PREDICTIVE REGULATION OF ASSOCIATIVE LEARNING IN A NEURAL NETWORK BY REINFORCEMENT AND ATTENTIVE FEEDBACK

Stephen Grossberg†
Center for Adaptive Systems
Boston University
111 Cummington Street
Boston, MA 02215

and

Daniel Levine
Department of Mathematics
University of Texas
Arlington, TX 76019

and

Nestor Schmajuk‡
Center for Adaptive Systems
Boston University
111 Cummington Street
Boston, MA 02215

September, 1987

To appear in a special issue on Memory of the
International Journal of Neurology

---

† Supported in part by the Air Force Office of Scientific Research (AFOSR F49620-86-C-0037 and AFOSR F49620-87-C-0018) and the National Science Foundation (NSF IRI-84-17756).

‡ Supported in part by the National Science Foundation (NSF IRI-84-17756).

Acknowledgements: We wish to thank Cynthia Suchta and Carol Yanakakis for their valuable assistance in the preparation of the manuscript and illustrations.

ABSTRACT

A real-time neural network model is described in which reinforcement helps to focus attention upon and organize learning of those environmental events and contingencies that have predicted behavioral success in the past. Computer simulations of the model reproduce properties of attentional blocking, inverted-U in learning as a function of interstimulus interval, primary and secondary excitatory and inhibitory conditioning, anticipatory conditioned responses, attentional focussing by conditioned motivational feedback, and limited capacity short term memory processing. Qualitative explanations are offered of why conditioned responses extinguish when a conditioned excitor is presented alone, but do not extinguish when a conditioned inhibitor is presented alone. These explanations invoke associative learning between sensory representations and drive, or emotional, representations (in the form of conditioned reinforcer and incentive motivational learning), between sensory representations and learned expectations of future sensory events, and between sensory representations and learned motor commands. Drive representations are organized in opponent positive and negative pairs (e.g., fear and relief), linked together by recurrent gated dipole, or READ, circuits. Cognitive modulation of conditioning is regulated by adaptive resonance theory, or ART, circuits which control the learning and matching of expectations, and the match-contingent reset of sensory short term memory. Dendritic spines are invoked to dissociate read-in and read-out of associative learning and to thereby design a memory which does not passively decay, does not saturate, and can be actively extinguished by opponent interactions.

1. Introduction

A key problem in biological theories of intelligence concerns the manner in which external events interact with internal organismic requirements to trigger learning processes capable of focussing attention upon motivationally desired goals. The results reported herein further develop a neural theory of learning and memory (Grossberg, 1982, 1987) in which sensory-cognitive and cognitive-reinforcement circuits help to focus attention upon and organize learning of those environmental events that predict behavioral success.

The first set of results (Grossberg and Levine, 1987) describe computer simulations that show how the model reproduces properties of attentional blocking, inverted-U in learning as a function of interstimulus interval, anticipatory conditioned responses, secondary reinforcement, attentional focussing by conditioned motivational feedback, and limited capacity short-term memory processing. Conditioning occurs from sensory to drive representations ("conditioned reinforcer" learning), from drive to sensory representations ("incentive motivational" learning), and from sensory to motor representations ("habit" learning). The conditionable pathways contain long-term memory traces that obey a non-Hebbian associative law. The neural model embodies a solution of two key design problems of conditioning, the synchronization and persistence problems. This model of vertebrate learning has also been compared with data and models of invertebrate learning. Predictions derived from models of vertebrate learning have been compared with data about invertebrate learning, including data from *Aplysia* about facilitator neurons and data from *Hermissenda* about voltage-dependent Ca^{++} currents.

In the second set of results (Grossberg and Schmajuk, 1987), representations are expanded to include positive and negative opponent drive representations, as in the opponency between fear and relief. This expanded real-time neural network model is developed to explain data about the acquisition and extinction of conditioned excitors and inhibitors. Systematic computer simulations have been performed to characterize a READ circuit, which joins together a mechanism of associative learning with an opponent processing

circuit, called a *recurrent gated dipole*. READ circuit properties clarify how positive and negative reinforcers are learned and extinguished during primary and secondary conditioning. Habituating chemical transmitters within a gated dipole determine an affective adaptation level, or context, against which later events are evaluated. Neutral CS's can become reinforcers by being associated either with direct activations or with antagonistic rebounds within a previously habituated dipole. Neural mechanisms are characterized whereby conditioning can be actively extinguished, by a process called *opponent extinction*, even if no passive memory decay occurs.

READ circuit mechanisms are joined to mechanisms for associative learning of incentive motivation; for activating and storing internal representations of sensory cues in a limited capacity short term memory (STM); for learning, matching, and mismatching sensory expectancies, learning to the enhancement or updating of STM; and for shifting the focus of attention toward sensory representations whose reinforcement history is consistent with momentary appetitive requirements. This architecture has been used to explain conditioning and extinction of a conditioned excitor; conditioning and extinction of a conditioned inhibitor; properties of conditioned inhibition as a "slave" process and as a "comparator" process, including effects of pretest deflation or inflation of the conditioning context, of familiar or novel training or test contexts, of weak or strong shocks, and of preconditioning US-alone exposures. The same mechanisms have also been used (Grossberg, 1982, 1987) to explain phenomena such as unblocking, overshadowing, latent inhibition, superconditioning, partial reinforcement acquisition effect, learned helplessness, and vicious-circle behavior. The theory clarifies why alternative models have been unable to explain an equally large data base.

2. Neural Network Macrocircuits

Two types of macrocircuits control learning within the model.

Sensory-Cognitive Circuit: Sensory-cognitive interactions in the theory are carried out by an Adaptive Resonance Theory (ART) circuit (Carpenter and Grossberg, 1985, 1987a, 1987b; Grossberg, 1976, 1987). The ART architecture suggests how internal representations of sensory events, including conditioned stimuli (CS) and unconditioned stimuli (US), can be learned in stable fashion (Figure 1). Among the mechanisms used for stable self-organization of sensory recognition codes are top-down expectations which are matched against bottom-up sensory signals. When a mismatch occurs, an arousal burst acts to reset the sensory representation of all cues that are currently being stored in STM. In particular, 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.

Figure 1

Cognitive-Reinforcement Circuit: Cognitive-reinforcer interactions in the theory are carried out in the circuit described in Figure 2. In this circuit, there exist cell populations that are separate from sensory representations and related to particular drives and motivational variables (Grossberg, 1972, 1987). Repeated pairing of a CS sensory representation, S_{cs}, with activation of a drive representation, D, by a reinforcer causes the modifiable synapses connecting S_{cs} with D to become strengthened. Incentive motivation pathways from the drive representations to the sensory representations are also assumed to be conditionable. These $S \rightarrow D \rightarrow S$ feedback pathways shift the attentional focus to the set of previously reinforced, motivationally compatible cues (Figure 2). This shift of attention occurs because the sensory representations, which emit conditioned reinforcer signals and receive incentive motivation signals, compete among themselves for a limited capacity short-term memory (STM) via a shunting on-center off-surround anatomy. When

incentive motivational feedback signals are received at the sensory representational field, these signals can bias the competition for STM activity towards motivationally salient cues.

Figure 2

3. Attentional Blocking and Interstimulus Interval

The attentional modulation of Pavlovian conditioning is part of the general problem of how an information processing system can selectively process those environmental inputs that are most important to the current goals of the system. A key example is the blocking paradigm studied by Kamin (1969) (Figure 3). First, a 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, such as fear, develops. Then CS_1 and another stimulus CS_2, such as a light, are presented together, followed at the same time interval by the US. Finally, CS_2 is presented alone, not followed by a US, and no conditioned response occurs.

Figure 3

The blocking paradigm suggests four key subproblems of the selective information processing problem. These subproblems are: (1) How does the pairing of CS_1 with US in the first phase of the blocking experiment endow the CS_1 cue with properties of a conditioned, or secondary, reinforcer? (2) How do the reinforcing properties of a cue shift the focus of attention towards its own processing? (3) How does the limited capacity of attentional resources arise, so that a shift of attention towards one set of cues can prevent other cues from being attended? (4) How does withdrawal of attention from a cue prevent that cue from entering into new conditioned relationships?

The explanation of blocking also leads to an explanation of the inverted-U relationship between strength of the conditioned response (measured in one of several ways) and the time interval (ISI) between conditioned and unconditioned stimuli. Figure 4 gives an

example of experimental data on the effects of ISI from studies of Smith *et al.* (1969) and Schneiderman and Gormenzano (1964) of the rabbit nictitating membrane response. This is noteworthy because Sutton and Barto (1981) previously stated that the ISI data pose a difficulty for any network with associative synapses, that is, synapses whose efficacy changes as a function of the correlation between presynaptic and postsynaptic activities. They argued that a network with associative synapses should, to a first approximation, have an optimal ISI of zero because cross-correlation between two stimulus traces is strongest when the two stimuli occur simultaneously. To avoid this difficulty, other modellers introduced a delay in the CS pathway that was equal to the optimal ISI. But such a delay would delay the CR by an equal amount, and hence is incompatible with the so-called anticipatory CR that occurs before US onset. On this basis, Sutton and Barto suggested a different synaptic modification rule at the single-unit level.

Figure 4

Our simulations, by contrast, reproduce both the ISI data and the anticipatory CR without invoking a long delay in the CS pathway. Poor conditioning with CS and US simultaneous, or nearly so, is explained by a mechanism identical to the blocking mechanism except that CS_1 is replaced by US and CS_2 by CS. In both cases, the stimulus with more motivational significance inhibits the processing of the stimulus with less motivational significance. Poor conditioning with CS and US far apart in time occurs because by the time the US arrives, the CS representation has decayed in short-term memory to a level that is below the threshold for affecting efficacy of the appropriate synapses.

The answers to subproblems (1) to (4) are obtained from study of a network which includes modifiable associative links between sensory and drive representations (in both directions) and competitive links between different sensory representations (Figure 2). The associative links do not obey Hebb's postulate because cross-correlation is counteracted by decays; hence, synaptic efficacy can either increase or decrease with paired presynaptic and

postsynaptic activities (Grossberg, 1968, 1969, 1982), not just increase, as Hebb claimed (Hebb, 1949). Such an associative law has recently received direct neurophysiological support (Levy, Brassel, and Moore, 1983; Levy and Desmond, 1985; Rauschecker and Singer, 1979; Singer, 1983). The existence of drive representations was derived from an analysis of the *synchronization problem* (Grossberg, 1971); that is, of how a stable conditioned response can develop even if variable time lags occur between the CS and the US. These drive representations, separate from the sensory representations of particular stimuli, are what Bower has called *emotion nodes* (Bower, 1981; Bower, Gilligan, and Monteiro, 1981) and Barto, Sutton, and Anderson (1983) have called *adaptive critic elements*. A US unconditionally activates its drive representation if the drive level is sufficiently high. Repeated pairing of a CS with, for example, a food US causes pairing of stimulation of the CS sensory representation, denoted S_{CS}, with that of the representation for the hunger drive, denoted D_H. The answer to subproblem (1) therefore depends on the strengthening of $S_{CS} \rightarrow D_H$ synapses according to an associative rule.

Subproblem (2) is answered using $D_H \rightarrow S_{CS}$ incentive motivational feedback. In the blocking experment, S_{CS_1} is enhanced relative to S_{CS_2}. S_{CS_2} will thus tend to be suppressed due to competition between sensory representations that causes limited capacity of short term memory storage. Similarly, in the simultaneous phase of the ISI experiment, S_{US} is more enhanced than S_{CS}, so that S_{CS} is suppressed.

The limited capacity of short-term memory, which is needed to answer subproblem (3) arises from limited capacity properties of a recurrent on-center off-surround field, which was originally derived to satisfy a more basic processing requirement: the ability to process spatially distributed input patterns without irreparably distorting these patterns due to either noise or saturation (Ellias and Grossberg, 1975; Grossberg and Levine, 1975). Figure 2 schematizes a network with modifiable sensory-to-drive and drive-to-sensory association links and recurrent on-center off-surround links between sensory representations.

Our computer simulations, reported more completely in Grossberg and Levine (1987),

run through different stimulus conditions on the network of Figure 5, which is a variant of Figure 2 with three sensory representations, CS_1, CS_2, and US. For simplicity, there is only one drive representation, D, in our network. The $US \rightarrow D$ and $D \rightarrow US$ synapses are fixed at high value. The $CS \rightarrow D$ and $D \rightarrow CS$ synapses are strengthened by appearance of the US while the CS short term memory representation is active. In this variant of the network, sensory representations are divided into two successive stages. The activity x_{i1} of the ith first stage can activate conditioned reinforcer pathways, whereas the activity x_{i2} of the ith second stage receives conditioned incentive motivational pathways from D, and can thereupon activate x_{i1} and output motor pathways.

The same set of network parameters yielded both the ISI inverted-U curve in the case of only one CS present, and blocking in the case of two CS's. In both cases, the CR anticipated the US.

Figure 5

Our simulated ISI curves (Figure 6) were qualitatively compatible with experimental data on the rabbit's conditioned nictitating membrane response shown in Figure 4. For ISI's of fewer than 2 time units in the numerical algorithm, competition from the US representation prevented CS activity from staying above the $S_{CS} \rightarrow D$ pathway's threshold long enough to appreciably increase the pathway's strength while D was activated by the US. At long ISI's, the prior decay of the CS's short term memory trace prevented the $S_{CS} \rightarrow D$ pathway from sensing the later activation of D by the US.

Figure 6

In the blocking simulation (Figures 7a–7d), pairing of CS_1 with a delayed US enabled the long term memory trace of the $CS_1 \rightarrow D$ pathway to achieve an S-shaped cumulative learning curve. After CS_1 had become a conditioned reinforcer, it enhanced its own short term memory storage by generating a large $S_{CS_1} \rightarrow D \rightarrow S_{CS_1}$ feedback signal. As a result, when CS_1 and CS_2 were simultaneously presented, the short term memory activity

of S_{CS_2} was quickly suppressed by competition from CS_1. Consequently, the long term memory $S_{CS_2} \rightarrow D$ pathway did not grow in strength, preventing the CS_2 from being a conditioned reinforcer or eliciting a CR.

Figure 7

4. Comparison with *Aplysia* Conditioning Model

An alternative explanation of blocking, due to Hawkins and Kandel (1984), involved habituation of transmitter pathways. Based on invertebrate evidence, they developed a model whereby each US activates a *facilitator neuron* that presynaptically modulates CS pathways. They explain blocking (p.385) by saying that "the output of the facilitator neurons decreases when they are stimulated continuously". Thus after a CS_1 is paired with a US on a number of trials, subsequent presentation of a compound stimulus $CS_1 + CS_2$ with a US does not condition CS_2 because the facilitator neuron cannot fire adequately. Hawkins and Kandel's explanation, however, is incompatible with the fact (Kamin, 1969) that blocking can be overcome ("unblocked") if $CS_1 + CS_2$ is paired with either a higher or lower intensity of shock than CS_1 alone. Recent evidence (Matzel *et al.*. 1985) indicates that unblocking can also occur if the response to CS_1 is extinguished.

In our framework, the explanation for unblocking depends on *gated dipole* opponent processes that link together "positive" and "negative" drive representations (Figure 8). Positive and negative channels allow for a comparison between current and expected levels of positive or negative reinforcement. The more complete theory of Grossberg (1982, 1987) which includes gated dipoles has explained such unblocking results quantitatively.

Figure 8

In the remainder of the article, some of our computer simulation results using gated dipoles are summarized. A more systematic development is provided in Grossberg and Schmajuk (1987). Such gated dipoles are needed because, in the cognitive-reinforcement circuit, CS's are conditioned to either the onset or the offset of a reinforcer. In order to

explain how the offset of a reinforcer can generate an antagonistic rebound to which a simultaneous CS can be conditioned, gated dipoles were introduced by Grossberg (1972). A gated dipole is a minimal neural network which is capable of generating a sustained, but habituative, on-response to onset of a cue, as well as a transient off-response, or antagonistic rebound, to offset of the cue.

5. The READ Circuit: A Synthesis of Opponent Processing and Associative Learning Mechanisms

Although several varieties of a gated dipole circuit can describe the association between a CS with the onset and the offset of a reinforcer, a specialized gated dipole is needed to explain secondary inhibitory conditioning. Secondary inhibitory conditioning consists of two phases. In phase one, CS_1 becomes an excitatory conditioned reinforcer (e.g., source of conditioned fear) by being paired with a US (e.g., a shock). In phase two, the offset of CS_1 can generate an off-response which can condition a subsequent CS_2 to become an inhibitory conditioned reinforcer (e.g., source of conditioned relief). In order to explain secondary inhibitory conditioning, a gated dipole circuit must also contain internal feedback pathways, i.e., it should be recurrent. In addition, such a recurrent gated dipole must be joined to a mechanism of associative learning. The total circuit that we have analyzed is called a READ circuit, as a mnemonic for REcurrent Associative gated Dipole (Figure 9).

Figure 9

The equations for the READ circuit are as follows:

Arousal + US + Feedback On-Activation:

$$\frac{d}{dt}x_1 = -A_1x_1 + I + J + T(x_7) \tag{1}$$

Arousal + Feedback Off-Activation:

$$\frac{d}{dt}x_2 = -A_2x_2 + I + T(x_8) \tag{2}$$

On-Transmitter:

$$\frac{d}{dt}y_1 = B(1 - y_1) - Cg(x_1)y_1 \tag{3}$$

Off-Transmitter:

$$\frac{d}{dt}y_2 = B(1 - y_2) - Cg(x_2)y_2 \tag{4}$$

Gated On-Activation:

$$\frac{d}{dt}x_3 = -A_3x_3 + Dg(x_1)y_1 \tag{5}$$

Gated Off-Activation:

$$\frac{d}{dt}x_4 = -A_4x_4 + Dg(x_2)y_2 \tag{6}$$

Normalized Opponent On-Activation:

$$\frac{d}{dt}x_5 = -A_5x_5 + (E - x_5)x_3 - (x_5 + F)x_4 \tag{7}$$

Normalized Opponent Off-Activation:

$$\frac{d}{dt}x_6 = -A_6x_6 + (E - x_6)x_4 - (x_6 + F)x_3 \tag{8}$$

Total On-Activation:

$$\frac{d}{dt}x_7 = -A_7x_7 + G[x_5]^+ + L\sum_{k=1}^{n} S_k z_{k7} \tag{11}$$

Total Off-Activation:

$$\frac{d}{dt}x_8 = -A_8x_8 + G[x_6]^+ + L\sum_{k=1}^{n} S_k z_{k8} \tag{12}$$

On-Conditioned Reinforcer Association:

$$\frac{d}{dt}z_{k7} = S_k[-Hz_{k7} + K[x_5]^+] \tag{13}$$

Off-Conditioned Reinforcer Association:

$$\frac{d}{dt}z_{k8} = S_k[-Hz_{k8} + K[x_6]^+] \tag{14}$$

On-Output Signal:

$$O_1 = [x_5]^+ \tag{15}$$

Off-Output Signal:

$$O_2 = [x_6]^+, \tag{16}$$

where the notation $[x_i]^+$ denotes a linear signal above the threshold value zero; that is, $\max(x_i, 0)$.

In the equations, I denotes the tonic arousal level, J the US input, S_k the k^{th} CS, z_{k7} and z_{k8} the association of the k^{th} CS with the on- and the off-response, respectively. $A, B, C, D, E, F, G, H, K$, and L are parameter values, which were kept constant for all simulations. When $E = F$, x_5 and x_6 compute an opponent process and a ratio scale at the same time. Thus one key property of the READ circuit is associative averaging, rather than summation.

6. Opponent Extinction by Dissociating Long Term Memory Read-In and Read-Out at Dendritic Spines

A second key property of the READ circuit has been called *opponent extinction*. Although passive memory decay does not occur in the parameter ranges which we used, when the net signals in the on- and off-channels are balanced, then $x_5 = 0 = x_6$, and therefore z_{k7} and z_{k8} approach 0. The LTM traces hereby continually readjust themselves to the *net imbalance* between the on- and off-channels. Opponent extinction avoids the possible saturation at maximal values of both LTM traces z_{k7} and z_{k8}.

A third key property of the READ circuit is a dissociation between read-in and read-out of long-term memory (LTM), as in Figure 10. For example, in the on-channel, read-out

is proportional to $[x_7]^+$, whereas read-in is proportional to $[x_5]^+$. Grossberg (1975) proposed that such dissociation can be physiologically implemented by assuming that synaptic plasticity occurs at the dendritic spines of neural cells. Signal $[x_5]^+$ is assumed to cause a global potential change that invades all the spines inducing plastic changes throughout the dendritic column, as in equation (13). However, due to the geometry and electrical properties of the dendritic tree, an input that activates a particular dendritic branch may not be influenced by inputs that activate different dendritic branches. Activation at a particular dendritic branch would produce local potentials that propagate to the cell body where they influence axonal firing via potential x_7 in equation (11).

Figure 10

7. Computer Simulations of Primary and Secondary Conditioning

This section summarizes computer simulations in different classical conditioning paradigms. Although the simulations show the competence of the READ circuit in these paradigms, additional neural machinery (such as the ART circuit in Figure 1) is necessary to explain some difficult conditioning data.

Excitatory primary conditioning. Because the CS is presented in the presence of the US, it becomes associated with the on-response. Variable CS_1-ON describes conditioning of the LTM trace z_{17} within the pathway from the sensory representation of CS_1 to the on-channel. After 10 acquisition trials, presentations of CS_1 alone do not cause extinction of the CS_1-ON association (Figure 11). As explained later in the text, forgetting of CS_1–ON associations is due to the acquisition of CS_1-OFF associations.

Figure 11

Inhibitory primary conditioning. Because the CS is presented after the US offset, it becomes associated with the off-response. Variable CS_1-OFF describes conditioning of the LTM trace z_{18} within the pathway from the sensory representation of CS_1 to the off-channel. After 10 acquisition trials, presentations of CS_1 alone cause the CS_1-OFF

association to relax to a persistent remembered value (Figure 12). As explained later in the text, forgetting of the CS_1-OFF association is due to the acquisition of CS_1-ON associations.

Figure 12

In Grossberg and Schmajuk (1987), the following types of secondary conditioning phenomena are also simulated:

Excitatory secondary conditioning. The LTM trace CS_1-ON grows during the first 10 trials and is then used to induce the growth of the LTM trace CS_2-ON during the next 10 trials.

Inhibitory secondary conditioning. The LTM trace CS_1-ON grows during the first 10 trials and is then used, by presenting a CS_2 after CS_1 offset, to induce the growth of the LTM trace CS_2-OFF during the next 10 trials.

8. Qualitative Explanations of Extinction and Non-Extinction Data

This section presents qualitative explanations for some difficult conditioning data that require additional neural machinery, such as STM attentional modulation and STM reset by expectancy mismatch by an ART circuit.

Excitatory conditioning and extinction. When a CS is paired with an aversive US on successive conditioning trials, the sensory representation S_1 of CS_1 is conditioned to the drive representation D_{on} corresponding to the fear reaction, both through its conditioned reinforcer path $S_1 \rightarrow D_{on}$ and through its incentive motivational path $D_{on} \rightarrow S_1$. As a result, later presentations of CS_1 tend to generate an amplified STM activation of S_1, and thus CS_1 is preferentially attended. Due to the limited capacity of STM less salient cues tend to be attentionally blocked when CS_1 is presented.

As the cognitive-motivational feedback loop $S_1 \rightarrow D_{on} \rightarrow S_1$ is strengthened during conditioning trials, S_1 is also associated to a sensory expectation of the shock within an

ART circuit. During extinction, S_1 is presented on unshocked trials. Parameters of the READ cicuit are chosen to prevent passive decay of LTM traces from occurring on these trials. However, when the expected shock does not occur, a mismatch occurs with the learned expectation read-out by S_1, the STM activity of S_1 is reduced by the consequent STM reset, and an antagonistic rebound occurs in the off-channel of the READ circuit. Consequently, S_1 is associated to an antagonistic rebound at D_{off}. Because S_1 is smaller after reset than before, $S_1 \rightarrow D_{off}$ associations take place at a slower rate than during conditioning. After several learning trials, however, the pathway $S_1 \rightarrow D_{off}$ is as strong as the $S_1 \rightarrow D_{on}$ pathway, and opponent extinction occurs.

Inhibitory conditioning and non-extinction. Suppose that CS_1 has become a conditioned excitor, and that CS_1 and CS_2 are presented together in absence of the US. When CS_1 and CS_2 are simultaneously presented (Figure 13), S_1's activity is amplified by positive feedback through the strong conditioned $S_1 \rightarrow D_{on} \rightarrow S_1$ pathway. As a result of the limited capacity of STM, the STM activity of S_2 is blocked at time T_1. When the expected US does not occur at time T_2, the mismatch with S_1's sensory expectation causes both S_1 and S_2 to be reset, and S_1's STM activity decreases while S_2's STM activity increases. Due to S_1's decrease, a rebound occurs at D_{off}. Consequently, the unexpected nonoccurrence of the shock enables S_2 to become associated with D_{off} in both the pathways $S_2 \rightarrow D_{off}$ and $D_{off} \rightarrow S_2$. These are the primary cognitive-motivational conditioning events that turn CS_2 into a conditioned inhibitor.

Figure 13

According to the READ circuit, when presented alone the conditioned value of $CS_2 \rightarrow D_{off}$ persists. No further extinction occurs because the CS_2 sensory expectation predicts the absence of the US. Thus when presented alone, CS_2 does not disconfirm its sensory expectation, and S_2's STM activity is not reset.

9. Conclusion

At least four types of learning processes are relevant in the present paper: learning of conditioned reinforcement, incentive motivation, sensory expectancy, and motor command. These several types of learning processes, which operate on a slow time scale, regulate and are regulated by rapidly fluctuating limited capacity STM representations of sensory events. The theory suggest how nonlinear feedback interactions among these fast information processing mechanisms and slow learning mechanisms participate in different conditioning paradigms, and actively regulate learning and memory to generate predictive internal representations of external environmental contingencies.

References

Barto, A.G., Sutton, R.S., and Anderson, C.W. (1983). Neuron-like adaptive elements that can solve difficult learning control problems. *IEEE Transactions*, **SMC-13**, 834–846.

Bower, G.H. (1981). Mood and memory. *American Psychologist*, **36**, 129–148.

Bower, G.H., Gilligan, S.G., and Monteiro, K.P. (1981). Selectivity of learning caused by adaptive states. *Journal of Experimental Psychology: General*, **110**, 451–473.

Carpenter, G.A. and Grossberg, S. (1985). Category learning and adaptive pattern recognition: A neural network model. *Proceedings of the Third Army Conference on Applied Mathematics and Computing*, ARO Report 86-1, 37–56.

Carpenter, G.A. and Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37**, 54–115.

Carpenter, G.A. and Grossberg, S. (1987b). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, in press.

Carpenter, G.A. and Grossberg, S. (1987c). Neural dynamics of category learning and recognition: Structural invariants, reinforcement, and evoked potentials. In M.L. Commons, S.M. Kosslyn, and R.J. Herrnstein (Eds.), **Pattern recognition and concepts in animals, people, and machines**. Hillsdale, NJ: Erlbaum.

Ellias, S. And Grossberg, S. (1975). Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biological Cybernetics*, **20**, 69–98.

Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences*, **60**, 758–765.

Grossberg, S. (1969). On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics*, **1**, 319–350.

Grossberg, S. (1971). On the dynamics of operant conditioning. *Journal of Theoretical Biology*, **33**, 225–255.

Grossberg, S. (1972). A neural theory of punishment and avoidance, I: Qualitative theory. *Mathematical Biosciences*, **15**, 39–67.

Grossberg, S. (1975). A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology*, **18**, 263–327.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, **23**, 187–202.

Grossberg, S. (1982). **Studies of mind and brain: Neural principles of learning, perception, development, cognition, and motor control.** Boston: Reidel Press.

Grossberg, S. (Ed.) (1987). **The adaptive brain, I: Cognition, learning, reinforcement, and rhythm.** Amsterdam: Elsevier/North-Holland.

Grossberg, S. and Levine, D.S. (1975). Some developmental and attentional biases in the contrast enhancement and short term memory of recurrent neural networks. *Journal of Theoretical Biology*, **45**, 341–380.

Grossberg, S. and Levine, D.S. (1987). Neural dynamics of attentionally-modulated Pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. *Applied Optics*, in press.

Grossberg, S. and Schmajuk, N.A. (1987). A neural network architecture for attentionally-modulated Pavlovian conditioning: Conditioned reinforcement, inhibition, and opponent processing. *Psychobiology*, in press.

Hawkins, R.D. and Kandel, E.R. (1984). Is there a cell-biological alphabet for simple forms of learning? *Psychological Review*, **9**, 375–391.

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

Kamin, L.J. (1969). Predictability, surprise, attention, and conditioning. In B.A. Campbell and R.M. Church (Eds.), **Punishment and aversive behavior**. New York: Appleton-Century-Crofts.

Levy, W.B., Brassel, S.E., and Moore, S.D. (1983). Partial quantification of the associative synaptic learning rule of the dentate gyrus. *Neuroscience*, **8**, 799–808.

Levy, W.B. and Desmond, N.L. (1985). The rules of elemental synaptic plasticity. In W.B. Levy, J. Anderson, and S. Lehmkuhle (Eds.), **Synaptic modification, neuron selectivity, and nervous system organization**. Hillsdale, NJ: Erlbaum, pp.105–121.

Matzel, L.D., Schachtman, T.R., and Miller, R.R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation*, **16**, 398–412.

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

Schneiderman, N. and Gormenzano, I. (1964). Conditioning of the nictitating membrane response of the rabbit as a function of the CS-US interval. *Journal of Comparative and Physiological Psychology*, **57**, 188–195.

Singer, W. (1983). Neuronal activity as a shaping factor in the self-organization of neuron assemblies. In E. Basar, H. Flohr, H. Haken, and A.J. Mandell (Eds.), **Synergetics of the brain**. New York: Springer-Verlag.

Smith, M.D., Coleman, S.R., and Gormenzano, I. (1969). Classical conditioning of the rabbit's nictitating membrane response at backward, simultaneous, and forward CS-US intervals. *Journal of Comparative and Physiological Psychology*, **69**, 226–231.

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

FIGURE CAPTIONS

Figure 1. Anatomy of an adaptive resonance theory (ART) circuit: (a) Interactions between the attentional and orienting subsystems. Code learning takes place at the long term memory (LTM) traces within the bottom-up and top-down pathways between levels F_1 and F_2. The top-down pathways can read-out learned expectations, or templates, that are matched against bottom-up input patterns at F_1. Mismatches activate the orienting subsystem A, thereby resetting short term memory (STM) at F_2 and initiating search for another recognition code. Subsystem A can also activate an orienting response. Sensitivity to mismatch at F_1 is modulated by vigilance signals from drive representations. (b) Trainable pathways exist between level F_2 and the drive representations. Learning from F_2 to a drive representation endows a recognition category with conditioned reinforcer properties. Learning from a drive representation to F_2 associates the drive representation with a set of motivationally compatible categories. (Adapted from Carpenter and Grossberg, 1987c.)

Figure 2. 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_i} elicit conditioned signals to drive representations and motor command representations. Learning from an S_{cs_i} to a drive representation D is called conditioned reinforcer learning. Learning from D to S_{cs_i} is called incentive motivational learning. Signals from D to S_{cs_i} are elicited when the combination of external sensory plus internal drive inputs is sufficiently large. In the simulations reported herein, the drive level is assumed to be large and constant.

Figure 3. A blocking paradigm. The two stages of the experiment are discussed in the text.

Figure 4. Experimental relationship between conditioned response strength (measured by percentage of trials on which response occurs) and interstimulus interval in the rabbit nictitating membrane response. (Reprinted with permission from Sutton and Barto,

1981.)

Figure 5. Simulated network: Each sensory representation possesses two stages with STM activities x_{i1} and x_{i2}. A CS or US input activates its corresponding x_{i1}. Activation of x_{i1} elicits unconditionable signals to x_{i2} and conditioned reinforcer signals to D, whose activity is denoted by y. Incentive motivational feedback signals from D activate the second stage potentials x_{i2}, which then send feedback signals to x_{i1}. Conditionable long-term memory traces are designated by hemi-disks.

Figure 6. Plot of CR acquisition speed as a function of ISI. This speed was computed by the formula 100 × (number of time units per trial)/(number of time units to first CR).

Figure 7. Blocking simulation: In (a)-(d), the ISI = 6 between CS_1 and US onset. Five trials of CS_1-US pairing are followed by five trials of $(CS_1 + CS_2)$-US pairing. Then CS_2 is presented alone for one trial. (a) Activity x_{11} of S_{CS_1} through time; (b) Activity x_{21} of S_{CS_2} through time; (c) LTM trace z_{11} from S_{CS_1} to D through time; (d) LTM trace z_{21} from S_{CS_2} to D through time.

Figure 8. Example of a feedforward gated dipole: A sustained habituating on-response (top left) and a transient off-rebound (top right) are elicited in response to onset and offset, respectively, of a phasic input J (bottom left) when tonic arousal I (bottom center) and opponent processing (diagonal pathways) supplement the slow gating actions (square synapses). See text for details.

Figure 9. A READ I circuit: This circuit joins together a recurrent gated dipole with an associative learning mechanism. Learning is driven by signals S_k from sensory representations S_k which activate long term memory (LTM) traces z_{k7} and z_{k8} that sample activation levels at the on-channel and off-channel, respectively, of the gate dipole. See text for details.

Figure 10. A possible microarchitecture for dissociation of LTM read-in and read-out: Individual LTM-gated sensory signals $S_k z_{k7}$ are read-out into local potentials which are

summed by the total cell body potential x_7 without significantly influencing each other's learned read-in. In contrast, the input signal x_5 triggers a massive global cell activation which drives learned read-in at all active LTM traces abutting the cell surface. Signal x_5 also activates the cell body potential x_7.

Figure 11. Computer simulation of primary excitatory conditioning and extinction with slow habituation and large feedback in a READ I circuit: CS_1 is paired with the US during the first 10 simulated trials, and CS_1 is presented in the absence of the US in the next 10 simulated trials. The numbers above each plot are the maximum and minimum values of the plot. Parameters are $A = 1, B = .005, C = .00125, D = 20, E = 20, F = 20, G = .5, H = .005, K = .025, L = 20, M = .05$.

Figure 12. Computer simulation of primary inhibitory conditioning and extinction with slow habituation and large feedback in a READ I circuit: CS_1 is presented after the US offset during the first 10 simulated trials, and CS_1 is presented in the absence of the US in the next 10 simulated trials. The same parameters were used as in Figure 11.

Figure 13. Presentation of CS_1 and CS_2 when CS_1 has become a conditioned excitor and the compound stimulus is followed by no-shock: During the no-shock interval between times T_1 and T_2, S_1 is actively amplified by positive feedback and S_2 is blocked. Nonoccurrence of the expected shock causes both S_1 and S_2 to be reset. S_1's STM activity decreases and S_2's STM activity increases. Due to S_1's increase, D_{on} also decreases, thereby causing a rebound at D_{off}. This rebound becomes associated with the increased activity of S_2.

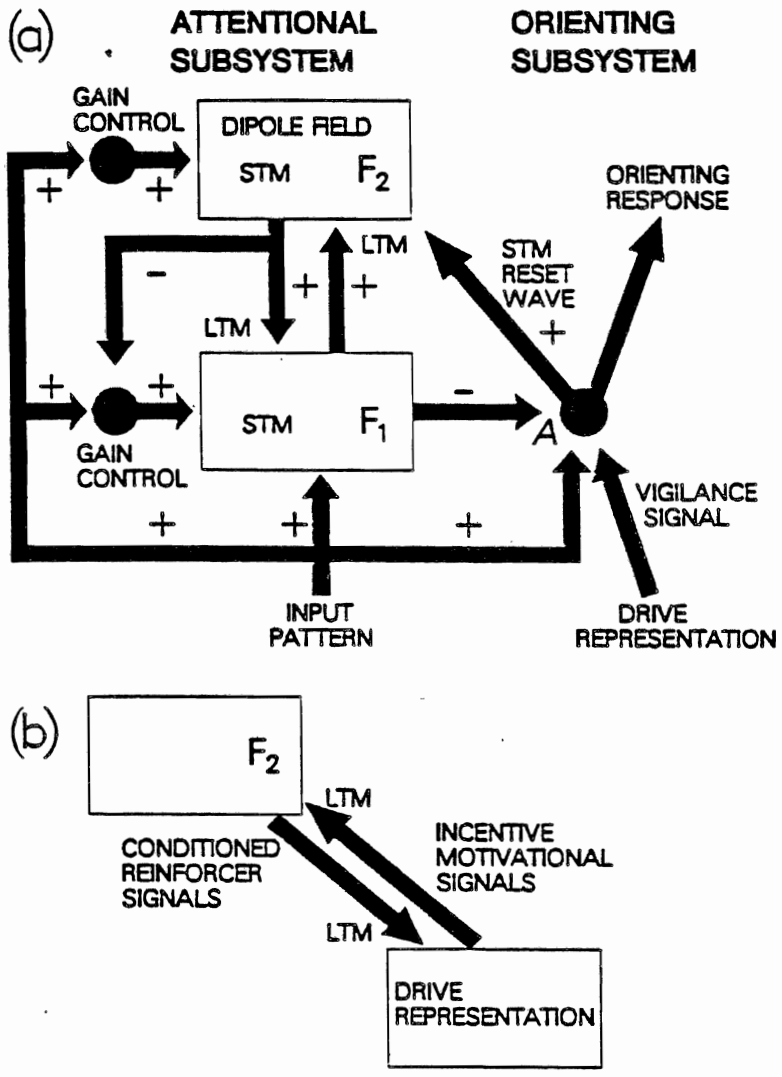
(a)
ATTENTIONAL SUBSYSTEM
ORIENTING SUBSYSTEM
GAIN CONTROL
DIPOLE FIELD
STM F_2
ORIENTING RESPONSE
LTM
STM RESET WAVE
LTM
STM F_1
A
GAIN CONTROL
VIGILANCE SIGNAL
INPUT PATTERN
DRIVE REPRESENTATION
(b)
F_2
LTM
INCENTIVE MOTIVATIONAL SIGNALS
CONDITIONED REINFORCER SIGNALS
LTM
DRIVE REPRESENTATION

Figure 1

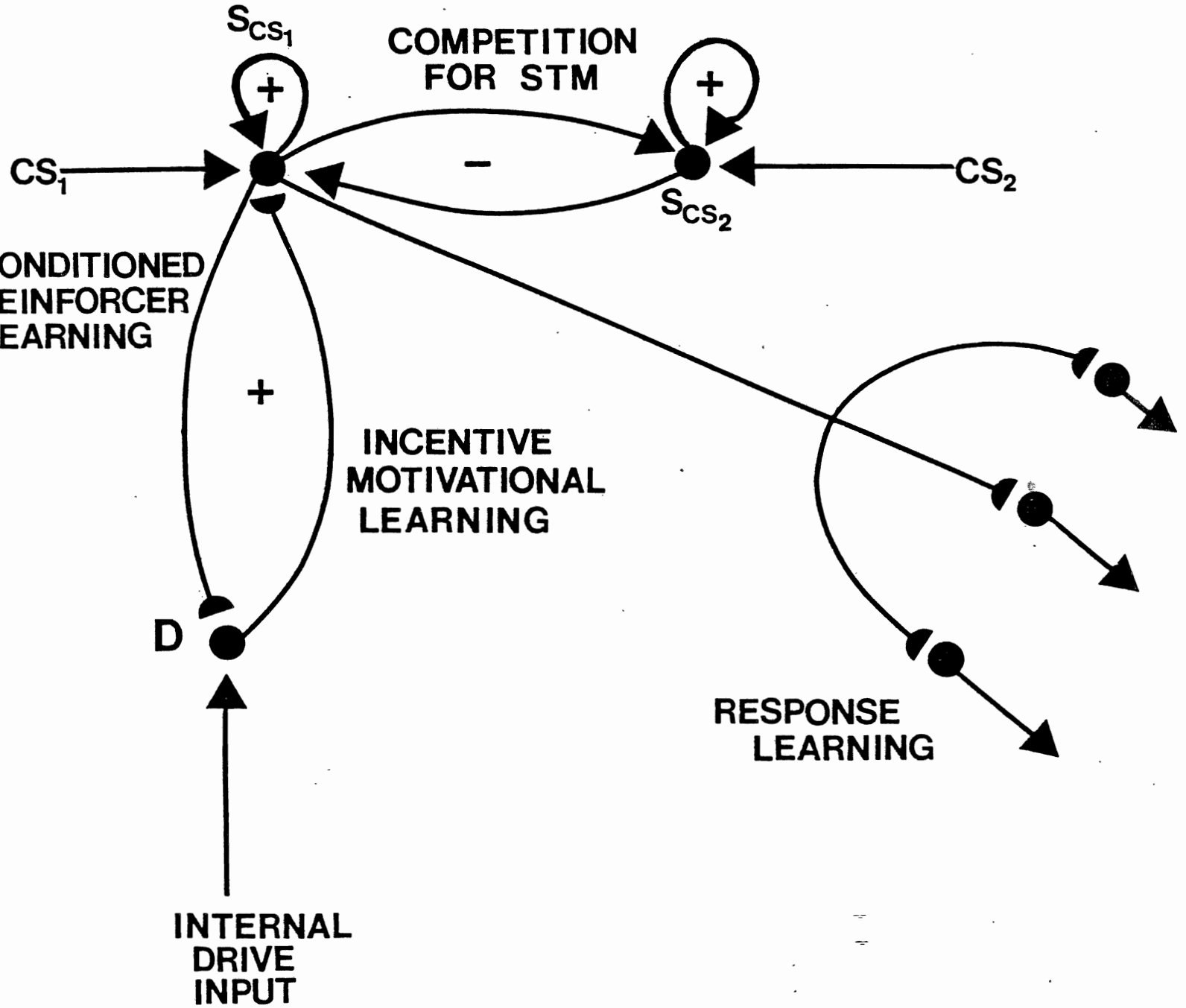
S_{CS_1}
COMPETITION
FOR STM
+
+
−
CS_1
CS_2
S_{CS_2}
ONDITIONED
EINFORCER
EARNING
+
INCENTIVE
MOTIVATIONAL
LEARNING
D
RESPONSE
LEARNING
INTERNAL
DRIVE
INPUT

Figure 2

1. $CS_1 \text{ —— } US$

$CS_1 \longrightarrow CR$

2. $CS_1 + CS_2 \text{ —— } US$

$CS_2 \not\longrightarrow CR$

Figure 3

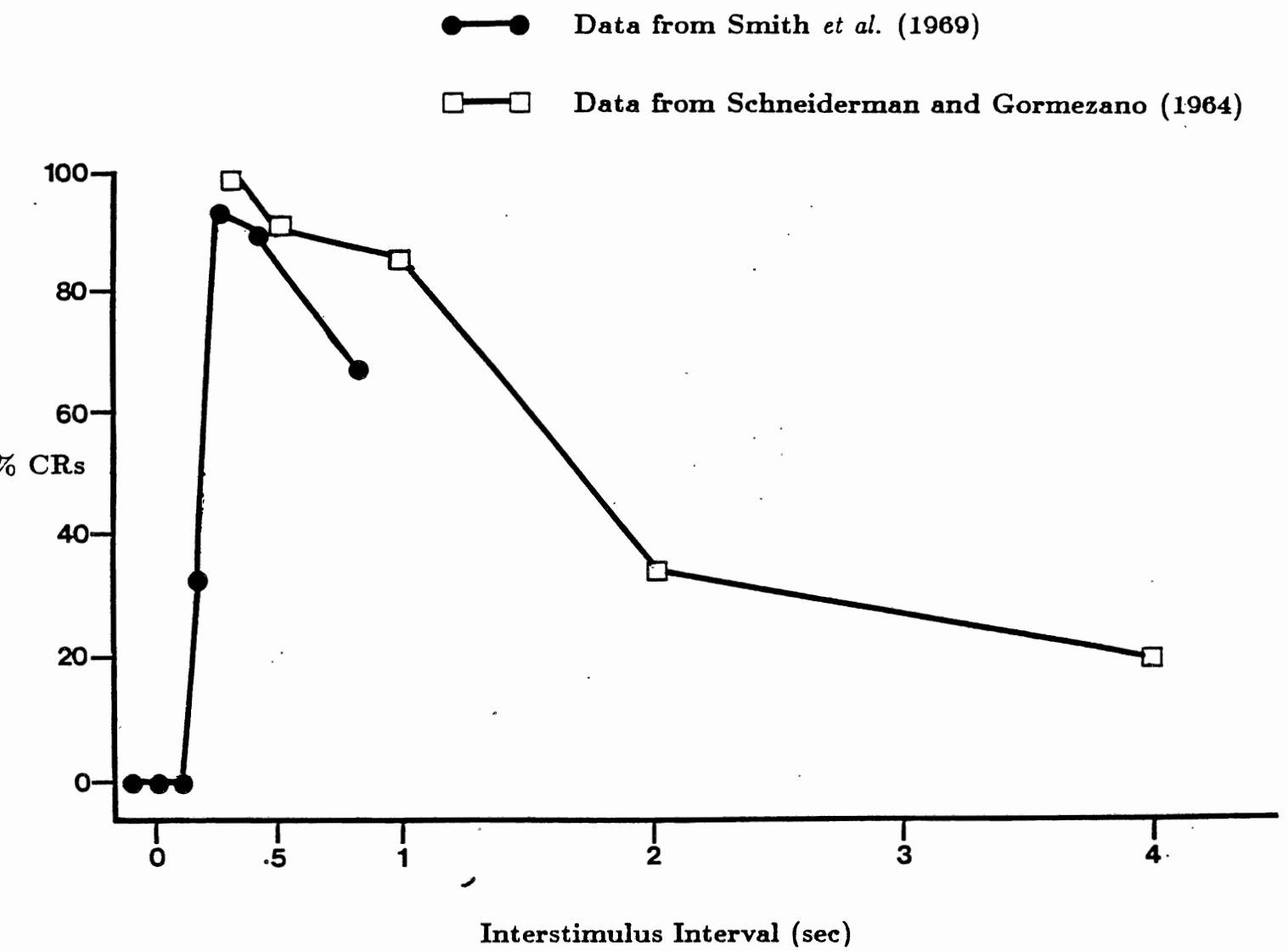
Data from Smith et al. (1969)
Data from Schneiderman and Gormezano (1964)
100
80
60
% CRs
40
20
0
0
.5
1
2
3
4
Interstimulus Interval (sec)

Figure 4

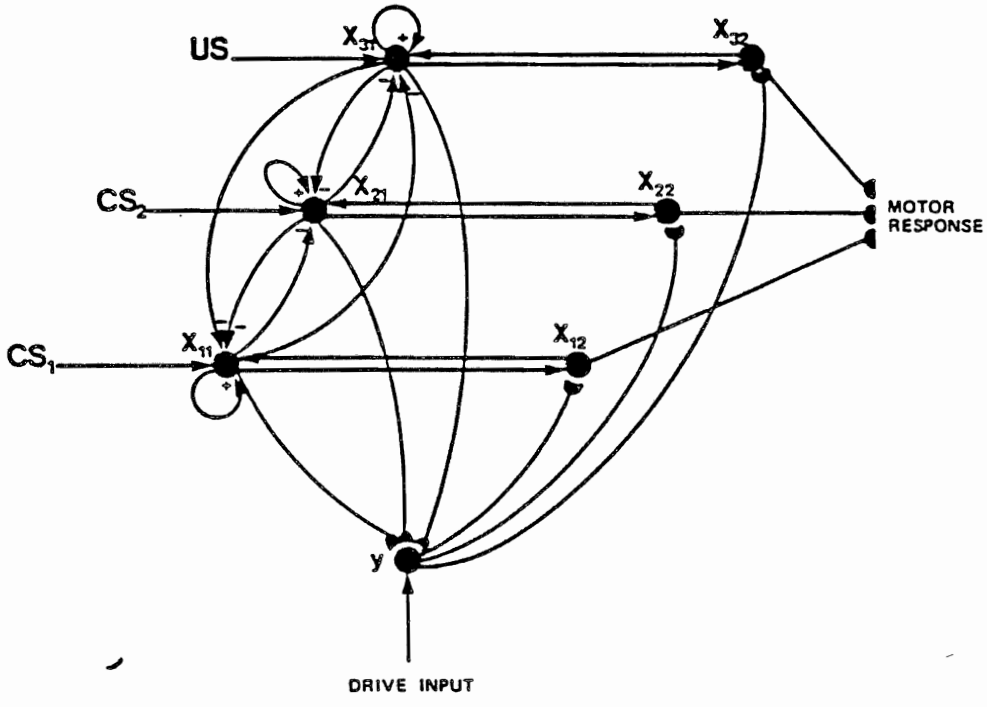
US
X_{31}
X_{32}
CS_2
X_{21}
X_{22}
MOTOR RESPONSE
CS_1
X_{11}
X_{12}
y
DRIVE INPUT

Figure 5

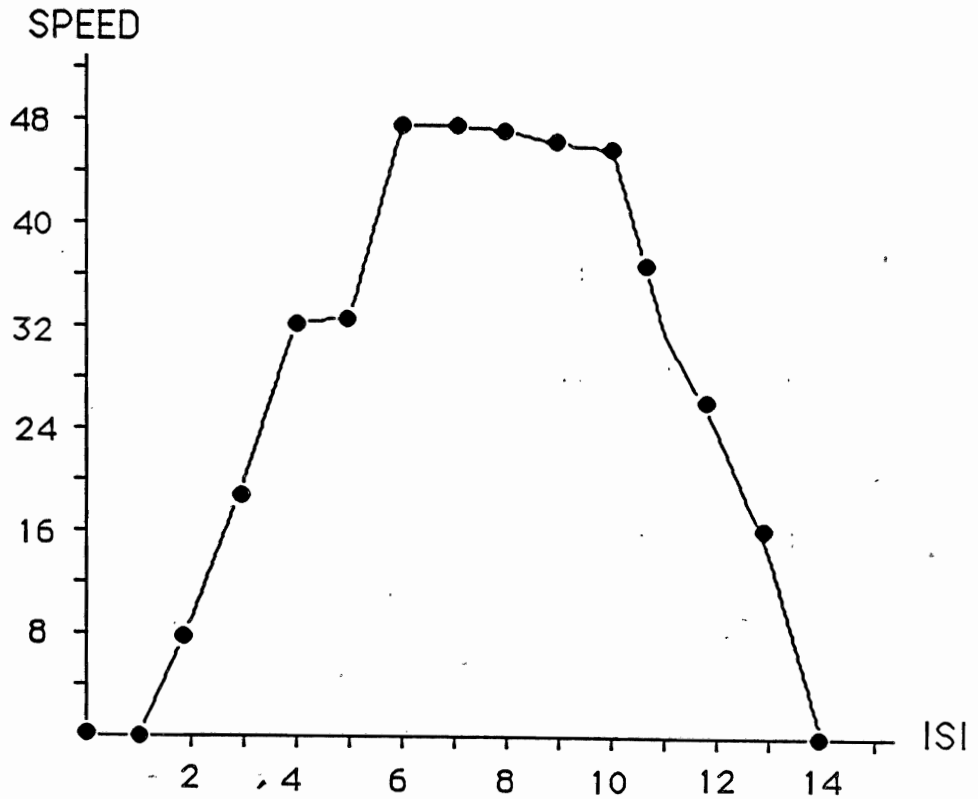

Figure 6

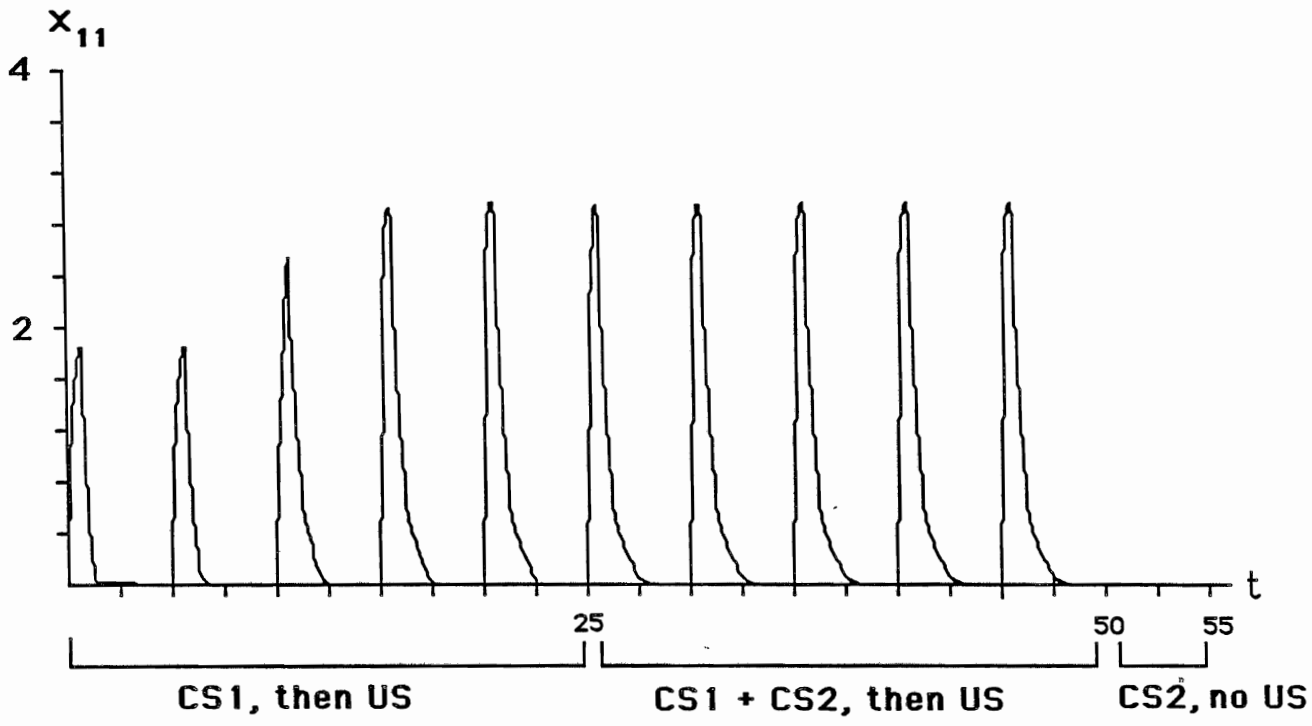

Figure 7a

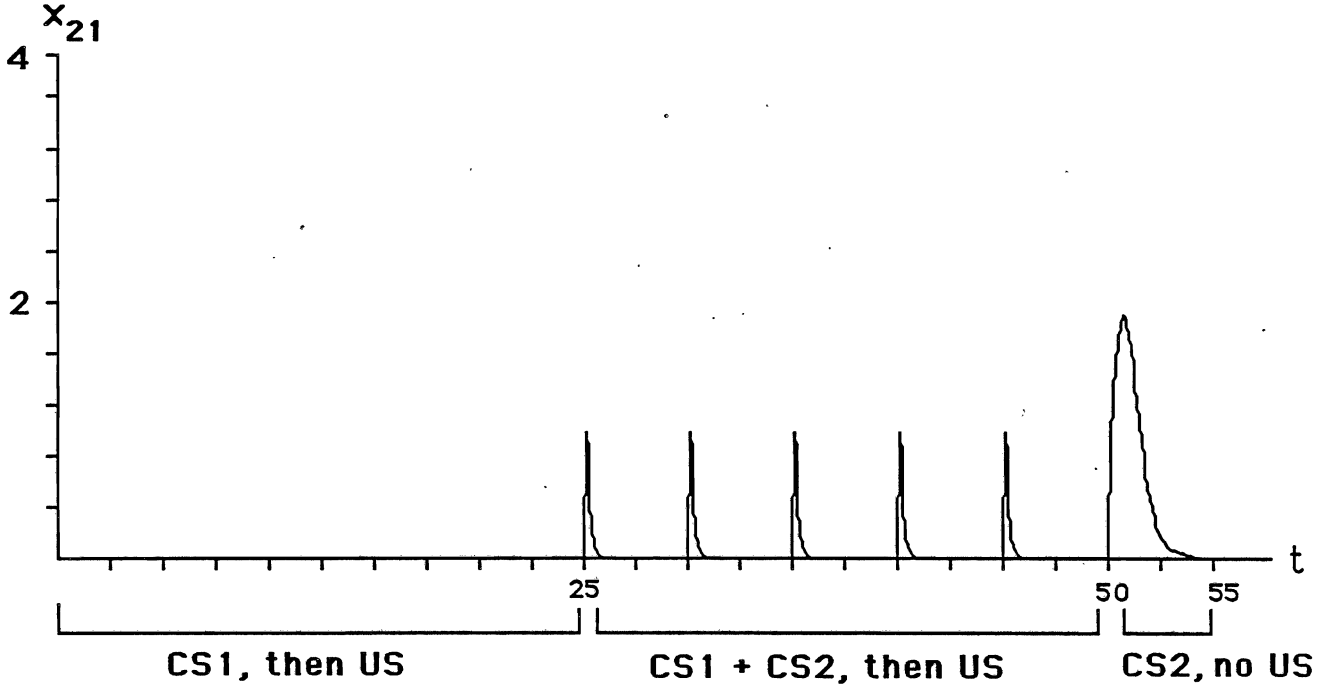

Figure 7b

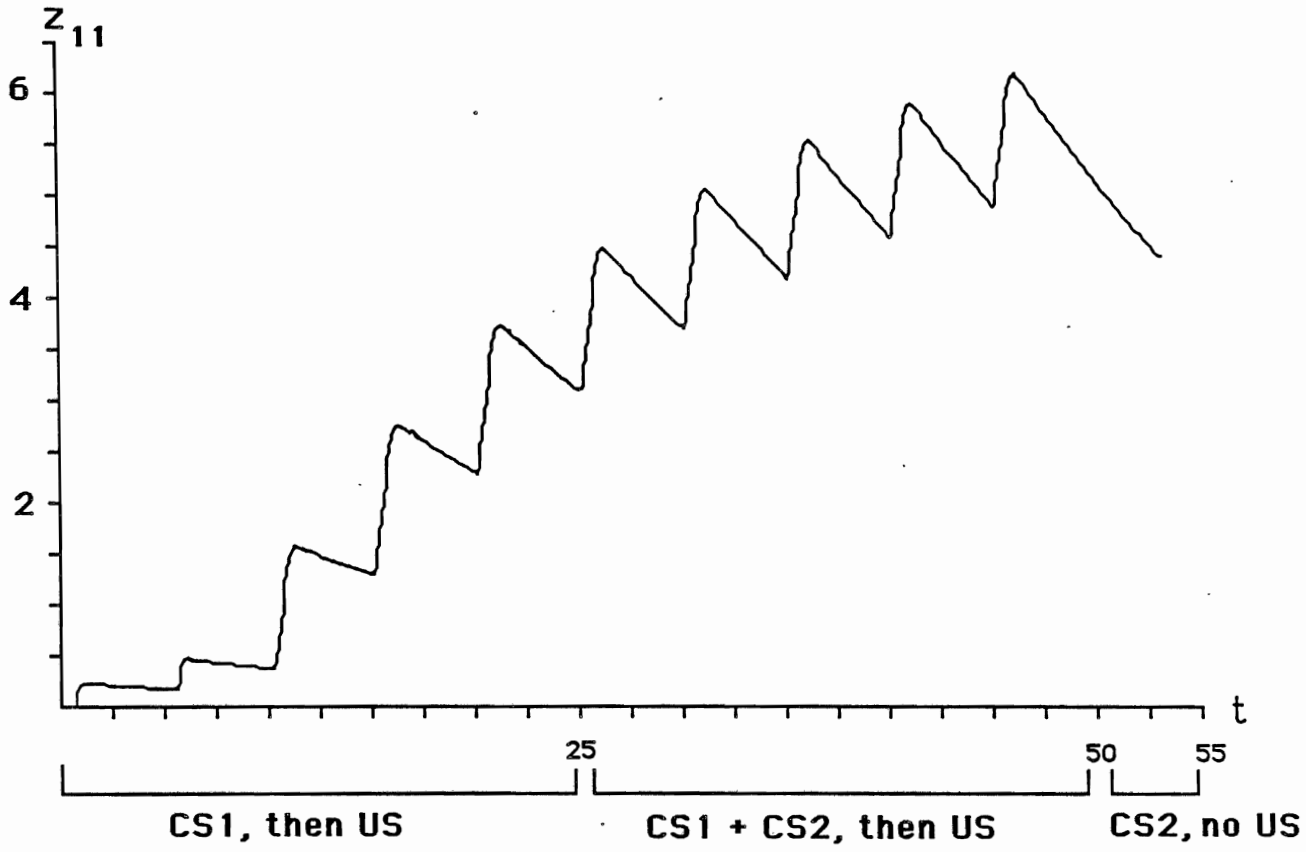

Figure 7c

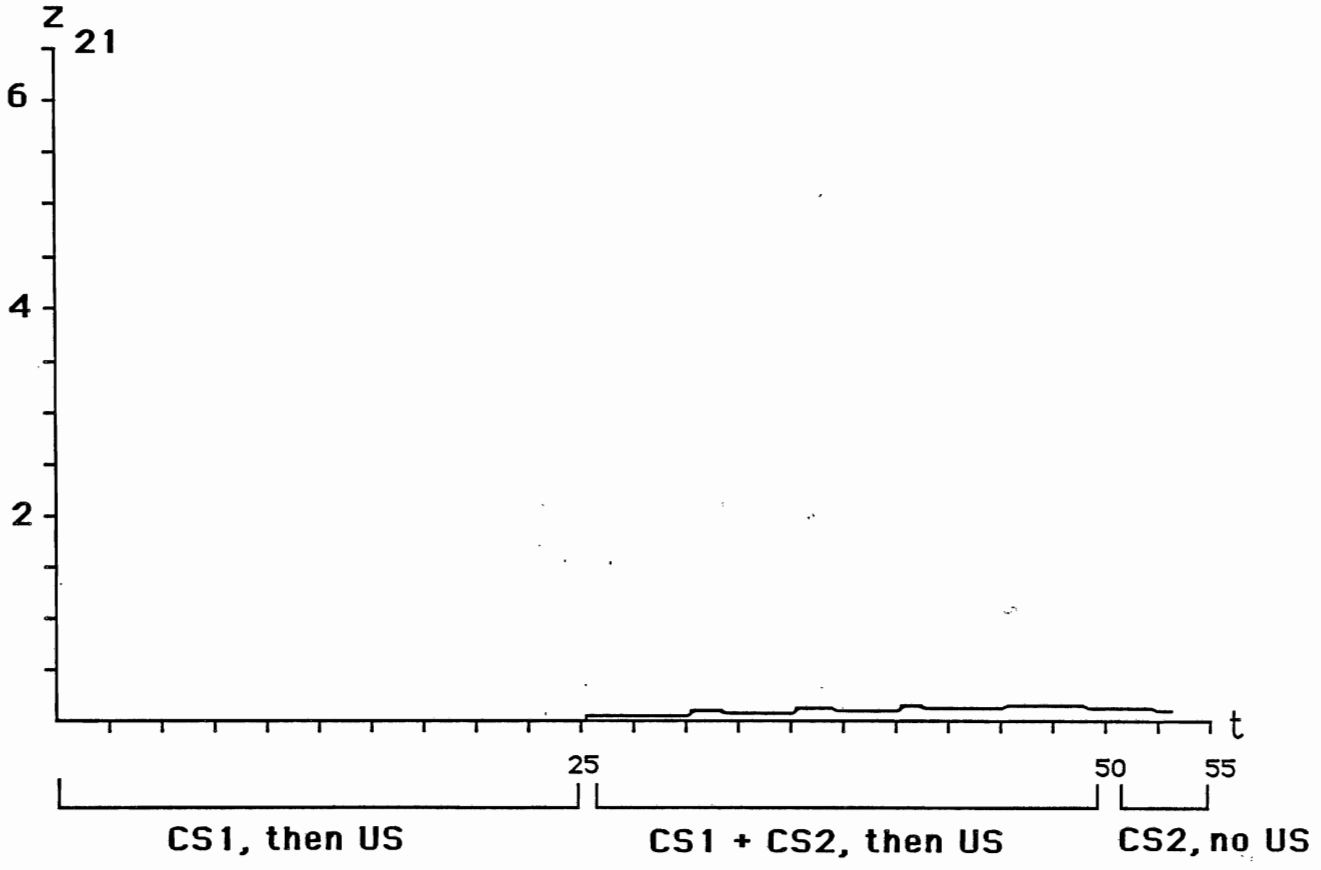

Figure 7d

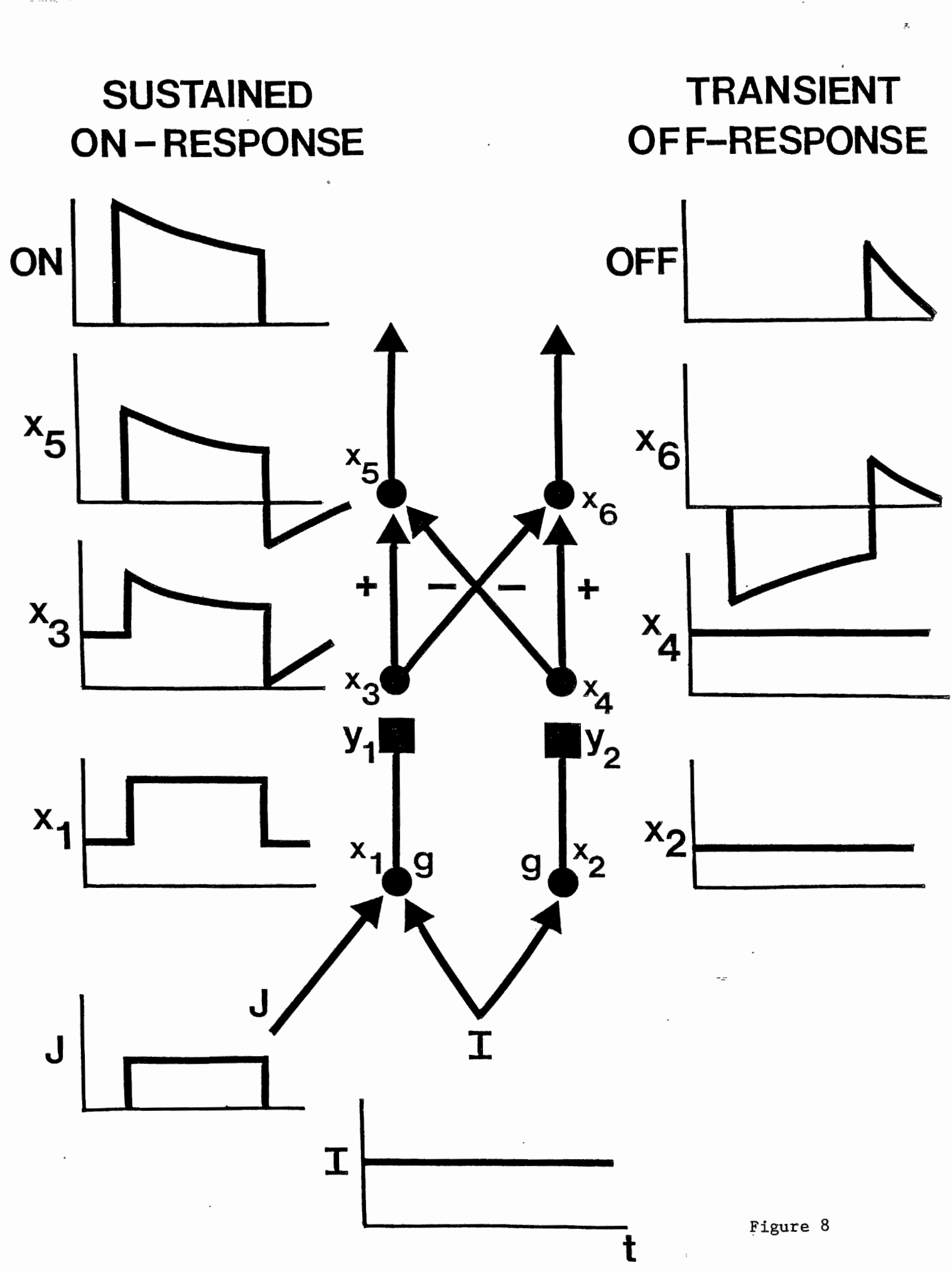
SUSTAINED
ON-RESPONSE
TRANSIENT
OFF-RESPONSE
ON
OFF
x5
x6
x3
x4
x1
x2
J
I
t
+
−
−
+
g
g

Figure 8

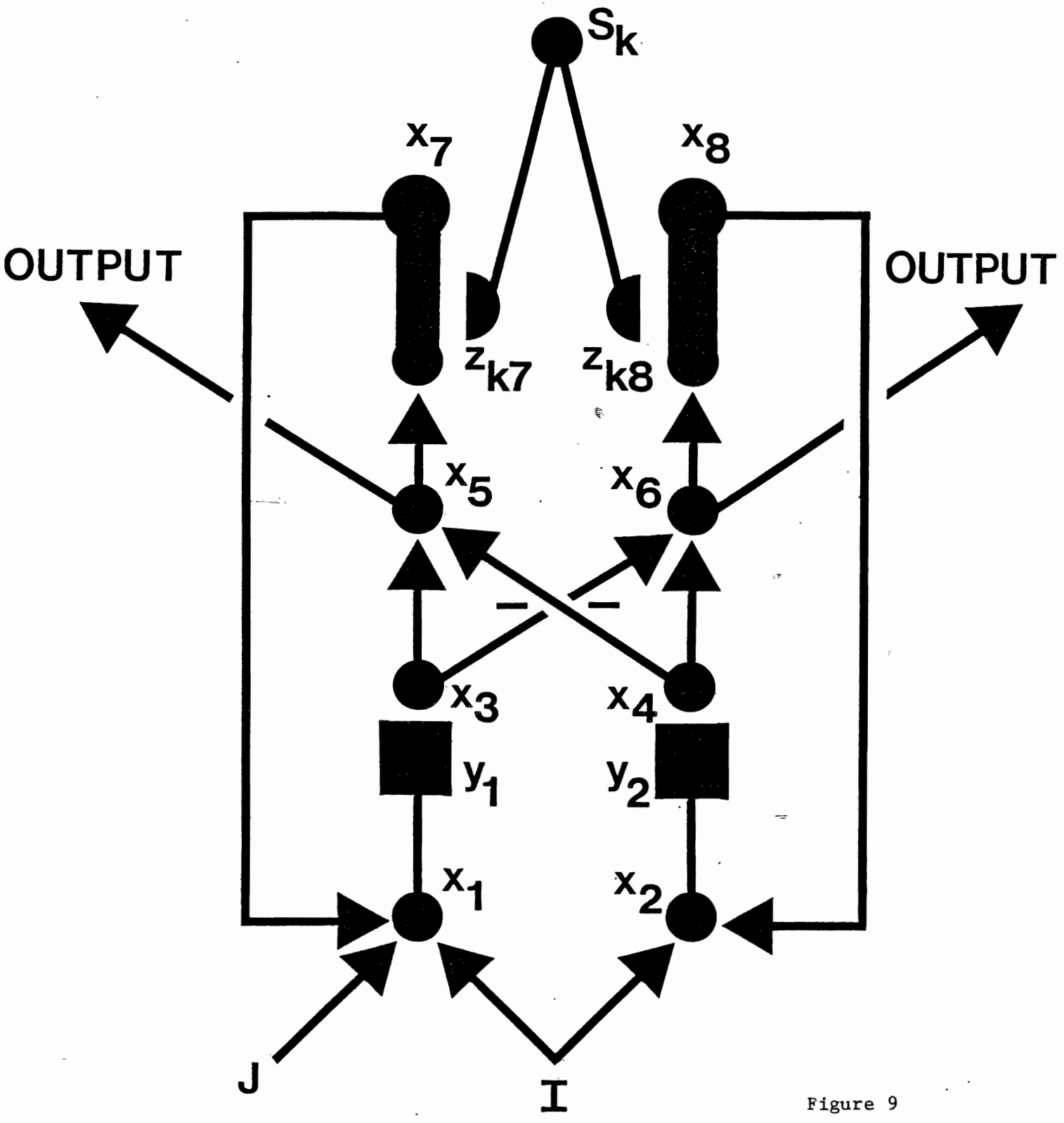

Figure 9

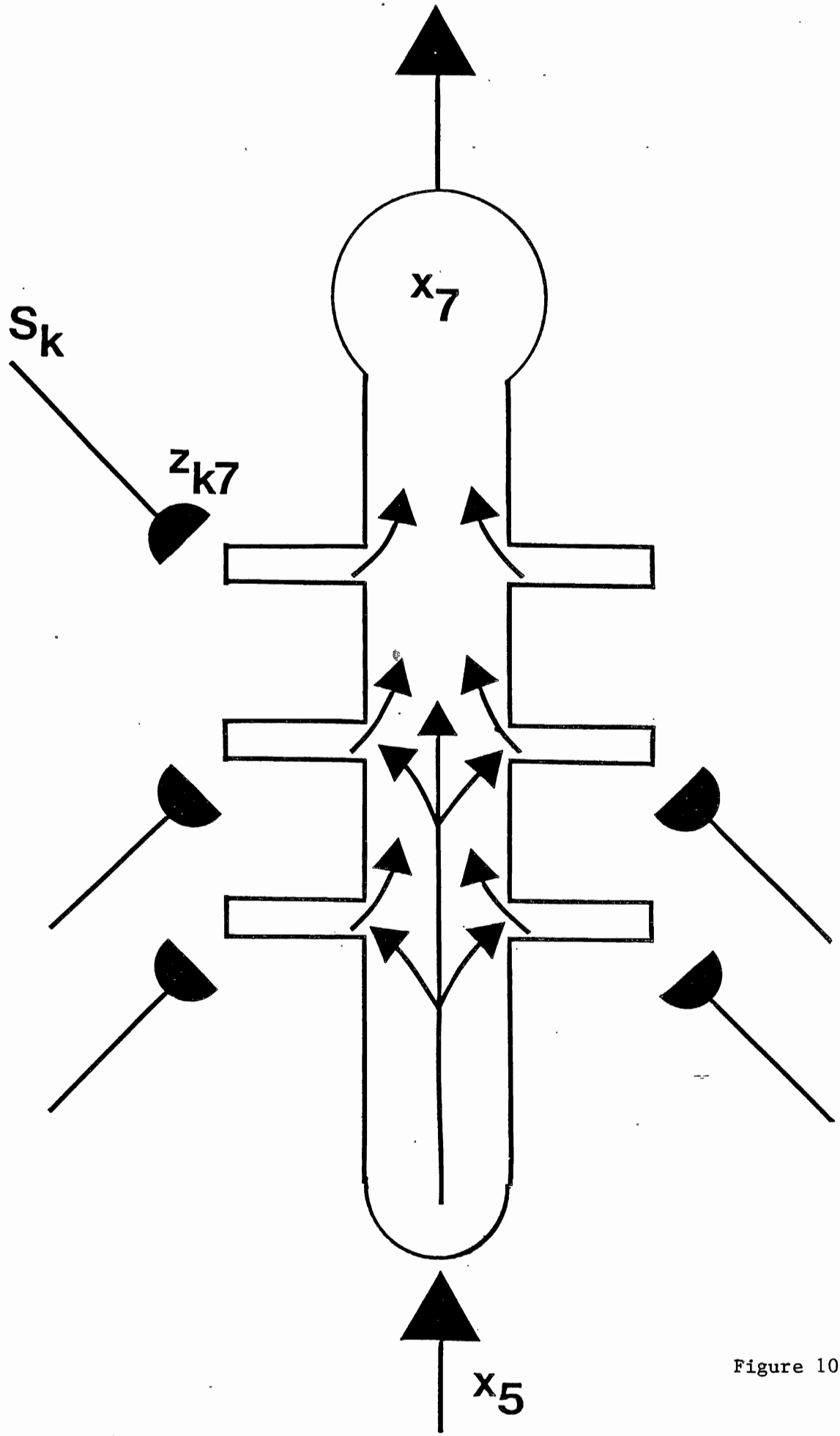

Figure 10

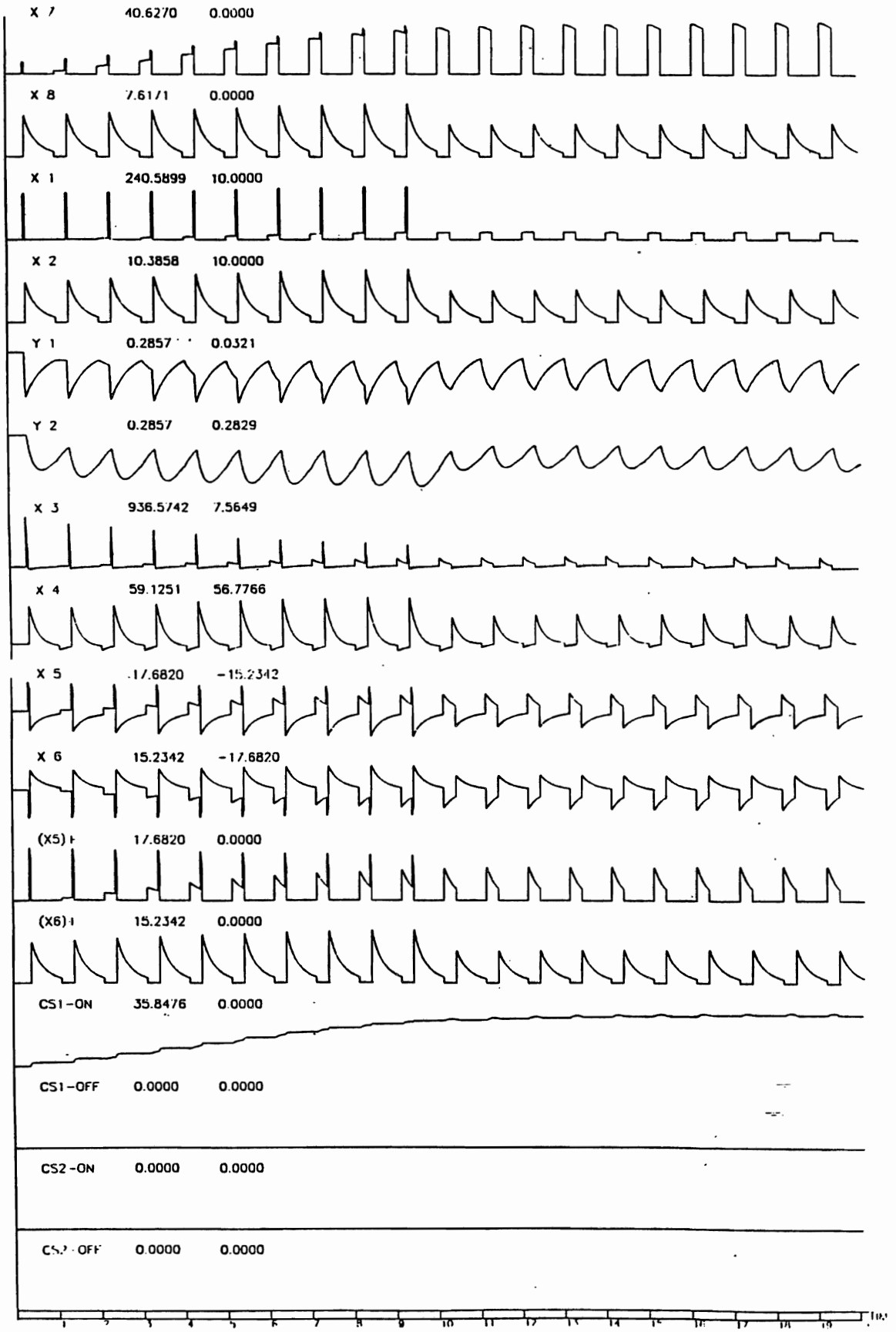
X 7 40.6270 0.0000
X 8 7.6171 0.0000
X 1 240.5899 10.0000
X 2 10.3858 10.0000
Y 1 0.2857 0.0321
Y 2 0.2857 0.2829
X 3 936.5742 7.5649
X 4 59.1251 56.7766
X 5 17.6820 -15.2342
X 6 15.2342 -17.6820
(X5)H 17.6820 0.0000
(X6)H 15.2342 0.0000
CS1-ON 35.8476 0.0000
CS1-OFF 0.0000 0.0000
CS2-ON 0.0000 0.0000
CS2-OFF 0.0000 0.0000

Figure 11

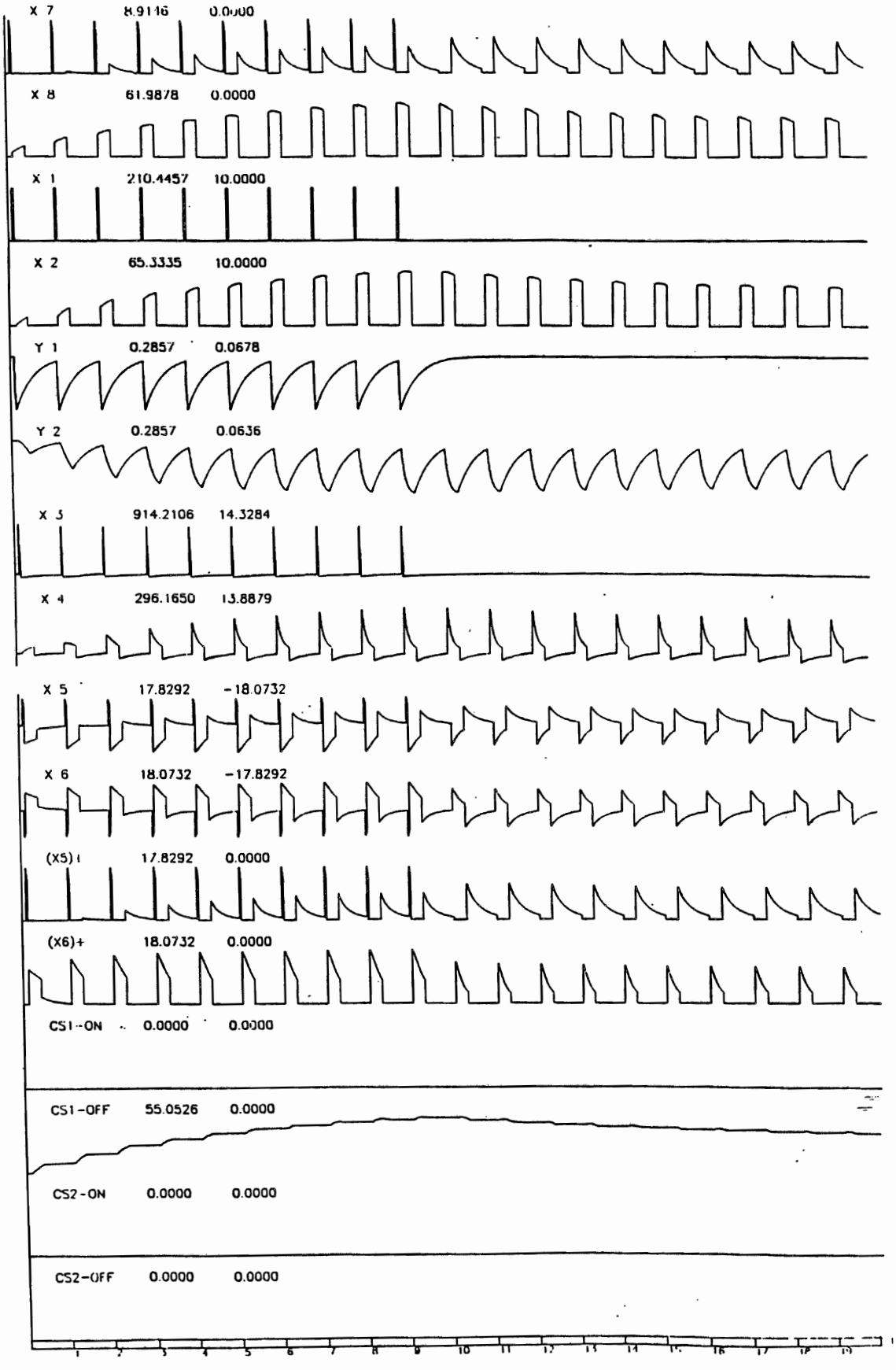
X 7 8.9146 0.0000
X 8 61.9878 0.0000
X 1 210.4457 10.0000
X 2 65.3335 10.0000
Y 1 0.2857 0.0678
Y 2 0.2857 0.0636
X 3 914.2106 14.3284
X 4 296.1650 13.8879
X 5 17.8292 -18.0732
X 6 18.0732 -17.8292
(X5)+ 17.8292 0.0000
(X6)+ 18.0732 0.0000
CS1-ON 0.0000 0.0000
CS1-OFF 55.0526 0.0000
CS2-ON 0.0000 0.0000
CS2-OFF 0.0000 0.0000

Figure 12

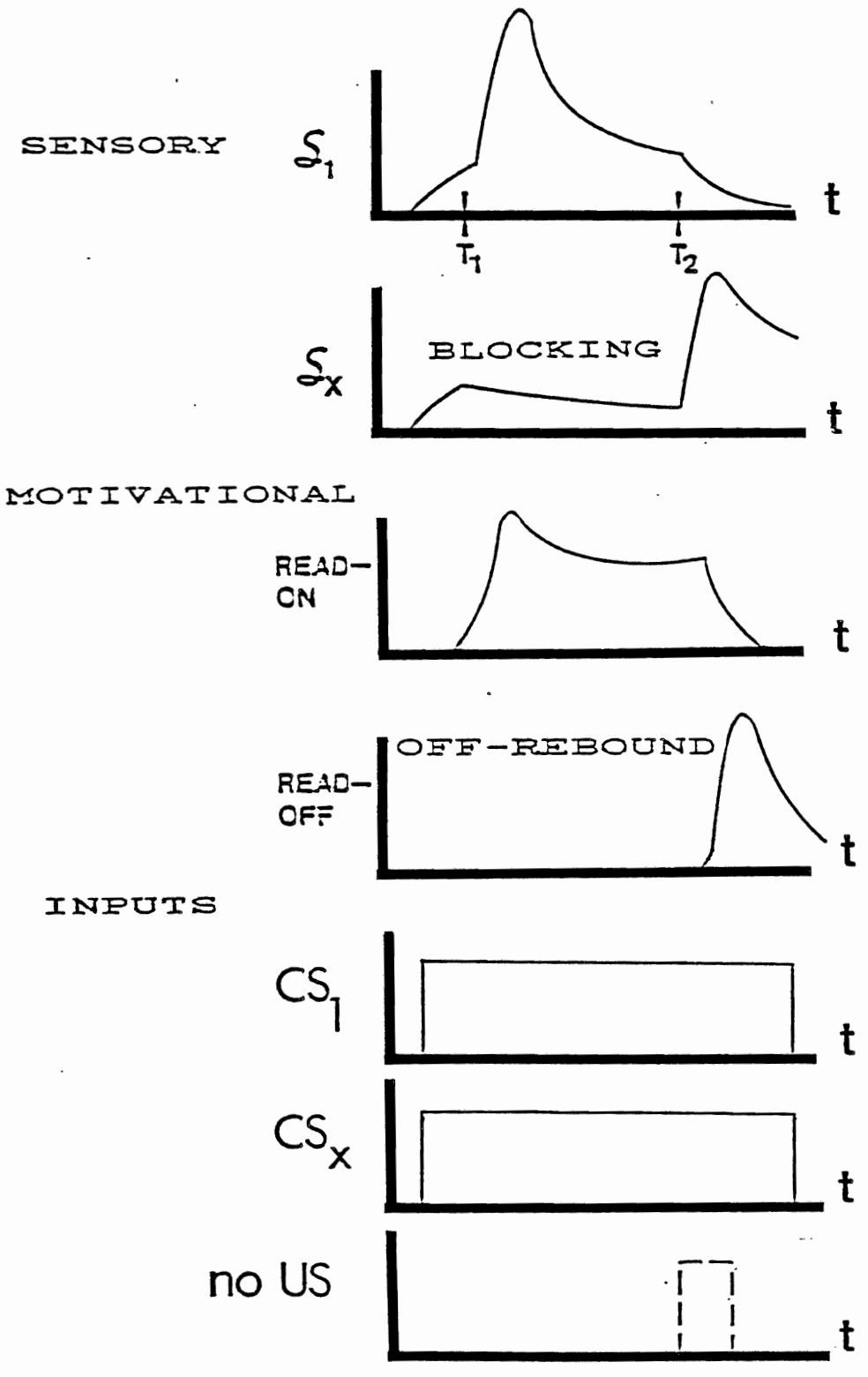
SENSORY
S_1
t
T_1
T_2
S_x
BLOCKING
t
MOTIVATIONAL
READ-ON
t
READ-OFF
OFF-REBOUND
t
INPUTS
CS_1
t
CS_x
t
no US
t

Figure 13