figure 17b shows that the model’s behavior is unaffected by changes in the parameter which controls the speed at which the competition among sensory representations $S$ take place. Figure 17c shows that the circuit’s qualitative behavior is robust against large accelerations or decelerations of the rate at which $\mathcal{D}$ generates the Now Print signal.

![Figure 17](image)

The adaptive timing circuit learns accurately even when the behavior of some other part of the model is qualitatively altered. In Figures 8 through 17, the parameters controlling the STM representations $S$ were chosen so that STM can store more than one item. In Figure 18, the parameters were chosen so that only one sensory representation can remain active through time. This has a dramatic effect upon the singly-gated signals $f(x_{ij})y_{ij}$ within the model, since their support from $S_1$ does not persist when $S_0$ is large on training trials, but it has little effect upon the timing of the circuit, which again reaches maximal total activity $R$ at around the time the $US$ is expected.

![Figure 18](image)

18. Inverted U and Weber Law

Figure 1 shows the average nictitating membrane topographies of animals trained with CS's of 50 msec duration and ISI's of 125, 250, 500, or 1000 ms, depending upon the group to which the animal belonged. Figure 19 displays the outputs of the model at the corresponding ISI's. The model mimics the data pattern of rapid growth of the peak amplitude as the ISI increases through small values, followed by a gradual fall-off of peak amplitude for larger values of the ISI; and an increase in output width across time as the ISI increases.

![Figure 19](image)

19. Comparison with Data using Multiple ISI's

Figure 20 summarizes data from experiments, reported in Millenson et al. (1977), in which rabbits were conditioned in a nictitating membrane response paradigm. The ISI was one of two different values: 200ms or 700ms, with the different ISI's being presented at differing frequencies to different groups. In the group $P_{\frac{1}{2}}$ which received equal numbers of each ISI, the animals' average NM extension on test trials shows a double peak for the longer trials. The two peaks also exhibit the Weber law property. These double peak experiments provide strong evidence that a spectrum of possible times exists that is tuned by experience. Figure 21 summarizes a computer simulation of that condition, which also exhibits two peaks that obey the Weber law at the two times at which the $US$ would have been delivered.

![Figure 20](image)

![Figure 21](image)

20. ISI Shift Experiments

In Coleman and Gormezano (1971), animals were conditioned in a paradigm whose temporal characteristics were shifted either gradually or abruptly, from a 200 msec ISI to a 700 msec ISI, or conversely, during the course of the experiment. The animals' behaviors
across learning trials are summarized in Figure 22. Figure 23 summarizes a set of computer simulations that qualitatively mimic the conditions of the original experiments.

Figure 22
Figure 23

21. Partial Reinforcement Experiments

The classical conditioning circuit depicted in Figure 5 forms part of a larger model neural system that is capable of explaining many data about operant conditioning (see Grossberg, 1982a, 1982b, 1984, 1987a and Part II for further discussion). Correspondingly, many operant conditioning data share similar properties with classical conditioning data. For example, the experiments of Roberts (1981) used an operant rat lever-pressing task in which frequency of reinforcement was varied but the ISI was fixed. This manipulation altered the terminal level of responding to the stimulus, without changing the peak time of responding; that is, partial reinforcement affects the likelihood, but not the timing, of the response. The results are shown in Figure 24.

Figure 24

A computer simulation of the same paradigm is shown in Figure 25. As in the Roberts (1981) data, only the level, not the time at which the output peaked, was affected by the probability of reinforcement.

Figure 25

22. Time Averaging in Response to Multiple Stimuli

Holder and Roberts (1985) examined the effect of combining the timed responses to two different CS stimuli, a tone and a light, using rats in a lever-pressing task. If each of these stimuli has acquired the ability to elicit a conditioned response, and if they are presented sequentially, the resulting response is timed neither as the former nor as the latter stimulus would have required, but rather as an average. A simulation that qualitatively replicates this averaging property is shown in Figure 26. This figure was generated with the model parameters set so that more than one sensory representation could be active in STM at one time. When the two stimuli were presented, the resulting output produced a peak that averages between the two expected times of arrival.

Figure 26

PART II

REINFORCEMENT, RECOGNITION, AND MOTOR LEARNING

The spectral timing part of the circuit in Figure 5 is new. The remainder of the circuit is part of a larger theory concerning the neural substrates of reinforcement, recognition, attention, memory search, and motor control. Relevant parts of the theory are summarized below. They are used to clarify how the adaptive timing circuit interacts with other types of brain circuits, and to show how recent neurobiological data support the existence of each of the model’s processing stages.

23. Reinforcement Learning in Vertebrates and Invertebrates

19
In Section 4, a drive representation $\mathcal{D}$ was defined as a population of cells at which sensory, reinforcement, and homeostatic, or drive, signals converge to regulate reinforcement learning, emotional reactions, and motivational decisions. Figure 27 depicts the type of model circuit in which drive representations were described in Grossberg (1971) to explain vertebrate conditioning data. A similar model (Figure 28) has recently been used to explain invertebrate conditioning data from experiments on Aplysia (Buonomano, Baxter, and Byrne, 1990; Grossberg, 1984b). The use of a similar circuit by such different species is clarified by the fact that it is the simplest solution of two general learning problems, called the synchronization problem and the persistence problem, that all animals capable of classical conditioning need to solve (Grossberg, 1971, 1982b, 1987a).

Figure 27

Figure 27 contains pathways that were omitted from Figure 5 for simplicity. As noted in Section 6, during classical conditioning, pairing of a $CS_1$ sensory representation $S_{CS_1}$ with activation of a drive representation $\mathcal{D}$ by a reinforcer $US$ causes the modifiable synapses connecting $S_{CS_1}$ with $\mathcal{D}$ to become strengthened. This conditioning event converts $CS_1$ into a conditioned reinforcer. Figure 27 shows reciprocal conditionable pathways from the drive representations $\mathcal{D}$ to the sensory representations $S$. Conditioning of these pathways is called incentive motivational learning. Activation of conditioned $S \rightarrow \mathcal{D} \rightarrow S$ feedback pathways by $CS_1$ can shift attention towards the set of all previously reinforced sensory cues that are motivationally consistent with $\mathcal{D}$.

Figure 28

This shift of attention occurs because the sensory representations, in addition to emitting conditioned reinforcer signals and receiving incentive motivation signals, compete among themselves (Figures 27) for a limited capacity short-term memory (STM). When strong incentive motivational feedback signals are received at the sensory representational field, these signals can bias the competition for STM activity towards the set of motivationally preferred cues.


The feedback signals $S \rightarrow \mathcal{D} \rightarrow S$ generate a resonant state of activation between levels $S$ and $\mathcal{D}$ that focuses attention upon recognition codes of events which have led to reinforcing consequences during past experiences. Such attentional modulation enables a biological information processing system to selectively process those environmental inputs that are most important to its current goals. A typical example of such selective processing is illustrated by the blocking paradigm shown in Figure 29 (Kamin, 1969). First, a conditioned stimulus $CS_1$, such as a tone, is presented several times, followed at a given time interval by an unconditioned stimulus $US$, such as electric shock, until a conditioned response $CR$, such as fear, develops. Then $CS_1$ and another conditioned stimulus, $CS_2$, such as light, are presented simultaneously, followed at the same time interval by the $US$. After conditioning, $CS_2$ is presented alone, not followed by a $US$, and no conditioned response occurs. Intuitively, $CS_1$ "blocks" conditioning of the simultaneously presented $CS_2$ because $CS_1$, by itself, perfectly predicts its consequence, the $US$. The $CS_2$ is thus redundant and unpredictable, hence does not get conditioned to the $US$.

Figure 29

The blocking property may be explained in terms of four properties of the network in Figure 26: (1) Pairing of a $CS_1$ with a $US$ in the first phase of the blocking experiment endows the $CS_1$ cue with properties of a conditioned, or secondary, reinforcer; that is, the positive feedback pathway $S_1 \rightarrow \mathcal{D} \rightarrow S_1$ between the drive representation $\mathcal{D}$ and the sensory
representation $s_1$ of $CS_1$ is strengthened due to learning. (2) These reinforcing properties of a $CS_1$ shift the focus of attention towards its own processing at $S_1$. (3) The processing capacity of attentional resources is limited, as a result of the competition between sensory representations $S$. Thus a shift of attention towards one set of stimuli can prevent other stimuli, such as $CS_2$, from being attended (Figure 28). Withdrawal of attention from the sensory representation $S_2$ of the stimulus $CS_2$ prevents that representation from entering new conditioned relationships, by attenuating learning from $S$ to $D$ and from $D$ to $S$. Learning is attenuated when the activity $S_2$ of $S_2$ becomes small, because it is regulated by an activity-dependent gate, as in (4).

Just as simultaneous occurrence of a conditioned reinforcer $CS_1$ with a new $CS_2$ can block conditioning of $CS_2$, so too can simultaneous occurrence of a primary reinforcer $US$ with a new $CS$ block conditioning of $CS$. This latter property helps to explain why $US$ onset needs to occur after $CS$ onset in order for effective conditioning to occur (Buonomano, Baxter, and Byrne, 1990; Grossberg, 1982b; Grossberg and Levine, 1987).

One way to verify whether a neural model has processing stages that correlate well with brain circuits is to test if a formal model lesion has effects similar to those of a corresponding brain lesion on behavioral properties. Grossberg (1975, Figure 24) suggested that a final common path within (an expanded model of) a drive representation $D$ includes the hippocampal formation. Eliminating the “hippocampal formation” in the model would therefore weaken $D \rightarrow S$ feedback signals, and thus the model’s mechanism of blocking. Hippocampal lesions do, in fact, prevent blocking from occurring. Both $CS_1$ and $CS_2$ can be conditioned in a blocking experiment performed on a hippocampectomized animal (Le Doux, 1989; Rickert, Bennett, Lane, and French, 1978; Schmajuk, Spear, and Isaacson, 1983; Solomon, 1977). Likewise, hippocampectomized animals find it hard to actively ignore nonreinforced cues (Pribram, 1986).

These experiments also showed that hippocampal lesions do not interfere with emotional conditioning. Although such a dissociation could not be explained in the model of Figure 27 and 28, it can be explained using the model of Figure 4, which distinguishes the circuit for adaptive timing from the circuit for emotional conditioning. A circuit which combines the components of Figure 4 with those of Figure 26 is shown in Figure 30.

Figure 30

25. Subcortical Fear Conditioning and a Cortical Role in Extinction

The circuit in Figure 30 includes sensory representations $S$ that process incoming signals in two successive processing stages $S^{(1)}$ and $S^{(2)}$. In Grossberg (1971, 1975, 1978, 1982b) it was shown that each sensory representation $S$ need to be broken into two successive stages $S^{(1)}$ and $S^{(2)}$, as in Figure 30, such that $S^{(1)}$ projects to both $S^{(2)}$ and $D$, $D$ projects to $S^{(2)}$, and $S^{(2)}$ projects back to $S^{(1)}$. The pathway $S^{(1)} \rightarrow D$ can support emotional conditioning, thereby converting the stimuli represented at $S^{(1)}$ into conditioned reinforcers. The pathway $D \rightarrow S^{(2)}$ supports incentive motivational conditioning. It primes all sensory representations that are motivationally consistent with $D$. The multisynaptic pathway $D \rightarrow S^{(2)} \rightarrow S^{(1)}$ provides the feedback from $D$ that supports the blocking process. This expanded version of the model has been used to computationally simulate blocking data (Grossberg and Levine, 1987).

Grossberg (1975, 1978) interpreted the first stage of sensory processing $S^{(1)}$ as a thalamic representation, the second stage $S^{(2)}$ as a cortical representation, the first stages of drive representational processing $D$ as networks, such as hypothalamus and amygdala, that are involved in homeostatic and emotional processing, and the final stages of drive representational processing as including the hippocampal formation. With this interpretation, the
conditioning of $S^{(1)} \rightarrow D$ synapses in Figure 30 predicts that subcortical emotional conditioning is possible. This prediction has been supported by recent experiments which show that ablation of visual cortex does not interfere with acquisition of fear responses to visual stimuli, but does greatly prolong, if not prevent, extinction of fear responses (LeDoux, Romanski, Xagorarlis, 1989). Analogous results have been reported for auditory conditioning. Fear conditioning to acoustic stimuli is mediated by projections from auditory processing areas of the thalamus to the amygdala, which bypass the auditory cortex and provide a subcortical mechanism of emotional learning (LeDoux, Iwata, Cicchetti, and Reis, 1988).

An ablation of cortex that spares the thalamus would, in the model, remove $S^{(2)}$ but leave $S^{(1)}$ intact. The acquisition of fear responses is mediated by the intact conditioned reinforcer pathways $S^{(1)} \rightarrow D$. Extinction in the model is mediated by cognitive mechanisms, described in Sections 27 and 28, that act upon the ablated cortical $S^{(2)}$ representations. These $S^{(2)}$ representations alter the dynamics of the $S^{(1)}$ and $D$ representations via $S^{(2)} \rightarrow S^{(1)}$ and $S^{(1)} \rightarrow D$ interactions in a way that causes extinction (Grossberg, 1982b, 1984a). Thus a selective ablation of $S^{(2)}$ spares acquisition of fear conditioning but impairs extinction, as also occurs in the data. This explanation of normal acquisition with abnormal extinction is consistent with the fact that acquisition may itself be impaired in paradigms wherein the abnormal distractibility of hippocampectomized animals occurs (see Section 28). Such data provide support for the model's anatomical interpretation of the interactions $S^{(1)} \rightarrow D$ and $S^{(2)} \rightarrow S^{(1)}$. The prediction in Grossberg (1975) that the $D \rightarrow S$ pathway is a hippocampal to cortical pathway is consistent with experiments of Rosene and Van Hoesen (1977). Thus the main anatomical and functional properties of the model have neurobiological correlates that permit its operations to be subjected to additional neurobiological tests, and to be refined in the light of new neural data.

Further support for the model derives from its ability to provide a unified mechanistic explanation of a large data base about conditioned behavior, including data about unblocking, latent inhibition, overexpectation, behavioral contrast, vicious circle behavior, selective forgetting, hyperphagia, and analgesia (Grossberg, 1972a, 1972b, 1975, 1982b, 1984a). A recent summary of other predictions and their experimental support is found in Carpenter and Grossberg (1991a, 1991b) and Grossberg (1991).

One issue of particular interest concerns whether or not there exists an analog of the vertebrate thalamocortical pathway $S_1 \rightarrow S_2$ in certain invertebrate circuits. Such a pathway is not described in the invertebrate model of Figure 27. Another important issue concerns the anatomical sites that subserve the motor learning, or "habit strength", pathways in Figures 27 and 30. Various data suggest that these pathways include the cerebellum.

26. Conditioning in the Cerebellum

Thompson et al. (1984, 1987) reviewed many experiments which implicate the cerebellum as "an obligatory part of the learned response circuit for eyelid/NM conditioning" (Thompson et al., 1987, p. 353). In particular, they noted that ablation of deep cerebellar nuclei abolishes the previously learned response, prevents relearning of the response as a result of ipsilateral stimulation (although such stimulation continues to evoke unconditioned responses), but has no effect upon acquisition of the response as a consequence of contralateral stimulation. They also summarized data from their laboratory showing that "the dorsal accessoary olivary-climbing fiber projection is the necessary and sufficient US pathway ... the mossy fiber projection is the necessary and sufficient CS pathway ... and ... appropriately timed conjoint activation of mossy fibers as the CS and climbing fibers as the US yields normal learning of discrete, adaptive behavioral CSs" (Thompson et al., 1987, pp. 387-388). Clark et al. (1984) have shown, moreover, that lesions of the deep cerebellar nuclei not only abolish the conditioned behavioral response, but also abolish the hippocampal temporal model of the behavior (Part
I) in response to ipsilateral USs. Administration of contralateral USs quickly causes the reacquisition of both the neuronal model and the behavioral response. Taken together, these data strongly support the hypothesis that conditioning of the CS-activated discrete adaptive response occurs in the cerebellum and that these cerebellar signals are needed for expression of the hippocampal temporal model.

The hypothesis that the cerebellum helps to control motor learning has a long history. Brindley (1964) and Grossberg (1964) were among the first to model motor learning in the cerebellum at the synapses between cerebellar parallel fibers and Purkinje cell dendritic spines, using the climbing fibers as a teaching signal. Grossberg (1969c), Marr (1969), and Albus (1971) further modeled this concept. Marr (1969) suggested that these synapses increase in strength due to learning; Albus (1971) suggested that they decrease in strength; Grossberg (1969c) suggested that they may either increase or decrease in strength, depending upon the learning context. Subsequent models of cerebellar motor learning include those of Bullock and Grossberg (1991), Fujita (1982a, 1982b), Grossberg and Kuperstein (1986, 1989), Houk, Singh, Fisher, and Barto (1989), and Ito (1974, 1984). These cerebellar models have been used to analyze behavioral and neural data about eye and arm movements, such as the results of Ebner and Bloedel (1981), Gilbert and Thach (1977), Ito (1982), Optican and Robinson (1980), and Ron and Robinson (1973). In addition to their discussions of nictitating membrane and jaw movement conditioning, Thompson et al. (1984, 1987) also summarized experiments demonstrating motor learning in the cerebellum during classical conditioning of the limb flexion reflex. Thus the cerebellum plays a key role in conditioning motor responses of eye, arm, leg, nictitating membrane, and jaw movements, among others.

These recent data and models about cerebellar learning clarify how motor responses are adaptively controlled, and also suggest that motor learning differs from the types of conditioned reinforcer learning, incentive motivational learning, recognition learning, and adaptive timing that are depicted in Figures 4, 27, and 30.

27. Macrocircuit for Sensory-Cognitive Processing: Adaptive Resonance Theory

It remains to describe how the orienting system A in Figure 4 is controlled; in particular, how the unexpected nonoccurrence of a reinforcer can activate A and thereby cause orienting reactions, attention shifts, and emotional frustration. With this information in hand, the hypothesis that drive representations inhibit A in response to expected nonoccurrences can be better understood as a mechanism for preventing maladaptive reactions to predictive cues. It also remains to discuss how extinction is controlled via this process, as remarked in Section 25.

These types of reactions are modeled by sensory-cognitive circuits that are called Adaptive Resonance Theory, or ART models. ART models have been used to explain and predict a large body of cognitive and neural data about recognition learning, attention, and memory search (Carpenter and Grossberg, 1991a; Commons, Grossberg, and Staddon, 1991; Grossberg, 1982a, 1987a, 1987b, 1988). ART systems suggest a solution to a fundamental learning problem that is called the stability-plasticity dilemma: An adequate self-organizing recognition system must be capable of plasticity in order to learn about significant new events, yet it must also remain stable in response to irrelevant or often repeated events. In order to prevent the relentless degradation of its learned codes by the “blooming, buzzing confusion” of irrelevant experience, an ART system is sensitive to novelty. It is capable of distinguishing between familiar and unfamiliar events, as well as between expected and unexpected events.

The importance of expectancy-related processes in conditioning and cognitive processes has been extensively documented since the pioneering work of Tolman (1932) and Sokolov (1958, 1968). In ART, interactions between an attentional subsystem and an orienting subsystem, or novelty detector, enable the network to self-stabilize its learning, without an external teacher, as the learned recognition code becomes globally self-consistent and predictively accurate; in other words, as the system familiarizes itself with an environment
by categorizing the information within it in a way that leads to behavioral success. The
tentional subsystem undergoes both bottom-up learning and top-down learning within the
LTM-marked pathways between the processing levels denoted by $\mathcal{F}_1$ and $\mathcal{F}_2$ in Figure 31.
The top-down LTM process learns expectations. The network self-stabilizes its learning by
matching its top-down expectations against bottom-up input patterns and using the degree
of match or mismatch to regulate processes of learning or memory search, respectively. Thus
ART suggests how novelty-sensitive matching processes regulate the course of learning.

Figure 31

By itself, the attentional subsystem is unable simultaneously to maintain stable repre-
sentations of familiar categories and to learn new categories for unfamiliar patterns. An
isolated attentional subsystem may be either rigid and incapable of creating new recogni-
tion categories for unfamiliar patterns, or unstable and capable of ceaselessly recoding the
recognition categories of familiar patterns.

The orienting subsystem interacts with the attentional subsystem to maintain the stabil-
ity-plasticity balance. It resets the STM of the attentional subsystem when an unfamiliar
event occurs and thereby initiates a memory search within the attentional subsystem for
a better internal representation. The orienting subsystem is thus essential for expressing
whether an input pattern is familiar and well represented by an existing recognition code,
or unfamiliar and in need of a new recognition code.

Figure 32 illustrates a typical ART memory search cycle. As shown in Figure 32a, an
input vector $\mathbf{I}$ registers itself as a pattern $\mathbf{X}$ of activity, or short term memory (STM), across
level $\mathcal{F}_1$. The $\mathcal{F}_1$ output vector $\mathbf{S}$ is then transmitted through the multiple converging and
diverging adaptive pathways emanating from $\mathcal{F}_1$. This transmission event multiplies the
vector $\mathbf{S}$ by a matrix of adaptive weights, or long term memory (LTM) traces, to generate
a net input vector $\mathbf{T}$ to level $\mathcal{F}_2$. Lateral inhibitory interactions within $\mathcal{F}_2$ contrast-enhance
vector $\mathbf{T}$. A compressed activity vector $\mathbf{Y}$ is thereby generated across $\mathcal{F}_2$.

Figure 32

Activation of $\mathcal{F}_2$ nodes may be interpreted as "making a hypothesis" about an input
$\mathbf{I}$. When $\mathbf{Y}$ is activated, it generates a signal vector $\mathbf{U}$ that is transmitted along top-down
adaptive pathways. After multiplication of these top-down signals by a matrix of adaptive
weights, or LTM traces, a net vector $\mathbf{V}$ inputs to $\mathcal{F}_1$ (Figure 32b). Vector $\mathbf{V}$ plays the role
of a learned top-down expectation. Activation of $\mathbf{V}$ by $\mathbf{Y}$ may be interpreted as "testing
the hypothesis" $\mathbf{Y}$, or "reading out the category prototype" $\mathbf{V}$. ART networks are designed
to match the "expected prototype" $\mathbf{V}$ of the category against the active input pattern, or
exemplar, $\mathbf{I}$.

This matching process may change the $\mathcal{F}_1$ activity pattern $\mathbf{X}$ by suppressing activation
of all the feature detectors in $\mathbf{I}$ that are not confirmed by $\mathbf{V}$. The resultant pattern $\mathbf{X}^*$
encodes the pattern of features to which the network "pays attention". If the expectation
$\mathbf{V}$ is close enough to the input $\mathbf{I}$, then a state of resonance occurs as the attentional focus
takes hold. Damasio (1989) uses the term "convergence zones" to describe the process
whereby an activation pattern $\mathbf{X}^*$ across distributed features is bound together by resonant
feedback. The resonant state persists long enough for learning to occur; hence the term
adaptive resonance theory. ART systems learn prototypes, rather than exemplars, because
the attended feature vector $\mathbf{X}^*$, rather than the input $\mathbf{I}$ itself, is learned.

The criterion of an acceptable match is defined by a dimensionless parameter called vigil-
ance. The vigilance parameter is computed in the orienting subsystem $\mathcal{A}$, where it may
be increased by punishing events or other unexpected consequences (Carpenter and Gross-
berg, 1987a; Carpenter, Grossberg, and Reynolds, 1991; Carpenter, Grossberg, Markuzon,
Reynolds, and Rosen, 1991). Vigilance weighs how close the input exemplar $I$ must be to the top-down prototype $V$ in order for resonance to occur. Because vigilance can vary across learning trials, recognition categories capable of encoding widely differing degrees of generalization, or morphological variability, can be learned by a single ART system. Low vigilance leads to broad generalization and abstract prototypes. High vigilance leads to narrow generalization and to prototypes that represent fewer input exemplars. In the limit of very high vigilance, prototype learning reduces to exemplar learning. Thus a single ART system may be used, say, to recognize abstract categories of faces and dogs, as well as individual faces and dogs.

If the top-down expectation $V$ and the bottom-up input $I$ are too novel, or unexpected, to satisfy the vigilance criterion, then a bout of hypothesis testing, or memory search, is triggered. Memory search leads to selection of a better recognition code at level $F_2$ with which to represent input $I$ at level $F_2$. The orienting subsystem $A$ mediates the search process. During search, the orienting subsystem interacts with the attentional subsystem, as in Figures 32c and 32d, to enable the attentional subsystem to learn about novel inputs without risking unselective forgetting of its previous knowledge.

The search process prevents associations from forming between $Y$ and $X^*$ if $X^*$ is too different from $I$ to satisfy the vigilance criterion. The search process resets $Y$ before such an association can form. A familiar category may be selected by the search if its prototype is similar enough to the input $I$ to satisfy the vigilance criterion. The prototype may then be refined in light of new information carried by $I$. If $I$ is too different from any of the previously learned prototypes, then an uncommitted population of $F_2$ cells is selected and learning of a new category is initiated.

A network parameter controls how deeply the search proceeds before an uncommitted node is chosen. As learning of a particular category self-stabilizes, all inputs coded by that category access it directly, without the need for search. Familiar, consolidated memories can thus be accessed in a one-pass fashion, after resetting the previously active category. The category selected is the one whose prototype provides the globally best match to the input pattern. In a situation where a mixture of familiar and unfamiliar events are experienced, familiar inputs can directly activate their learned categories, while novel inputs continue to trigger adaptive memory searches for better categories, until the network's memory capacity is fully utilized.

These ART mechanisms include the processes that are needed to interpret the effects of $P \rightarrow A$ inhibition that were described in Section 1. These include a process whereby learned expectations may be mismatched by a sensory expectation at level $F_1$ of the attentional subsystem in Figure 31. When a mismatch of bottom-up exemplar and top-down prototype occurs, the orienting subsystem is activated, giving rise to a short term memory (STM) reset wave in the form of a nonspecific arousal burst to the attentional subsystem (Figure 32c). This arousal burst acts to reset the sensory representations of all sensory events that are currently active in STM within the attentional subsystem. Representations with high STM activation tend to become less active, representations with low STM activation tend to become more active, and the novel event which caused the mismatch tends to be more actively stored than it would have been had it been expected. Banquet and Grossberg (1987) have discussed experiments on human event-related potentials (ERPs) during probabilistic choice reaction time tasks that have tested the predicted chronometry of the mismatch-arousal-reset sequence in terms of the P120-N200-P300 sequence of ERPs. One effect of STM reset is to shift the focus of attention towards sensory representations which may better predict environmental contingencies. In a classical conditioning paradigm, such an attention shift can dishabituate, or unblock, sensory representations that were not attended before the STM reset event (Grossberg, 1982a, 1982b). Activation of the orienting subsystem also triggers orienting responses, such as the activation of motor reactions to orient towards the unexpected event (Figure 31).
This organization of learned expectations, attention shift mechanisms, and orienting mechanisms within ART allowed Grossberg and Schmajuk (1989) to hypothesize that activation of the drive representation \( D \) gates, or inhibits, the orienting subsystem \( A \). Activation of this inhibitory gate prevents reset of the attentional focus and release of orienting behaviors if an expected nonoccurrence is experienced. Such a gating operation does not, however, prevent a sensory match from being detected earlier than usual, because matches with learned expectations occur within the attentional subsystem, not the orienting subsystem. At times when the adaptive timing mechanism is inactive, the gate is open. Then activation of the orienting subsystem can trigger reset of STM and orienting reactions in response to unexpected nonoccurrences.

28. Hippocampal Lesions and Medial Temporal Amnesia

The division of labor within ART, between an attentional subsystem and an orienting subsystem, thus provides the type of processing substrate that is needed to instantiate adaptive timing heuristics. This division of labor has also been helpful in clarifying many other types of data. For example, Carpenter and Grossberg (1987a, 1988b) have pointed out that a lesion of the ART orienting subsystem creates a memory disturbance whose formal symptoms are similar to those of humans afflicted with medial temporal amnesia, including unlimited anterograde amnesia; limited retrograde amnesia; failure of consolidation; tendency to learn the first event in a series; abnormal reactions to novelty, including perseverative reactions; normal priming; and normal information processing of familiar events (Cohen, 1984; Graf, Squire, and Mandler, 1984; Lynch, McGaugh, and Weinberger, 1984; Mattis and Kownar, 1984; Squire and Butters, 1984; Squire and Cohen, 1984; Warrington and Weiskrantz, 1970, 1974; Zola-Morgan and Squire, 1990).

Unlimited anterograde amnesia occurs in the model because, without a functional orienting subsystem, the network cannot carry out the memory search and subsequent learning needed to establish a new recognition code. Limited retrograde amnesia occurs because familiar events can directly access their recognition codes, without activating the orienting subsystem. Before events become familiar, a period of memory consolidation occurs during which the orienting subsystem does play a role, as indicated in Figure 32c. This failure of consolidation does not prevent learning per se. Instead, learning is associated with the first recognition category that is activated by bottom-up processing, much as “amnesics are particularly strongly wedded to the first response they learn” (Gray, 1982, p. 253). Abnormal reactions to novelty, including perseverative reactions, occur. In an ART circuit, this happens because the orienting subsystem cannot carry out its normal function of STM reset, and therefore cannot inhibit sensory representations or top-down expectations that may be persistently mismatched by bottom-up sensory signals. The inability to search memory via its orienting subsystem prevents an ART system from discovering more appropriate stimulus combinations to which to attend. In a similar vein, Butters and Cermak (1975), p. 393) reported that “Korsakoff patients’ encoding deficits may be related to a general impairment in their ability to attend to relevant dimensions of stimuli.” Normal priming is possible in an ART model because it can be mediated entirely by the attentional subsystem, notably the top-down expectations of this subsystem. The close correspondence between the symptoms of medial temporal amnesia and the formal properties of an ART model with defective orienting subsystem is consistent with accumulating evidence for the hypothesis (Grossberg, 1975) that the in vivo analog of the ART orienting subsystem intersects, or is closely linked to, the hippocampal formation.

Similar behavioral problems have been identified in hippocampectomized monkeys. Gaffan (1985) noted, for example, that fornix transection “impairs ability to change an established habit... (there is) impaired learning ability when one habit is to be formed in one set of circumstances and a different habit is to be formed in a different set of circumstances that is similar to the first and therefore liable to be confused with it” (p. 94). A similar
problem occurs in an ART network with a defective orienting subsystem. Such a defect prevents STM reset, which normally leads to memory search and learning of different representations for the two similar events. Pribram (1986) calls such a process a "competence for recombinant context-sensitive processing" (p. 362). These ART mechanisms illustrate how memory consolidation and novelty detection may be mediated by the same neural structures (Zola-Morgan and Squire, 1990), and clarify why hippocampectomized rats have difficulty orienting to novel cues (O'Keefe and Nadel, 1978) and why there is a progressive reduction in novelty-related hippocampal potentials as learning proceeds in normal rats (Deadwyler, West, and Lynch, 1979; Deadwyler, West, and Robinson, 1981). In summary, localization of both orienting subsystem circuits and adaptive timing circuits in, or intimately related to, the hippocampal formation helps to explain a large body of neuropsychological data. Further hippocampal relationships to ART model mechanisms will be discussed below.


We are now ready to join together the sensory-cognitive ART network in Figure 31a with the cognitive-reinforcer and adaptive timing network in Figure 30. When this is done, a striking formal similarity between the different types of circuits may be discerned. This similarity suggests that cognitive and emotional processes in the brain share many design properties in common (Grossberg, 1982b, 1984a), unlike artificial intelligence models of problem solving.

Figure 33

The sensory representations $S$ in Figure 30 are recognition codes for sensory events. For definiteness, we identify them with the recognition codes at the level $F_3$ of the ART network in Figure 31a, as in Figure 31b. When this is done, Figure 30 may be redrawn in a way that reveals a striking homology with the ART recognition circuit in Figure 31. A comparison between Figure 33a and Figure 33b illustrates this homology. In Figures 33a and 33b, the sensory representation level $F_1$ and the drive representation level $D$ play an analogous role. In particular, both level $F_1$ and level $D$ send inhibitory signals to the orienting subsystem $A$. The inhibitory signals from $F_1$ prevent $A$ from resetting STM at level $F_2$ unless a sensory input pattern mismatches a top-down learned expectation at level $F_1$. The inhibitory signals from $D$ help to prevent $A$ from resetting level $F_2$ when a reinforced event is being attended. As noted in Section 24, such an attentive focus develops due to an exchange of positive feedback signals between levels $F_2$ and $D$, supported by conditioned $F_2 \rightarrow D \rightarrow F_2$ pathways.

30. Influences of Hippocampectomy on Conditioned Timing

This synthesis of cognitive and emotional networks enables the theory to explain a broad range of data concerning changes in conditioned timing that are due to hippocampectomy. The expanded model clarifies why the hippocampus is not needed for delay conditioning, but is needed for classical conditioning of more complex associations, such as reversal conditioning and trace conditioning. It has been shown that bilateral hippocampectomy severely disrupts the rate of reversal of two-tone discrimination (Berger and Orr, 1983), reversal of cross-modality discrimination, and tone-light discrimination (Weikart and Berger, 1986). Hippocampectomy does not, however, disrupt initial learning of the discrimination (Berger, Berry, and Thompson, 1986). The deficit in reversal conditioning is consistent with the explanation of perseverative behavior due to disrupted STM reset and memory search that was given in Section 28.

Hippocampectomy has a profound effect on NM response shape during trace conditioning; for example, if a 100 msec CS duration and a 500 msec ISI are used. Then small, short-latency responses occur to the CS, rather than the large, adaptively timed long-latency responses of control animals (Solomon, Vander Schaaf, Thompson, and Weisz, 1983). The removal of the spectral timing process clarifies why the timed responses are eliminated. A
detailed study of the model circuit also clarifies why some responses remain. As Figure 29 illustrates, the interactions between sensory representations $S$ and drive representations $D$ survive ablation of the adaptive timing circuit, so that certain aspects of motivated behavior remain intact. On the other hand, the role of the adaptive timing circuit in prolonging reactions to sensory cues, and in regulating the duration of motivated attention, are no longer available. This analysis also clarifies why properties of delay conditioning are altered by hippocampectomy (Berger, Berry, and Thompson, 1986; Port, Mikhail, and Patterson, 1985).

The model gains additional support from its ability to rationalize this pattern of conditioned behavioral changes due to hippocampectomy. Although these data strongly suggest that the hippocampus plays an important role in the control of timing, they do not imply that other brain regions do not also contribute to the hippocampally observed "temporal model". The very fact that hippocampectomy alters conditioned behavioral timing indicates that timing is conditioned at hippocampal sites, as well as at non-hippocampal sites, such as the sites that control the cerebellar conditioned reactions (Section 26).


The theory developed in this article provides a computational framework in which many behavioral and neural data about conditioning can be analyzed. By identifying several problems that a behaving organism needs to solve in order to survive, the theory has been able to distinguish between several functionally distinct learning processes, to model several of their main mechanisms, and to outline a system architecture within which they are combined. These learning processes include adaptive timing, and the way in which it selectively inhibits inappropriate reactions to expected nonoccurrences; reinforcement learning, notably emotional conditioning; incentive motivational learning, including the allocation of attention and the energizing of behavioral responses; recognition learning, including the bottom-up learning that initiates selection of recognition categories and the top-down learning of expectations that help to calibrate novelty and to control memory search; and response learning, including the conditioning of discrete defensive reflexes. This sort of integrative theory exhibits features that are still quite novel in computational neuroscience. This is particularly true where the theory links together several conceptual and organizational levels in order to experimentally support its hypotheses. The theory provides behavioral analyses that help to identify functionally distinct brain processes, mathematically precise circuits that model these processes, unifying design principles to tie these circuits together into a system architecture, neural markers in terms of identifiable anatomical and physiological processes, and computer simulations and predictions to test this architecture at multiple levels of behavioral and neural organization. With such a foundation in hand, every new datum creates a series of implications that may support or confront the theory at multiple points, thereby creating multiple constraints for propelling further theoretical tests, modifications, and refinements. Such theories seem necessary if the immense masses of behavioral and neural data already available are ever to achieve a rational explanation.
REFERENCES


APPENDIX

Simulation methods: All simulations were performed on an Iris-4D/240 superworkstation using double precision representations of all values. The Iris-4D series is based upon a microcomputer that conforms to the IEEE floating point standard for accuracy. Simulations were performed by integrating the dynamical system that defines the model. Integration of this system was performed using LSODA (The Livermore Solver for Ordinary Differential equations with Automatic method-switching for stiff and non-stiff systems; (Hindmarsh, 1983; Petzold, 1983)).

The time scale was chosen to be consistent with that used in Grossberg and Schmajuk (1989). Trials were set to be 2 “units” long. In Grossberg and Schmajuk (1989), trials were set to be 2000 “milliseconds” long. All parameters from their model were preserved as they appeared in the original, except that they were multiplied by the scale factor 1000.
FIGURE CAPTIONS

Figure 1. Conditioning data from a nictitating membrane learning paradigm. Mean topography of nictitating membrane response after learning trial 10 with a 50 msec CS, ISI's of 125, 250, 500, and 1000 msec, and different (1, 2, 4 mAmp) intensities of the shock US in each subsequent panel. Reprinted from Smith (1968) with permission.

Figure 2. Computer simulation of Weber law property and inverted U in learning as a function of ISI. The output signal functions \( R(t) = \sum_i f(x_i)y_iz_i \) are plotted on a test trial, in response to the CS alone, subsequent to 10 prior learning trials with CS-US separated by different ISI's. Successive curves from left to right were generated by ISI's of 0 (the lowest amplitude curve), 125, 250, 500, and 1000 msec using a US duration of 50 msec and an \( I_{US} \) intensity of 10 units. (Reprinted from Grossberg and Schmajuk (1989) with permission.)

Figure 3. A schematic of the training regim in fixed-interval delay paradigm employed in the experiments summarized in Figure 4. After the conditioned stimulus was turned on, the first key peck after a certain interval was rewarded. During testing trials, the conditioned stimulus was turned on and remains on for an interval far longer than the expected delay, so that the subjects’ behavior at delays greater than the expected ISI could be quantified.

Figure 4. Data on pigeon key pecking in a fixed-interval delay condition. All animals were trained to respond to two different conditioned stimuli, one visual and one auditory, each of which predicted reward for the first key peck after a fixed interval. For the animals in the group whose behavior is summarized in the graph on the left, the tone signalled availability of reward after a 15 second delay and the light signalled availability of reward after a 30 second delay. For animals in the group whose behavior is summarized in the graph on the right, the tone signalled availability of reward after a 30 second delay and the light signalled availability of reward after a 15 second delay. The times at which each response curve peaks correspond closely with the times at which each key peck is of maximal value. Also, within each stimulus modality, subjects’ responses exhibit a covariation of peak time and peak breadth, as in the Weber law property shown in Figures 1 and 2.

Figure 5 A START model that combines a spectral timing module with a reinforcement learning network to achieve adaptively timed reinforcement learning and inhibition of the orienting subsystem.

Figure 6. In a START model, STM storage of a brief CS is achieved by positive feedback within the sensory representation \( s \). CS attenuation by the US is dynamically controlled by the strength of recurrent inhibitory signals. (a) Input \( I_1 \) activated by CS; (b) Input \( I_0 \) activated by US; (c) STM activation of CS sensory representation; (d) STM activation of US sensory representation.

Figure 7. Four possible feedback signal functions \( f \) for STM storage by equation (2): (a) threshold-linear signal; (b) sigmoidal signal; (c) binary signal; (d) threshold-jump-linear signal.

Figure 8. Behavior of the Now Print module of the START model after many conditioning trials: (a) Activation of the sensory representation \( S_1 \) by the CS; (b) Activation of the sensory representation \( S_0 \) by the US; (c) The resultant activation \( D \) of the drive representation \( D \); (d) The resultant Now Print signal \( N \).

Figure 9. Generation of a Now Print signal: (a) The output of a drive representation \( D \) is converted into a Now Print signal \( N \) by passing this output through a fast excitatory
pathway and a slower inhibitory pathway, whose signals converge at \( \mathcal{N} \): (b) Simulation of the activity \( D \) of \( \mathcal{D} \) in response to two successive inputs, with the first response larger; (c) Activity \( \mathcal{N} \) of \( \mathcal{N} \) scales with the size of the increment in \( D \).

**Figure 10.** Spectral timing properties of a START model. The CS\(_1\) and US stimuli, were of intensity 10 for 0.05 time units and 0 otherwise. (a) The CS-activated spectrum \( f_{ij}(t) = f(x_{1j}(t)) \); (b) The habituating transmitter gates \( y_{ij}(t) \). (c) The singly-gated spectrum \( g_{ij}(t) = f(x_{1j}(t))y_{ij}(t) \). (d) The doubly-gated spectrum \( h_{ij}(t) = f(x_{1j}(t))y_{ij}(t)z_{ij}(t) \) arising after 5 trials. (e) The corresponding output signal \( R(t) \). Simulations were performed as described in the Appendix, with parameters and signal functions given by \( \gamma = 0.2, \alpha_y = 1.0, \beta_y = 125.0, \alpha_x = 1.0, \delta = 0.0, \epsilon = 0.02, \alpha_F = 240.0, \alpha_A = 1.2, \beta_A = 120.0, \gamma_A = 12.0, \alpha_D = 120.0, \beta_D = 120.0, \gamma_D = 0.0, f_D(S) = [S - 0.05]^+, \alpha_C = 0.5, \beta_C = 25.0, f_C(D) = [D - 0.05]^+, f_A(A) = [A - 0.1]^+, F_X(A) = [A - 0.7]^+, r_j = 10.125 / (0.0125 + j); \) and the intensities of the CS and US inputs \( I_i \) in (1) equal 2.

**Figure 11.** Selective learning within different spectral populations at a fixed ISI = 500 msec. Each three-image panel from (a) to (f) represents the gated signal \( g_{ij}(t) \) [top], long term memory trace \( z_{ij}(t) \) [middle], and doubly gated signal \( h_{ij}(t) = g_{ij}(t)z_{ij}(t) \) [bottom], at a different value of \( j \). In (a), \( j = 1 \); in (b), \( j = 17 \); in (c), \( j = 33 \); in (d), \( j = 49 \); in (e), \( j = 65 \); in (f), \( j = 81 \). The same parameters as in Figure 10 were used.

**Figure 12.** The effect of self-printing upon the output of the model. (a) A large threshold \( \epsilon \) in the Now Print signal abolishes self-printing and secondary conditioning. It generates the lower output \( R(t) \). (b) A smaller threshold allows secondary conditioning and self-printing without a loss of timing. It generates the larger output \( R(t) \).

**Figure 13.** START model output \( R(t) \) during secondary conditioning with varying ISIs between the first and second CS, and between the second CS and the US using the parameters of Figure 10. Notation ISI\(_1\) below denotes the ISI between CS\(_1\) and US, and ISI\(_2\) denotes the ISI between CS\(_2\) and CS\(_1\). On each learning trial either CS\(_1\) - US or CS\(_2\) - CS\(_1\) occur, but not CS\(_2\) - CS\(_1\) - US. The curves are drawn with CS\(_1\) - US pairings in the left column and CS\(_2\) - CS\(_1\) pairings in the right column. The vertical bars occur at successive .25 time unit intervals: (a), (b) ISI\(_1\) = .25; ISI\(_2\) = .25; (c), (d) ISI\(_1\) = .5; ISI\(_2\) = .25; (e), (f) ISI\(_1\) = .25, ISI\(_2\) = .5; (g), (h) ISI\(_1\) = .5, ISI\(_2\) = .5.

**Figure 14.** A possible synaptic realization of spectral timing operations: (a) Different rates \( r_j \) can be realized at different (populations of) dentate granule cells. Each CS\(_1\) activates branching pathways whose collaterals synapse at dendritic spines across a subset of cells that include all the rates \( r_j \). The Now Print signal \( N \) is delivered in a way that can influence all active synapses across all the dentritic spines. The successive stages \( x_{ij}, y_{ij}, \) and \( z_{ij} \) of cellular activation and gating can, in principle, occur either postsynaptically, as in (b); or through a combination of presynaptic and postsynaptic operations, as in (b).

**Figure 15.** Interpretation of the output signal \( R(t) \) in equation (11) in terms of convergence of dentate granule cell outputs on CA3 pyramidal cells.

**Figure 16.** Evolution of the model's output. Output R(t) on each of the first four conditioning trials, followed by the CS alone on the fifth trial. Output after (b) 50 learning trials and (c) 100 learning trials.
Figure 17. Stability of learned timing under modifications of model parameters. (a) Effect of choosing the spectral learning rate $\alpha_z$ in (10) to $\frac{1}{2}$, 2, and 5 times that in Figure 16; (b) Effect of setting the rates $\alpha_A$, $\beta_A$, and $\gamma_A$ in (1) of the sensory representations to $\frac{1}{2}$, 1, and 2 times their values in Figure 16; (c) Effect of proportionally changing the rates $\alpha_D$, $\beta_D$, $\gamma_D$, and $\alpha_F$ in (3) and (6) to 2 and 4 times that in Figure 16 in order to speed up the Now Print circuit.

Figure 18. Stability of learned timing when strengthened inputs to the sensory representation field no longer allows STM of the CS to remain active after the US is stored in STM: (a) CS STM activity $S_1$; (b) US STM activity $S_0$; (c) spectral LTM traces after the first learning trial and in response to a CS alone on trial 25; (d) output $R(t)$ under the conditions of (c). STM parameters for this run: $\alpha_A = 0.6, \beta_A = 60.0, \gamma_A = 60.0$, US inputs $I_1$ and $I_0$ in (1) have intensities equal to 10.

Figure 19. An inverted $U$ in output intensity as a functions of CS-US ISI as produced by the START model with parameters as shown in Figure 16. This figure was produced by plotting the outputs from the model to ISI’s of 0.0, 0.125, 0.250, 0.500, and 1.000 units on a single set of axes.

Figure 20. Conditioning data from the nictitating membrane response paradigm as reported by Millenson, Kehoe, and Gormezano (1977). Data shown are average NM extensions in CS-only trials with a tone CS of length 200 ms (left-hand panels) and 700 ms (right-hand panels) in animals presented with varying mixtures of training trials: 200 ms tone CS’s immediately followed by a 50 ms shock US and 700 ms tone CS’s immediately followed by a 50 ms shock US.

Figure 21. Output $R(t)$ of the START model in a mixed ISI condition, such as in Figure 19 (Panel $P_\frac{1}{2}$) in which 50% of all ISI’s were 0.2 units and 50% of all ISI’s were 0.7 units.

Figure 22. In the Coleman and Gormezano (1971) experiments, rabbits were initially conditioned to respond to a 90 dB. 1000 Hz. tone CS by pairing it to a 4 mA., 50 msec. 60 Hz., periorbital shock US. 72 young (80-100 day old) rabbits were divided into 6 groups. Members of three of these groups were initially conditioned to respond with an ISI of 200 msec. while members of the other three groups were conditioned to respond to stimuli with an ISI of 700 msec. After five days of initial conditioning, during which time all of the animals acquired a strong response to the CS beginning shortly after its onset and peaking at roughly the time of onset of the US, the inter-stimulus intervals of some of the subjects were changed. Two groups were exposed to a sudden change from one of the ISI’s to the other (200 $\rightarrow$ 700A and 700 $\rightarrow$ 200A). Two of the groups were exposed to a gradual shift from one of the ISI’s to the other, again symmetrically (200 $\rightarrow$ 700G and 700 $\rightarrow$ 200G). As controls, the inter-stimulus intervals of two groups (200C and 700C) were held constant. Subjects’ responses were recorded over a period of four days in the experimental condition. The nictitating membrane responses under these ISI shift conditions are displayed. (From Coleman and Gormezano (1971), reprinted with permission.)

Figure 23. Simulation of the model output $R(t)$ under ISI shift conditions similar to those used in Coleman and Gormezano (1971). In the first column of each panel, the model output $R(t)$ is displayed after 25 learning trials. Successive columns are displayed after a block of 8 more learning trials. The vertical lines denote the ISI on the corresponding trial: (a) Gradual increase of ISI from .2 to .7 time units on successive learning trials; (b) Abrupt increase of ISI from .2 to .7 time units; (c) Gradual decrease of ISI from .7 to .2 time units; (d) Abrupt decrease of ISI from .7 to .2 units.
Figure 24. Effects of differing probability of reinforcement upon the timing and level of response. There is no statistically significant difference between the time at which responding peaks. (Reprinted with permission from Roberts (1981).)

Figure 25. Model output in a condition simulating the partial reinforcement paradigm of Roberts (1981). In the upper curve, 80% of all presentations of the CS were followed by presentations of the US. In the lower curve, 20% of all presentations of the CS were followed by presentations of the US. Despite this difference, the time at which the peak outputs occur is roughly equal, and only the relative levels of output are different.

Figure 26. Simulated time-averaging behavior in response to two CSs: (a) Model outputs $R(t)$ in response to the two CSs presented individually; (b) Composite model output $R(t)$ in response to the two CSs presented on the same trail at the same relay onset times as in (a). The first vertical line designates the onset time of the first CS, and the second that of the second CS.

Figure 27. Schematic conditioning circuit: conditioned stimuli ($CS_i$) activate sensory representations ($S_{CS_i}$) which compete among themselves for limited capacity short-term memory activation and storage. The activated $S_{CS}$ elicit conditionable signals to drive representations and motor command representations. Learning from a $S_{CS_1}$ to a drive representation $D$ is called conditioned reinforcer learning. Learning from $D$ to a $S_{CS_1}$ is called incentive motivational learning. Signals from $D$ to $S_{CS_1}$ are elicited when the combination of external sensory plus internal drive inputs is sufficiently large.

Figure 28. A model for conditioning in Aplysia. $SN =$ sensory neuron, $FN =$ facilitatory neuron, $IN =$ inhibitory neuron. The $SN$ play the role of sensory representations $S$, the $FN$ the role of a drive representation $D$, and the $IN$ carry out the competition between sensory representations.

Figure 29. A schematic of the Pavlovian blocking paradigm. The two phases of the experiment are discussed in the text.

Figure 30. A START model with feedback pathways $D \rightarrow S^{(2)} \rightarrow S^{(1)}$ that are capable of focusing attention in an adaptively timed fashion on reinforcing events. The sensory representations $S$ are broken into two successive levels $S^{(1)}$ and $S^{(2)}$. Levels $S^{(1)}$ and $S^{(2)}$ interact via reciprocal excitatory pathways. The excitatory pathways $S^{(1)} \rightarrow D$ and $D \rightarrow S^{(2)}$ are adaptive. Representations in $S^{(2)}$ can fire only if they receive convergent signals from $S^{(1)}$ and $D$. Then they deliver positive feedback to $S^{(1)}$ and bias the competition to focus attention on their respective features. Thus, prior to conditioning, a CS can be stored in STM at $S^{(1)}$ and can subliminally prime $S^{(2)}$ and $D$ representations without supraliminally firing these representations. After conditioning, the CS can trigger positive $S^{(1)} \rightarrow D \rightarrow S^{(2)} \rightarrow S^{(1)}$ feedback and draw attention to itself as it activates the emotional representations and motivational pathways controlled by $D$.

Figure 31. Anatomy of an adaptive resonance theory (ART) circuit: (a) Interactions between the attentional and orienting subsystems: Learning of recognition codes takes place at the long term memory (LTM) traces within the bottom-up and top-down pathways between levels $\mathcal{F}_1$ and $\mathcal{F}_2$. The top-down pathways can read-out learned expectations, or templates, that are matched against bottom-up input patterns at $\mathcal{F}_1$. Mismatches activate the orienting subsystem $A$, thereby resetting short-term memory (STM) at $\mathcal{F}_2$ and initiating search for
another recognition code. Output from subsystem $A$ can also trigger an orienting response. Sensitivity to mismatch at $F_1$ is modulated by vigilance signals from the drive representations. (b) Trainable pathways exist between level $F_2$ and the drive representations. Learning from $F_2$ to a drive representation $D$ endows a recognition category with conditioned reinforcer properties. Learning from $D$ to $F_2$ associates $D$ with a set of motivationally compatible categories. In general, more than two processing levels $F_i$ are used in ART systems.

**Figure 32.** ART search for an $F_2$ code: (a) The input pattern $I$ generates the specific STM activity pattern $X$ at $F_1$ as it nonspecifically activates the orienting subsystem $A$. Pattern $X$ both inhibits $A$ and generates the output signal pattern $S$. Signal pattern $S$ is transformed into the input pattern $T$, which activates the STM pattern $Y$ across $F_2$. (b) Pattern $Y$ generates the top-down signal pattern $U$ which is transformed into the prototype pattern $V$. If $V$ mismatches $I$ at $F_1$, then a new STM activity pattern $X^*$ is generated at $F_1$. The reduction in total STM activity which occurs when $X$ is transformed into $X^*$ causes a decrease in the total inhibition from $F_1$ to $A$. (c) If the matching criterion fails to be met, $A$ releases a nonspecific arousal wave to $F_2$, which resets the STM pattern $Y$ at $F_2$. (d) After $Y$ is inhibited, its top-down prototype signal is eliminated, and $X$ can be reinstated at $F_1$. Enduring traces of the prior reset lead $X$ to activate a different STM pattern $Y^*$ at $F_2$. If the top-down prototype due to $Y^*$ also mismatches $I$ at $F_1$, then the search for an appropriate $F_2$ code continues.

**Figure 33.** A schematic representation showing the close homology between the sensory-cognitive ART circuit shown in Figure 31 and the cognitive-emotional circuit shown in Figure 30. (a) The sensory-cognitive circuit consists of a level $F_1$ for representing activation of sensory features. Level $F_1$ interacts with a sensory representation level $F_2$ that encodes learned chunks, or compressed representations, of the sensory features. Level $F_1$ interacts with level $F_2$ via reciprocal pathways that are adaptive and excitatory. Level $F_1$ also inhibits the orienting subsystem $A$. (b) The sensory-drive ART circuit consists of a drive representation level that interacts with a sensory representation level $S$ via reciprocal pathways that are adaptive and excitatory. Level $D$ also inhibits the orienting systems $A$. The circuits (a) and (b) are combined by incorporating level $F_2$ into level $S$ as described in the text.
Figure 1
\[ \sum f(x_i) y_i z_i \]
TRAINING TRIAL

CS

US

TESTING TRIAL

CS

US

Figure 3
Figure 6
Figure 10
Figure 11
Figure 14
Figure 16
(a) ISI 0.2 → 0.7 gradual

(b) ISI 0.2 → 0.7 abrupt

(c) ISI 0.7 → 0.2 gradual

(d) ISI 0.7 → 0.2 abrupt
Blocking Paradigm

Phase 1

Phase 2

Figure 29
Figure 31
Figure 32
Figure 33