NETWORK MODELS OF HABITUATION

James C. Stanley

COINS Technical Report 76-13

(August 1976)

A Dissertation Presented

By

JAMES CARL STANLEY

DEPT. OF COMPUTER & INFORMATION SCIENCE
GRADUATE RESEARCH CENTER
UNIVERSITY OF MASSACHUSETTS
AMHERST, MASSACHUSETTS 01002

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 1976

Computer and Information Science

This work was supported in part by NIH Grant No. 5 R01 NS09755-03 COM awarded to Dr. M.A. Arbib, Department of Computer and Information Science, University of Massachusetts, Amherst, Massachusetts.

NETWORK MODELS OF HABITUATION

A Dissertation Presented

By

James Carl Stanley

Approved as to style and content by:

William L. Kilmer, Chairperson of Committee

M. A. Arbib, Member

D. N. Spinelli, Member

C. E. Hutchinson, Member

R. M. Graham, Department Chairman
Computer and Information Science

ACKNOWLEDGEMENTS

I wish to express my thanks to my advisor, Dr. W.L. Kilmer, for many hours of fascinating and fruitful discussions and for his guidance and critical evaluation of my work. I am grateful to Dr. M.A. Arbib for editorial comments on my dissertation, and for his efforts in creating an atmosphere in which such work could be done. I also thank the other members of my committee, Dr. D.N. Spinelli and Dr. C.E. Hutchinson for their help during the course of my research.

Special thanks go to Dr. R.F. Thompson for his continued interest in and support of my work, and to Dr, S.- I. Amari for his help with the mathematics of the population model of habituation presented here.

Most grateful thanks go to Dr. Fanya Montalvo for her constant faith in me and encouragement of my efforts. I also thank my parents for their continued and loving support. To these three people, I dedicate this thesis, without whom it would not have been possible.

Finally, I thank Donna Jerkovich and Janet Turnbull for their patient efforts in typing this manuscript.

This work was supported in part by NIH Grant No. 5 RO1 NS09755-03 COM awarded to Dr. M.A. Arbib, Department of Computer and Information Science, University of Massachusetts, Amherst, Massachusetts.

ABSTRACT

Network Models of Habituation

(September 1976)

James C. Stanley, B.A., Purdue University
M.S., Stanford University, Ph.D., University of Massachusetts

Directed by: Professor William L. Kilmer

How nervous tissues mechanize behavioral plasticities in animals is for the most part unkown. A number of theoretical models of plasticity have been presented that seek to explore the link between neural activity and behavior. In order to be relevant, such studies require a detailed specification of the behavioral aspects of the plasticity to be modeled, together with constraints on the possible circuit and synaptic effects that underlie the observed plasticity.

Habituation is a well-defined form of plasticity that has been studied at both the physiological and behavioral levels in a wide variety of organisms. Habituation is defined as a reversible decrement of response due to repeated stimulation. A more specific operational definition consisting of nine detailed characteristics of habituation is due to Thompson and Spencer. The wealth of data concerning habituation and its continued interest to experimenters make this form of plasticity an ideal subject for relevant theoretical modeling. In this dissertation, I present two sets of analytical models dealing with networks that exhibit habituation using simple synaptic modification rules.

A number of general theories of habituation exist that can serve as frameworks for the construction of circuit-level models. The most widely-accepted of these theories are the two-process theory (TPT) of Groves and Thompson and Sokolov's stimulus model theory (SMT). The SMT deals with complicated responses of complex systems, and has not been developed in great enough detail to allow relevant circuit-level modeling to proceed. The TPT has been expressed from its inception in terms of the actions of cells and synapses, and provides a sound basis for mathematical and computer modeling. Thompson and co-workers have elaborated the TPT to include explanations of a variety of the characteristics of habituation. Analytical modeling enables these explanations to be checked in detail.

In the first part of the dissertation, I describe a circuit model based on a circuit put forth by Groves and Thompson. In my model, synaptic decrement and facilitation governed by simple differential equations driven by presynaptic activity realize the processes of central habituation and sensitization. I show how this system can account for the affects on habituation of stimulus intensity and frequency. Generalization of habituation and the long-term and below-zero effects are dealt with through simple extensions of this model. The model is further elaborated in a statistical treatment of the input-output characteristics of a population of cells with thresholds. It is shown how two forms of experimentally observed stimulus intensity effect can arise in such a system. Together, these results give detailed support for the TPT and supply the experimenter with tools that will allow him to investigate par-

ticular forms of synaptic plasticity that might underlie habituation.

One characteristic of habituation whose underlying mechanism has not been dealt with in detail in the TPT is temporal conditioning. The TPT and the SMT deal in somewhat similar terms with this characteristic. Each theory posits that an organism constructs a model of the temporal characteristics of a regularly repeated stimulus. The model is then used to sensitize the organism to the stimulus or to alterations of the stimulus. Plausible circuit models of this phenomenon have not been offered by proponents of either theory of habituation.

The work of Vinogradova and her colleagues suggests that the hippocampus is a brain region in which such temporal modeling may occur. The hippocampus has been implicated in learning and memory, and has been shown to display plasticity at the levels of unit activity and synaptic operation. Accordingly, in the second part of this dissertation, I present a temporal sequence memory network based on the known structure of the dentate gyrus portion of the hippocampal complex. This model network is so designed that once it has been sufficiently exposed to a regularly repeated sequence of inputs, it can repeat that sequence with the proper timing when cued by an initial portion of the sequence. The network employs traveling waves of activity to generate a dynamic representation of the stimulus that codes both the set of cells excited by the input and the interval that has elapsed since the stimulation. Representations are linked by a process of association to allow recall to

occur. Simulation studies are presented which elucidate the network's storage and recall capabilities. This network can be used as part of an habituation system exhibiting temporal conditioning of the form required by either the SMT or the TPT.

TABLE OF CONTENTS

TABLE OF CONTENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

LIST OF FIGURES

CHAPTER I

HABITUATION: MODELS AND THEORIES

Introduction

One of the most intriguing aspects of animal behavior is its plasticity. Organisms must constantly adapt themselves to a changing environment in order to survive. Behavioral plasticity is studied by a variety of disciplines and at a number of levels. Some psychologists study overt behavior in order to characterize learning and other forms of plasticity at phenomenological levels. Psychologists and physiologists also deal with the neural substrates of behavior in order to identify changes in the nervous system that lead to changes of response.

The notion that mathematics can be applied to the biological sciences has led to a host of theoretical investigations into plasticity that seek to complement and extend the experimental studies. Some theoretical models deal with association learning and habituation with little regard for direct relevance to biological studies (Amari, 1972; Andersen, 1972; Kohonen, 1974; Nilsson, 1965). These models explore the possibilities inherent in particular circuit anatomies or in given forms of plasticity. They also serve to extend the repertoire of mathematical and computer tools available to the modeler. The extension of modeling methods is particularly important, since interesting models often involve nonlinear or probabilistic elements that are difficult to analyse. Studies of the sort mentioned above can help to develop concepts to guide

modeling at a level closer to experimental data. More concrete models are aimed at an understanding of a body of data pertaining to a particular system, phenomenon, or mechanism (Kilmer and Olinski, 1974; Marr, 1970, 1971; Wigström, 1974; Uttley, 1975; Hebb, 1949).

In some of the studies mentioned above, intuitive arguments are given for the ways the models function. Other models, however, are described using mathematical analysis or computer simulation. The analytical approach to theories and models of physiological systems has several advantages. Complex models are difficult to understand on an intuitive basis, and even in simple models, factors can combine in unexpected ways to yield unforseen effects. Mathematics and computer simulations act in these cases as active bookkeeping devices that take all factors properly into account. The discipline necessarily involved in turning the intuitive beginnings of a theory into a rigorous mathematical system or a computer program can also uncover hidden assumptions and inconsistencies in the theory. For example, it may be found that no range of parameters exists within which the model operates as desired, forcing it to be revised or rejected.

Analytical modeling can make three main contributions to theories and models that do what intuition says they should. First, ranges of parameters can be found within which a model behaves properly. Knowledge of these parameter ranges has a number of consequences for verification of the model. If the critical para-

meter ranges are too narrow, the model might be called into question as being unrealistic. Model parameter values can be compared with experimentally determined values and the model can be rejected if the measured values fall outside the model's workable range. Analysis might also predict that unmeasured parameters must fall within particular ranges. The values of those parameters can then be sought experimentally to verify the model.

Secondly, analytical models allow direct, quantitative comparison of different explanations of a given phenomenon. Models that are found to require different parameter values can be differentiated experimentally on that account. Finally, experiments can easily be performed on a model in order to generate predictions that support or refute the theory upon which the model is based.

In order to be relevant to a course of experimentation, analytical modeling must be guided and constrained by experimental data. Models less tied to experimental results can show what behaviors are possible in given anatomies employing particular plastic mechanisms, but such models may be of use primarily in extending the available catalog of modeling techniques, as discussed above. The work of modeling begins with the casting of circuit operation and the characteristics of circuit elements into forms amenable to mathematical analysis and computer simulation. The forms chosen must reflect the available data and the kinds of measurements that can reasonably be made on the system being modeled. The use of imaginary or inaccessable variables may lead to an elegant theory, but its elegance is likely never to be tarnished by experimental validation. As discussed above, the operation of a relevant analytical model should be governed by parameters that allow it to be tested against other models and against the physiological system in question.

A useful model of plasticity therefore requires data on a number of levels. First, a detailed description of the characteristics of the observed plasticity is necessary to supply criteria against which to judge a model. Second, reasonable constraints must be given for the circuit anatomies and plastic mechanisms that may underlie the observed changes of response. Within these constraints, the modeler must develop a parsimonious mathematical or computer description of the desired phenomenon. This description should tell how various characteristics of the modeled plasticity arise and how they depend on the characteristics of the underlying circuits and circuit elements. As above, the model should be testable in terms of predictions concerning parameter ranges and untried experiments.

One form of plasticity that has been widely studied at both the behavioral and the physiological levels is habituation. Habituation is defined as a reversible decrement of response due to repeated stimulation (Horn, 1970). The many organisms and preparations that display habituation have made this phenomenon especially interesting to workers in physiology and psychology. The detailed definition of habituation on the behavioral level and the wealth of data concerning its physiological underpinnings give the modeler

the guidelines and constraints necessary for the construction of relevant models. These models can be used to guide further experiments into the nature of habituation and can also illustrate and extend the current repertoire of modeling techniques.

In what follows, I present two classes of network models of habituation. In each case it is seen how simple synaptic modification rules, together with reasonable circuit anatomies, can be made to realize some of the detailed characteristics of habituation. Analytical and computer investigations of these models lead to explanations of experimental results and to testable predictions for further experiments. In the rest of this chapter, I first discuss data concerning the physiology of habituation, then discuss two major habituation theories. The first of these theories, the two-process theory of Groves and Thompson (1970), forms the basis of network models presented in Chapters III and IV. In Chapter V, I present a memory model that may be used within the framework of either general theory to realize the characteristic of temporal conditioning.

The Physiology of Habituation

Habituation to more or less natural stimulation is displayed by a wide variety of organisms. The earthworm exhibits a defense response to vibration that involves contracting its body and hooking its tail. Upon repeated stimulation, the body contraction rapidly wanes, while the tail hook disappears more slowly (Gardner,

1968, cited in Wyers, et. al., 1973). In the aquatic snail Aplysia, a gill withdrawal response may be elicited by stimulation of the gill mantle or siphon with jets of sea water (Kandel, Castellucci, Pinsker and Kupfermann, 1970). Upon repeated stimulation, the gill is withdrawn less and remains withdrawn for shorter periods of time. The spider Uloborus responds to stimulation of its web with a vibrating needle by turning toward the vibrating strand, then running along it to the point of stimulation (Szlep, 1964). If on each trial the needle is removed before the spider reaches it, repeated trials lead first to a cessation of the running, then to cessation of the turning. Human beings, when exposed to low or moderate intensity stimulation display an orienting response described by Sokolov (1961) as a complex of changes of the autonomic nervous system directed toward heightened perception. After a few presentations of a nonsignificant stimulus, many components of this response wane and drop out.

While organisms from earthworm to man show response decrements to repeated stimulation, not every decrement is an instance of habituation. Habituation must be differentiated from sensory and motor fatigue or accommodation. An habituated response may be elicited by a stimulus sufficiently different from the habituating stimulus, ruling out effector fatigue as the basis of the decrement. Brief presentation of a novel stimulus may result in renewed response to an habituated stimulus, ruling out sensory fatigue as the cause of habituation. This effect is known as dishabituation, and goes hand in hand with habituation itself. Dishabitua-

tion has been demonstrated in all of the organisms referred to above. The specificity of habituation to the habituating stimulus and to similar stimuli, and the presence of dishabituation, serve to differentiate between true habituation and fatigue or accommodation effects.

Habituation is of obvious use to the behaving organism. In shutting out the effects of repeated, nonsignificant stimuli, the organism is more able to cope with important stimuli. At the same time, the ability to dishabituate allows the organism to respond to a hitherto insignificant stimulus suddenly made significant by a change in the stimulus itself or in the stimulus context.

Given the variety of organisms that display habituation, it should be no surprise that the characteristics of habituation have been studied in detail. These studies were brought together by Thompson and Spencer (1966) in an operational definition of habituation consisting of the following nine parts:

1. Given that a stimulus elicits a response, repeated stimulation lends to decreased response.

This characteristic is naturally the backbone of the definition of habituation (though the definition applies to invertebrates, as well). Figure 1.1 illustrates response decrement in the case of the flexion reflex of the spinal cat (Thompson and Spencer, 1966). As stimulation proceeds, the response decreases with a negative ex-

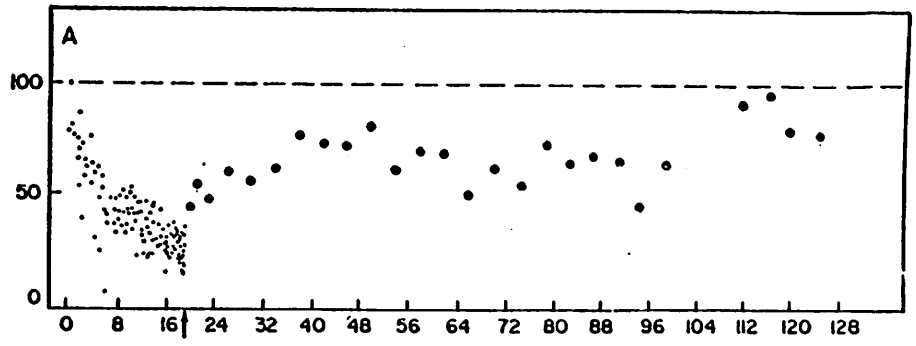


Figure 1.1 Response decrement of flexion reflex of spinal cat (Thompson and Spencer, 1966). Abscissa is marked in terms of percentage of control response, ordinate in minutes. At arrow, habituation stimuli ceased and recovery test stimuli began. Reprinted with the permission of the author and publisher.

ponential form. Note that the response does not fall smoothly, but is quite variable. In order to smooth the habituation curves and make the decrement more apparent, responses are often averaged over several trials and curves are plotted giving the average in each block of trials.

The response in Figure 1.1 is measured in terms of percentage of a control response taken before habituation training is started. For inputs spaced at great-enough intervals the response remains constant, and thus can serve as a level with which the response during habituation training can be compared. In using a relative response measure, individual vaiations in response are factored out, facilitating comparisons of data taken from different preparations. The absolute decrement of response is also used to measure habituation. As is discussed in point 6 below, these two measures of habituation behave differently with respect to stimulus intensity.

2. Presentation of a novel stimulus results in recovery of the response.

This is the phenomenon of dishabituation. An example of dishabituation of the cat flexion response is shown in Figure 1.2, taken from the work of Spencer, Thompson and Neilson (1966). Following habituation with a low intensity stimulus, a brief high intensity stimulus is presented. The response to the low intensity stimulus is then retested and found to be increased for a short

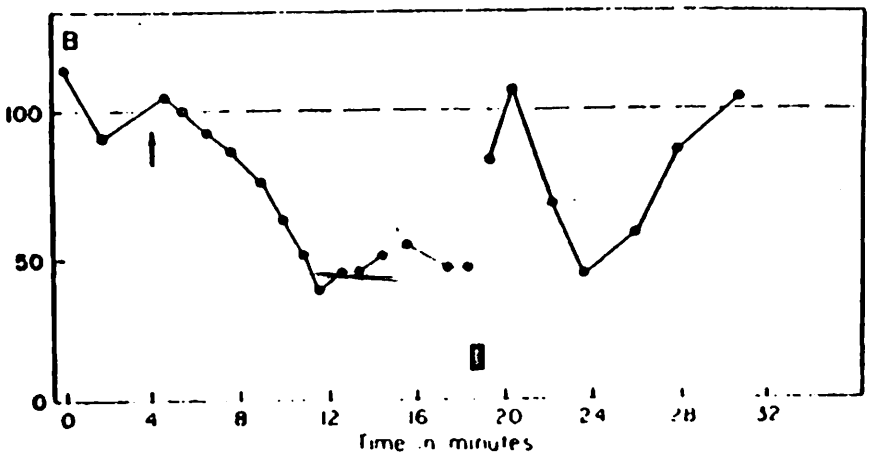


Figure 1.2 Dishabituation of flexion reflex of spinal cat (Thompson and Spencer, 1966). Abscissa and ordinate as in Figure 1.1. Dishabituating stimulus given at point of black bar. Test stimuli follow that. Reprinted with permission of the author and publisher. Copyright 1966 by the American Psychological Association.

time. Another example is found in the work of Rowell (1970) on the locust visual system. After habituation to a visual stimulus, the response of a unit in the tritocerebrum is reestablished after a brief shock applied to an afferent nerve. A dishabituating stimulus need not be strong or aversive, as will be discussed later in conjunction with Sokolov's theory of habituation.

3. If the stimulus is withheld, the response recovers with time.

Response recovery is tested with application of the habituating stimulus after rest intervals of various lengths following habituation to asymptote. At longer test intervals the response is closer to the control level. Recovery of the gill withdrawal response of Aplysia is illustrated in Figure 1.3, taken from the work of Kandel and his associates (Pinsker, Kupfermann, Castellucci and Kandel, 1970). As shown there, recovery proceeds exponentially with time from the point that the habituation training is stopped. Note that in this case, recovery appears to proceed rapidly for a short time, then more slowly. A further characteristic of recovery is illustrated in point 5 below.

4. The more rapid the stimulus frequency, the more rapid or pronounced the decrement.

This effect is illustrated in Figure 1.4 taken from the response of the ventral root of the isolated frog spinal cord to lateral column stimulation (Farel, et. al., 1973). For the same num-

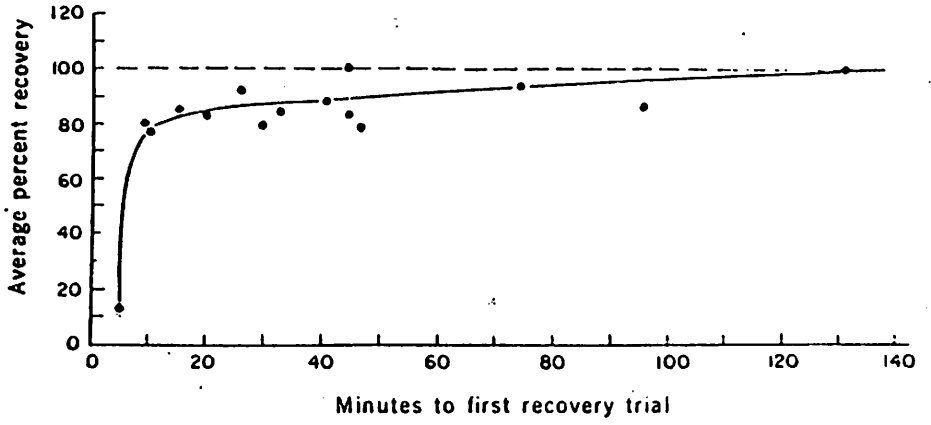


Figure 1.3 Recovery of habituation of the gill withdrawal response of Aplysia (Pinsker, et. al., 1970). Reprinted with permission of the author and publisher. Copyright 1970 by the American Association for the Advancement of Science.

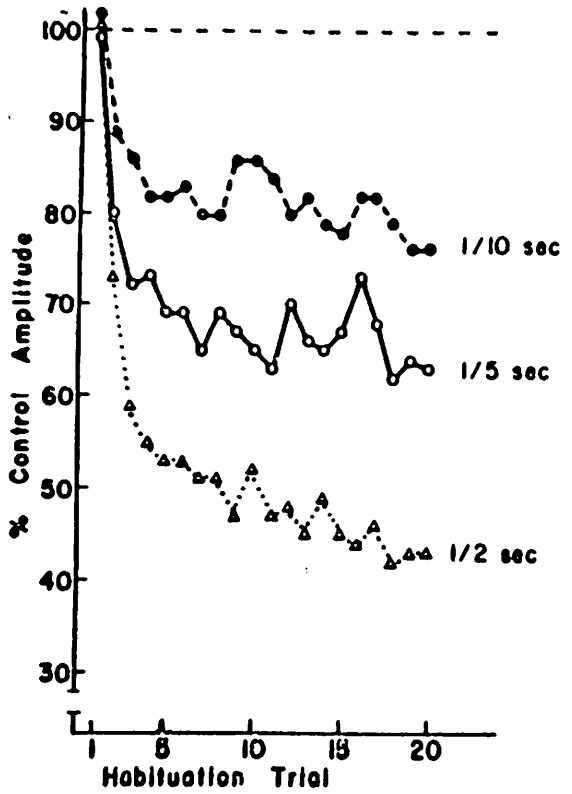


Figure 1.4 Effect of stimulus frequency on habituation (Farel, et. al., 1973). Reprinted with permission of the author and publisher.

ber of stimuli, as stimulus frequency increases so does the relative response decrement.

5. The effect of habituation may proceed beyond the asymptotic response level.

If enough habituating stimuli are given, the response generally falls to an approximately steady asymptotic level. The asymptote may be zero, or some non-zero level. In some systems, continued stimulation leads to a slowed recovery. This effect is illustrated in Figure 1.5, taken from the work of Farel and coworkers (1973) on the ventral root response to electrical stimulation of the lateral column of the isolated frog spinal cord.

6. Weaker stimuli produce greater habituation; strong stimuli may produce none at all.

Figure 1.6 taken from Thompson and Spencer's (1966) work on the flexion reflex of the spinal cat illustrates this effect. As stimulus intensity increases, relative habituation decreases. Farel and coworkers (1973) demonstrate that the lateral column to ventral root response of spinal frogs also displays this effect. In the dorsal root to ventral root response of the frog spinal cord, the absolute response decrement increases with increasing stimulus intensity (Farel and Thompson, 1972). The statement of this habituation characteristic, then, refers to relative habituation, rather than to absolute habituation.

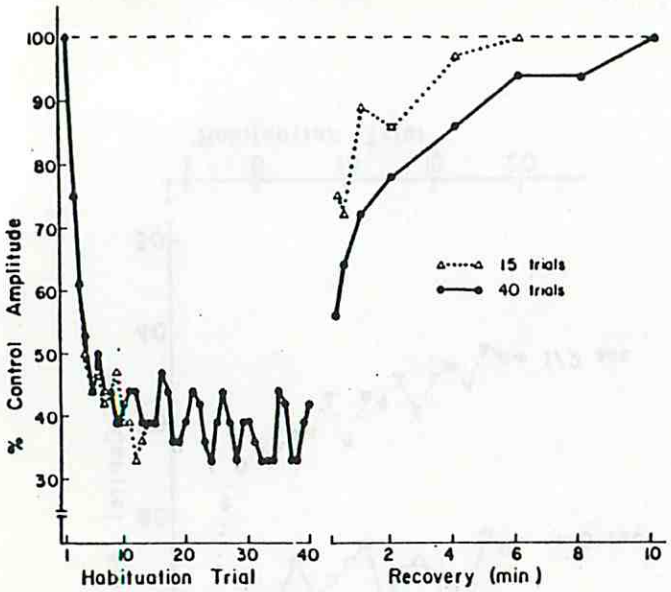


Figure 1.5 Illustration of the below-zero effect. Response recovery is slowed following prolonged habituation (Farel, et. al., 1973). Reprinted with permission of the author and publisher.

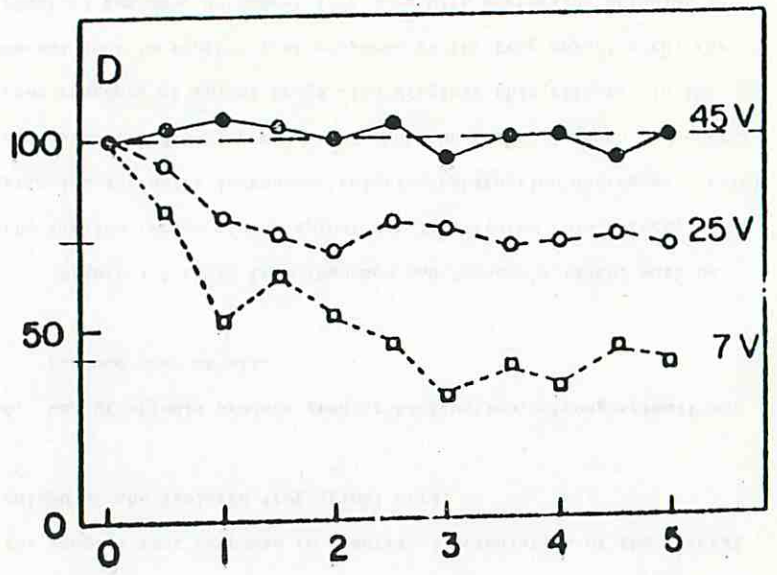


Figure 1.6 The intensity effect in habituation of the cat flexion reflex (Thompson and Spencer, 1966). Reprinted with permission of the author and publisher. Copyright 1966 by the American Psychological Association.

7. Repeated series of habituation lead to more rapid or pronounced habituation.

This is called the long-term effect. As illustrated in Figure 1.7, the gill withdrawal response of Aplysia becomes more pronounced with each successive series of habituation trials followed by a period of recovery (Carew, Pinsker and Kandel, 1972). This effect in Aplysia can last up to two weeks. A similar long-term effect has been reported in the chronic spinal frog (Farel, 1971), and in the orienting response of cats (Sharpless and Jasper, 1956).

8. Dishabituation habituates upon repeated elicitation.

Figure 1.8, taken from the work of Farel and Thompson (1972) on the dorsal root to ventral root response of the isolated frog spinal cord shows that the peak response reached upon dishabituation with a shock becomes smaller with successive shocks. This may be thought of as a long-term habituation effect, as the period between dishabituation trials here is greater than that between habituation trials.

9. Habituation due to one stimulus exhibits generalization to other stimuli.

In testing habituation generalization, the response to one stimulus is checked before and after habituation with a different stimulus. If the response to the first stimulus is decreased fol-

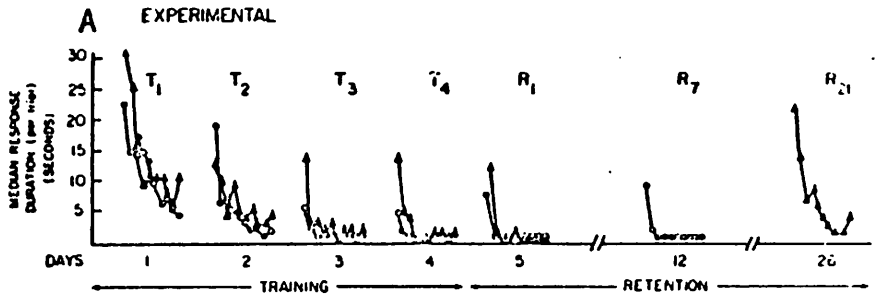


Figure 1.7 The long-term effect, illustrated by the gill withdrawal reflex of Aplysia (Carew, et. al., 1972). Habituation is greater with each successive training session. Reprinted with permission of the author and publisher. Copyright 1972 by the American Association for the Advancement of Science.

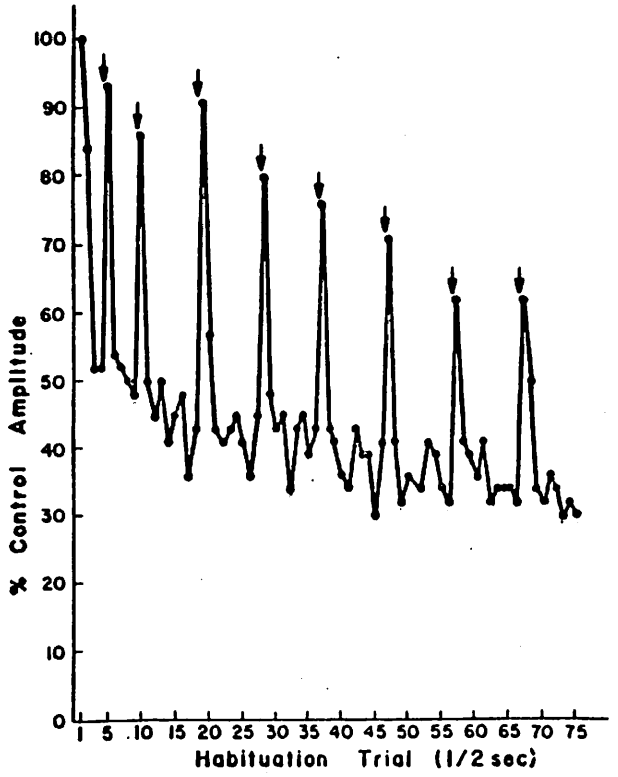


Figure 1.8 Habituation of dishabituation in the frog spinal cord (Farel and Thompson, 1972). Dishabituating shocks applied at arrows. Effect of successive shocks decreases. Reprinted with permission of the author and publisher.

lowing habituation to the second, then habituation is said to have generalized from the second to the first. An example of this effect is provided in Figure 1.9, taken from Thompson and Spencer's (1966) work on the flexion reflex of the cat. Two different branches of the same afferent nerve were used to elicit the response. Habituating stimuli (that is, stimuli given at a rate great enough to produce habituation) were given to one branch and test stimuli were presented at a lower rate to the other branch of the nerve. The result shown in Figure 1.9 indicates that habituation transferred to the test branch as stimulation proceeded.

Petrinovich (1973) has suggested the addition of a tenth commonly-seen characteristic of habituation to the list of nine basic features given above. This characteristic is due to Hinde (1954):

10. Habituation proceeds more rapidly with spaced rather than massed trials.

Working with chaffinches, Hinde (1954) found that the mobbing response to owl models habituated more to short presentations on successive days than to a single long presentation. A similar result was found in the gill withdrawal response of Aplysia (Carew, Pinsker and Kandel, 1972). Note that the conditions under which this effect is tested (widely-spaced habituation series with recovery between series) are the same as those under which the long-term effects are studied. This effect, then, should probably be considered in conjunction with characteristic 7 above.

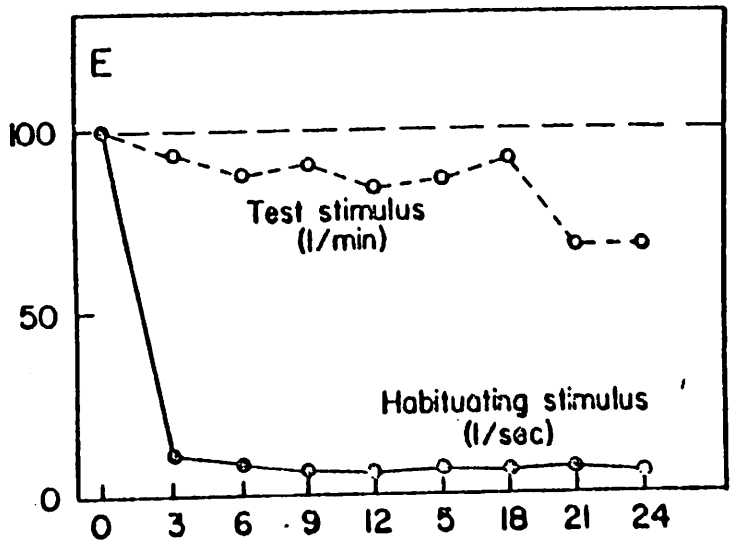


Figure 1.9 Generalization of habituation in the cat spinal cord (Thompson and Spencer, 1966). Test and habituating stimuli given to two different branches of an afferent nerve. Response to test decreases during habituation of other input. Reprinted with permission of the author and publisher. Copyright 1966 by the American Psychological Association.

This operational definition of habituation has been widely but not universally accepted. In particular, Hinde (1970) has questioned it on the basis that these characteristics are not detailed enough to be of use or general enough to cover the whole range of behavioral and physiological experiments. On the other hand, Thompson and others have demonstrated the applicability of this definition in a wide variety of experiments. It is probably best, then, to consider these ten characteristics as guidelines to the phenomenon of habituation that may be used in the study of a given preparation. A preparation that displays a number of these characteristics may be said to habituate, and may then oe used to verify or modify the operational definition of habituation itself.

A number of preparations have been studied at a level that reveals something of the neurophysiological mechanisms that underlie habituation. One of the best-studied examples of habituation is the gill withdrawal reflex of the marine snail Aplysia (Pinsker, Kupfermann, Castellucci and Kandel, 1970; Kupfermann, Castellucci, Pinsker and Kandel, 1970; Castellucci, Pinsker, Kupfermann and Kandel, 1970; Epstein and Tauc, 1970; Jacklet and Lukowiak, 1975; Carew, Pinsker and Kandel, 1972). When a jet of seawater is used to stimulate the Aplysia gill mantle or the siphon, gill and siphon are withdrawn for protection. Upon repeated stimulation, the gill is withdrawn less and for a shorter period of time, as illustrated in Figure 1.10. This decrement of response exhibits eight of the ten characteristics of habituation, lacking only the below-zero effect and generalization to different points of stimulation. Dis-

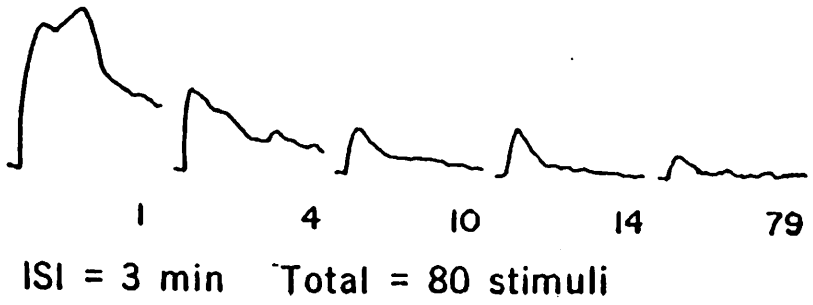


Figure 1.10 Response decrement in the gill withdrawal response of Aplysia (Pinsker, et. al., 1970). Height of each curve gives size of response. Number of stimulus giving rise to each curve is indicated beneath the curve. Reprinted with permission of the author and publisher. Copyright 1970 by the American Association for the Advancement of Science.

habituation of the response takes place following strong stimulation of the animal's head, rather than stimulation of the gill mantle or siphon.

The circuitry serving the gill withdrawal response in Aplysia has been worked out in detail by Kandel and his co-workers (Kupfermann, et. al., 1970). Figure 1.11 illustrates their findings. The muscles of the gill and mantle are contracted by the actions of a number of motoneurons in the animal's abdominal ganglion. These neurons are in turn activated by two tactile pathways. In one pathway, the tactile sensory cells make direct contact with the motoneurons. An interneuron is interposed between the sensory cell and the motoneuron in the second pathway. There are other (nontactile) inputs to these motoneurons, but those inputs do not show habituation. Weak stimuli activate the direct path to the motoneurons, while stronger stimuli result in the firing of the interneurons as well.

In a series of experiments on the direct pathway, Castellucci and his colleagues (Castellucci, et. al., 1970) demonstrated that habituation is a result of a decrement of the EPSP produced by the synapse connecting the sensory cell with the motoneuron. Effector fatigue and sensory adaptation were excluded as the basis of the response decrement. Owing to the presence of other nonhabituating synapses on the motoneurons, changes in the overall excitability of those cells could not be the cause of habituation. Presynaptic inhibition due to activation of nearby afferent fibers was ruled out

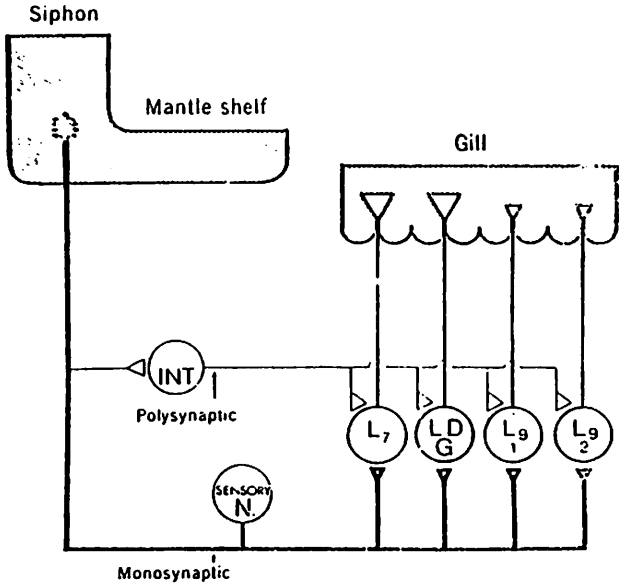


Figure 1.11 Circuitry mediating gill withdrawal in Aplysia (Kupfermann, et. al., 1970). Sensory neuron activated by tactile stimulation of the siphon makes synapses with motoneurons and with interneuron. Motoneuron action withdraws the gill and siphon. Reprinted with permission of the author and publisher. Copyright 1970 by the American Association for the Advancement of Science.

through direct electrical stimulation of a single afferent. These results suggest that habituation of the gill withdrawal response of Aplysia is due either to a decrement in the amount of transmitter released by the synapses connecting sensory cells with motoneurons, or to a decrease in the sensitivity of the postsynaptic membrane. The results of Castellucci and his colleagues (Castellucci, et. al., 1970) show that the EPSP evoked on the motoneurons by tactile stimulation continues to decline following failure of the motoneuron spike, suggesting that the decrement is due primarily to activity in the presynaptic fiber.

Studies of dishabituation of this response (Castellucci, et. al., 1970) indicate that it is a form of sensitization; stimulation of the nerves connecting the cerebral ganglion with the abdominal ganglion leads to an increase in EPSP's generated by both habituated and nonhabituated synapses. Several authors have suggested that this sensitization is a form of heterosynaptic facilitation (Castellucci, et. al., 1970; Kandel and Tauc, 1965, a, b; Epstein and Tauc, 1970). Epstein and Tauc (1970) have demonstrated that facilitation functions independently of activity in the cells that originate and receive the facilitated synapse.

Therefore, these results indicate that the depression of response seen upon repeated stimulation of the gill of Aplysia is most likely due to decreased transmission at the synapses connecting the tactile sensory cells with the motoneurons controlling the gill muscles. Dishabituation of the response may be a form of pre-

synaptic facilitation caused by activity in cells outside the direct stimulus-response link.

Another example of response decrement in a simple system is habituation of the crayfish tactile-defense reflex. The crayfish exhibits a tail-flip response to stimulation of abdominal sensory cells. The tail-flip acts as a defense mechanism that carries the animal away from the source of stimulation (Wine and Krasne, 1972). Electrophysiological and anatomical studies of this response (Zucker, 1972, a, b; Wine and Krasne, 1969; Krasne and Roberts, 1967) have shown that it is mediated by a pair of lateral giant fibers activated by the abdominal sensory cells through interneurons. The efficacy of the synapses connecting the sensory cells with the interneurons decreases with stimulation at rates as low as one per five minutes, resulting in decreased reflex response. Similarly, the electrical junctional potential generated in the abdominal muscles by the action of the giant fibers decreases with stimulation at rates of about one per minute (Bruner and Kennedy, 1970). Stimulation at an intermediate rate (1/sec) leads to an increase of the junction potential, while with still greater stimulus rates (5/sec), the potential increases then decreases. These effects appear to be due to changes in the presynaptic element only (Bruner and Kennedy, 1970).

Habituation due to depression of synaptic efficacy has been reported in the stellate ganglion of the squid (Horn and Wright, 1970), in the parietal ganglion of the land snail Helix (Pakula and

Sokolov, 1973), and in the abdominal ganglia of the cockroach (references quoted in Eisenstein and Peretz, 1973). Each of these preparations was studied by stimulating a nerve bundle or fiber and recording the resulting monosynaptic EPSP. Some controversy exists as to the mechanism of habituation in Helix. Holmgren and Frenk (1961) report a build-up of inhibition recorded intracellularly in neurons of the parietal ganglion concurrent with the development of the response decrement. Sokolov and Pakula (1973), however, point out that stimulus rates employed by Holmgren and Frenk were greater than those that cause behavioral habituation. This indicates that build-up of inhibition may be a factor in response decrement at rates great enough to cause addition of successive IPSP's, while synaptic decrement may be the main source of response decrement at lower stimulus rates.

Two habituation systems that are more complex than those mentioned above are the cat and frog spinal cords. Each of these systems exhibits responses with habituation characteristics suggestive of synaptic depression. The spinal cord of the bullfrog may be removed from the animal and kept in a physiological state for several hours (Brookhart and Fadiga, 1960; Machne, Fadiga and Brookhart, 1959). During this time the isolated cord is readily accessable for stimulation and recording. Two systems within the cord display habituation to electrical stimulation. The first system is a set of descending fibers called the lateral column that runs the length of the cord and makes monosynaptic contact with motoneurons in the ventral roots (Brookhart and Fadiga, 1960; Machne, et. al., 1959).

The response recorded in the ventral root to stimulation of the lateral column displays eight of the ten characteristics of habituation (Farel, Glanzman and Thompson, 1973). Farel and co-workers (1973) have indicated that the response decrement is due to decreased efficacy of single synapses.

The second habituating response in the isolated frog spinal cord is the electrically recorded response in the ventral root due to dorsal root stimulation. This system involves at least one interneuron (Machne, et. al., 1959), so is more complex than the lateral column system. The ventral root response to dorsal root stimulation displays eight of the ten habituation characteristics (Farel and Thompson, 1972), but has not been studied at a level of detail that would allow the mechanism of habituation to be proposed.

The flexion reflex of the cat is mediated by the animal's spinal cord. The reflex is studied in spinal cats--animals in which the midbrain is transected in order to remove descending influences --through mechanical stimulation of the skin or through direct electrical stimulation of a nerve entering the dorsal root (Thompson and Spencer, 1966; Wickelgren, 1967, a, b; Spencer, Thompson and Neilson, 1966). This reflex has been shown to display the nine basic characteristics of habituation (Thompson and Spencer, 1966). Thompson and his colleagues have indicated that habituation of the flexion reflex is due to decreased efficacy of synaptic transmission, and that dishabituation is the result of sensitization, or increased transmission, at other synapses. This notion will be discussed in

greater detail below. Other authors (Wickelgren, 1967, a, b; Wall, 1970) have suggested instead that build-up of inhibition is responsible for the decrement. Again, this system has not been studied in great enough detail to resolve the controversy.

In systems simple enough or accessible enough to be studied in detail on a neurophysiological level, then, habituation generally appears to be the result of homosynaptic depression, or self-generated depression (Horn, 1967). "Homo" here indicates that only activity in the stimulated pathway itself is altered by the stimulation; "hetero" indicates that activity in one pathway affects the conduction properties of others (Kandel, et. al., 1970). Circuits involving pre- or postsynaptic inhibition mediated by interneurons or afferent collaterals have been rendered less likely candidates for the mechanism of habituation in the simpler systems mentioned above (Pakula and Sokolov, 1973; Castellucci, et. al., 1970; Farel, et. al., 1973). Though such inhibitory effects certainly exist even in simple systems (Holmgren and Frenk, 1960; Tauc, 1965) and so will act to shape overall response, the notion of homosynaptic depression will guide the circuit modeling presented in later chapters.

Two Habituation Theories

A variety of theories exist that seek to explain habituation in general. Two theories that have wide appeal are Sokolov's model theory and Thompson's two-process theory. The model theory was de-

rived from experiments on the human orienting response (OR) (Sokolov, 1961, 1975). The OR is released by a wide range of stimuli and causes changes in such variables as heart rate, blood pressure, GSR, and EEG. These changes are believed to be produced by signals originating in the reticular formation. In Sokolov's theory, repeated presentation of a stimulus causes a neural model of the stimulus to be constructed in cerebral cortex. This model is then compared with incoming stimuli. If the model and the stimulus match, signals are generated that inhibit sensory input to the reticular formation, causing a diminution of the OR. A mismatch caused by presentation of a new stimulus restores the OR and presumably alters the model or loosens its inhibitory hold over the reticular formation's inputs, resulting in dishabituation.

Two experiments crucial to the development of this theory involve the response to a change in stimulus intensity and to the withholding of a regularly repeated stimulus. Sokolov found that after habituation to a stimulus of low intensity, the OR would return when a stimulus of still lower intensity was presented. This result indicated to him that habituation could not simply be due to a decrease in synaptic transmission or to a build-up of inhibition, since a weaker stimulus could not be expected to break through these decremental barriers to reestablish the OR. It was also found that if one of a regular train of stimuli was withheld, the OR would reappear at about the time the stimulus was expected. A theory that involves only a decremental process cannot deal with this effect. However, a theory that posits a neural model that encodes the temporal properties of a stimulus, and match-mismatch circuits able to inhibit and to generate arousal can easily account for this effect. Sokolov's model theory, then, relies on complex coding of stimulus properties in higher centers, or in the reticular formation itself (Hernandez-Peon, 1960), together with inhibitory blocking of arousal to account for habituation.

Thompson's two-process theory was derived from experiments on the hindlimb flexor reflex in spinal cats (Groves and Thompson, 1970). This reflex may be elicited by stimulation of the skin on the limb, and involves a general withdrawal of the limb from the point of stimulation. In 1906, Sherrington studied the habituation of the flexor reflex of the dog. It has since been studied in a variety of animals and under a number of conditions (See references in Griffin, 1970). Thompson and Spencer (1966) showed that habituation of the flexor reflex of the spinal cat displays the nine characteristics of their operational definition. They postulated that the response decrement is caused by a decrease in transmission through chains of interneurons linking cutaneous afferents with motoneurons. Dishabituation in their scheme is caused by a transitory, superimposed sensitization of transmission. The two processes of decrement and sensitization arise and develop independently of one another, and their effects add on a final common path to produce the observed characteristics of reflex habituation. In the case of spinal habituation, Thompson and Spencer consider the decrement to be produced by failure of transmission at single synapses, while others (Wall, 1970; Pearson and MacDonald, 1973; Wickelgren, 1967,

a, b) feel that an inhibitory process is involved. Sensitization, likewise, is felt by Thompson to be the result of increased synaptic transmission or increased cell excitability.

Two experiments crucial to this theory involve the interaction between habituation and dishabituation. First, it can be shown that a strong stimulus causes subsequent sensitization of the response to a nonhabituated weaker stimulus (Spencer, Thompson and Neilson, 1966; Thompson and Spencer, 1966), as illustrated in Figure 1.12. Some form of sensitization occurs in the spinal response, then. In the second experiment it was first established that a stimulus of a given intensity applied at a low frequency caused no habituation, while the same stimulus applied at a higher frequency led to response decrement. This stimulus could then be used at high frequencies to induce habituation and at lower frequencies to test the extent of habituation without disturbing the system. The response was first habituated, then dishabituated with a stronger stimulus. The test stimulus was then used to see how the response recovered.

The result of this experiment is shown in Figure 1.2, taken from Thompson and Spencer (1966). After a brief increase, the response settled back to about the level of maximum habituation, then slowly recovered. If dishabituation is an actual disruption of the decrement that produces habituation, then it should be necessary to present more habituation stimuli to return the response to a lower level. With no further habituation stimuli, the response to a test stimulus should remain at a higher constant level, or increase to-

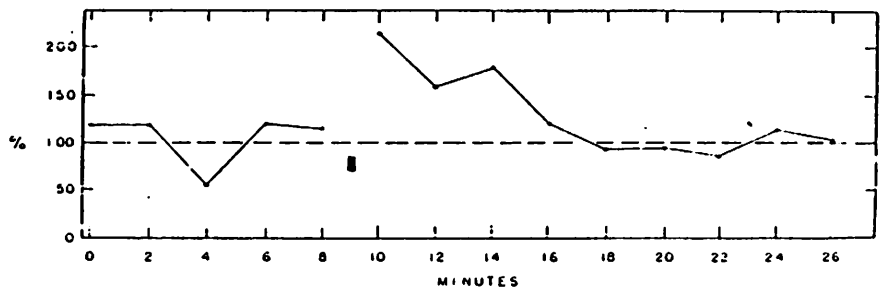


Figure 1.12 Sensitization of response in cat spinal cord. Following a brief shock, applied at the point of the black bar, the response to a less intense stimulus is heightened for a time (Spencer, Thompson, and Neilson, 1966). Abscissa in percentage of control response ordinate in minutes. Reprinted with permission of the author and publisher.

ward the control level as the system recovers. If dishabituation is instead a superimposed response sensitization that decays more rapidly than habituation itself, the response will return to the habituated level with no further stimulation, as occurred in Figure 1.2.

Although these two theories deal on different levels with different responses, they are in fact similar in some respects. Sokolov's theory demands that a stimulus model be encoded in some neural system that is able to influence direct stimulus response channels to produce decrement and sensitization. The decrement, in particular, is due to inhibition produced by signals indicating a match between incoming stimuli and the stimulus model. In Thompson's theory, decrement is produced at the synapses of cells that form the stimulus-response channels themselves. As pointed out by Groves and Thompson (1970) and by Segundo and Bell (1970), depleted synapses are a sort of model of previous stimuli, and further stimuli are compared with this model by means of simple overlap of sets of activated synapses. To Sokolov, new stimuli trigger mismatch signals that result in renewed response. In Thompson's scheme, new stimuli excite undepleted synapses, also resulting in renewed response by a kind of mismatch operation.

Despite this similarity, the two theories differ in ways due primarily to the differences in their origins. Sokolov's theory was derived from studies of the orienting response, and deals primarily with that response. The OR together with associated defensive reflexes form a complex interacting system. In humans, the OR has been shown to be sensitive to cognitive activity (Sokolov, 1961; Graham, 1973) as well as to simpler stimulation. It seems reasonable to employ this concept of stimulus model in conjunction with so complex a system.

It is not easy, however, to generalize Sokolov's theory to simpler organisms or responses. The stimulus model has been defined by Sokolov (1975) as a "matrix of potentiated synapses" which code various stimulus properties. Neuronal circuits have been proposed to encode relevant stimulus parameters (Sokolov, 1975) and to perform model matching operations (Horn, 1967). This neural machinery does not fit the fact that very simple organisms and preparations display habituation. The proposed circuits and the current elaboration of the model theory, do not touch directly on such issues as the relation between habituation and dishabituation, and the realization of the intensity and long-term effects. Sokolov's theory, then can be useful in considering habituation in complex systems, but much work is needed before useful circuit modeling of a particular system can proceed.

The two-process theory of Groves and Thompson (1970, 1973) was derived from experiments on simple responses in simpler systems. The notion of additive interaction between a process of decrement and one of facilitation unifies data from a number of preparations that display habituation. Thompson and his colleagues (Thompson,

et. al., 1973) have gone far in employing the TPT in explaining a variety of experimental results at the behavioral and physiological levels. These explanations have from the start been couched in terms of plausible neuronal circuitry. For these reasons of completeness and of closeness to physiological data, the TPT is attractive to the modeler.

Two features of the TPT leave work for the modeler, however. First, the notion of the "state system" as employed in the TPT is vague as to physiological substrates. The state of the organism is defined as its general level of arousal or readiness to respond to stimuli. Activity in the state system is to increase and then decrease as habituation training proceeds. On this basis, cells possibly involved in the state system of the cat spinal cord have been identified (Groves and Thompson, 1973). In any given system--theoretical or experimental--the state or facilitatory components must likewise be identified on the basis of their behavior during habituation training.

Though the TPT appears to deal satisfactorily with the nine basic characteristics of habituation, there is one phenomenon associated with habituation in relatively complex systems that is not properly explained. This is the phenomenon of temporal conditioning, whereby dishabituation occurs if the duration or period of a regularly repeated stimulus is altered following habituation to that stimulus. Workers both with human OR (Sokolov, 1961, 1975; Pendergrass and Kimmel, 1968; Graham, 1973) and with single-cell responses

in a variety of brain regions of animals (Vinogradova, 1975) have reported findings of temporal conditioning. The effect may be weak, however (Vinogradova, 1975), and its exact workings are controversial (Graham, 1973).

In dealing with temporal conditioning, Groves and Thompson (1970) relegate the effect to entrainment of the state system to the regularly repeated stimulus. Such entrainment requires that the state system be capable of supporting long-period oscillations (up to 40 sec. in the case of the study by Pendergrass and Kimmel (1968)). Entrainment of that sort is in effect a temporal model of the stimulus, kept by the state system. Sokolov (1975) has presented neural circuits employing time delays to realize temporal conditioning, though it is not clear how long time delays are to arise in neural systems. In the case of this effect, then, the TPT makes a much closer approach to the model theory of habituation. Both the TPT and the model theory require the invention of plausible neural circuitry to deal with temporal conditioning. Whether this circuitry is to reside entirely within the state system, as suggested by Groves and Thompson (1970), or is to be used with model matching circuits as required by Sokolov (1975) will depend upon the particular system being modeled.

This consideration of the two major habituation theories suggests that relevant analytical modeling can proceed in two main directions. Both the TPT and the model theory require more work concerning temporal conditioning. One brain region whose responses

display temporal conditioning is the hippocampus (Vinogradova, 1975). The hippocampus is a region of cerebral cortex that has been implicated in learning and memory (see discussion and references in Chapter V); as such it is a prime candidate for modeling of plasticity. In Chapter V, I detail a model of a temporal sequence memory based on the structure of a part of the hippocampus called the dentate gyrus. This memory circuit may be used to allow the hippocampus to habituate to the temporal qualities of stimuli in a manner suggested by Vinogradova (1975), or to become more aroused when presented with regularly repeated stimuli, as in the TPT (Groves and Thompson, 1970). The main feature of this model is its use of propagating waves of activity to obtain long time delays.

The TPT is reasonably complete in its handling of the other characteristics of habituation, as elaborated in the work of Groves and Thompson (1970, 1973; Thompson, et. al., 1973). As mentioned above, this theory has been presented in terms of simple neuronal circuitry. Groves and Thompson (1970) illustrated the TPT with the circuit shown in Figure 1.13. In this circuit, synaptic depression occurs in the direct stimulus-response channel, as indicated by synapses marked H. Facilitation occurs in the state portion at synapses marked S. The state system also displays significant circuit dynamics, as indicated by the cycle of cells in Figure 1.13.

Though it is reasonable that this circuit should be able to duplicate the properties of habituation discussed by Groves and Thomp-

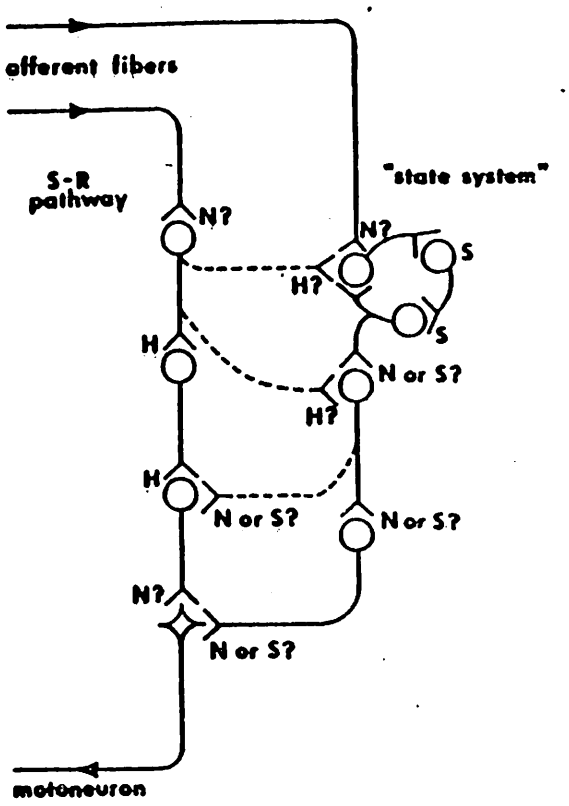


Figure 1.13 Two-process habituation circuit of Groves and Thompson (1970). Circles represent cells, N denotes nonplastic synapse, S denotes sensitizing synapse, D denotes depleting synapse. Reprinted with permission of the author and publisher. Copyright 1970 by the American Psychological Association.

son (1970), its operation remains to be explored quantitatively. As discussed previously, an analytical model of such a two-process circuit would help to tell what is demanded of its components in order that it function properly. Some of the properties of habituation may be seen to be crucial in guiding the modeling effort--and hence may be crucial to experimental investigations of habituation--while others may follow directly from the modeling of the more crucial properties. It may be found that some of the characteristics of habituation can be modeled independently of the more vital properties, and that in this regard they do not emerge as fundamental properties of habituation.

Chapters II through IV of this work, then, deal with analysis and simulation of a simplified version of the two-process circuit of Figure 1.13. It will be shown that a simple mathematical framework can embrace the properties of response decrement, facilitation, and spontaneous recovery. These response properties are related to properties of synaptic modification through simple rules of cell or cell population operation. In Chapter II, I show how one form of experimentally observed intensity effect and a form of frequency effect can arise in a two-process circuit. Modeling of other habituation characteristics is sketched there more briefly. In Chapter III, I deal with the intensity effect in a more detailed circuit employing a population of cells and afferent fibers. I show there that different intensity effects can arise, depending on circuit parameters and on the forms of the synaptic modification rules. This work is summarized in Chapter IV. Together, these results

show how some simple central mechanisms can give rise to observed properties of response habituation.

CHAPTER II

LUMPED MODELS OF THE TWO-PROCESS THEORY

I have chosen the two-process theory of Groves and Thompson (1970) as the basis of this study of habituation networks because the TPT is more complete in its handling of habituation data than Sokolov's (1961) model theory, and more detailed in terms of plausible neuronal realizations. A network is to be found that embodies the assumptions of the TPT and that is amenable to mathematical analysis and computer simulation. Using this network, it should be possible to show ways the two central processes of decrement and sensitization can interact to produce the observed properties of response habituation. Furthermore, it should be possible to relate the ways the processes vary with continued stimulation to plausible characteristics of cell operation and synaptic variation.

Not all of the habituation characteristics should be included in the general process-level model. Trying to model all ten characteristics of the operational definition of habituation would most likely lead to a system complex enough to obscure its own workings. As noted in Chapter I, it seems best to concentrate on those characteristics that appear most crucial to the workings of habituation. Other characteristics may arise naturally in a model of the more vital properties. Still other properties of habituation may be included in a basic model through the use of additional operational mechanisms. Characteristics that do not seem vital to the workings of the basic model will therefore not be included in the

initial modeling effort. Models that include these characteristics will be presented following discussion of a basic habituation circuit.

Which properties of response habituation seem crucial to the form of the network model? Response decrement and sensitization form the basis of habituation in the TPT. Proper modeling of these effects will allow realization of habituation, dishabituation, and their spontaneous recovery with time. It will be seen later how the effects of stimulus frequency come naturally out of the characteristics of spontaneous recovery. Of the remaining properties of habituation, the effect of stimulus intensity seems the most crucial. It is this effect that determines the basic anatomy of the two-process circuit, and which necessitates the use of synaptic modification rules with particular forms. Within the framework of a circuit that properly realizes the desired intensity effect, the other characteristics can easily be realized.

The intensity effect actually takes a number of forms in different physiological systems. Figure 1.6 illustrates one form of intensity effect demonstrated by the flexion reflex of the spinal cat. In this case, relative habituation decreases as stimulus intensity is increased. A similar effect is shown in the ventral root response to lateral column stimulation in the isolated frog spinal cord (Farel, et. al., 1973), as illustrated in Figure 2.1.

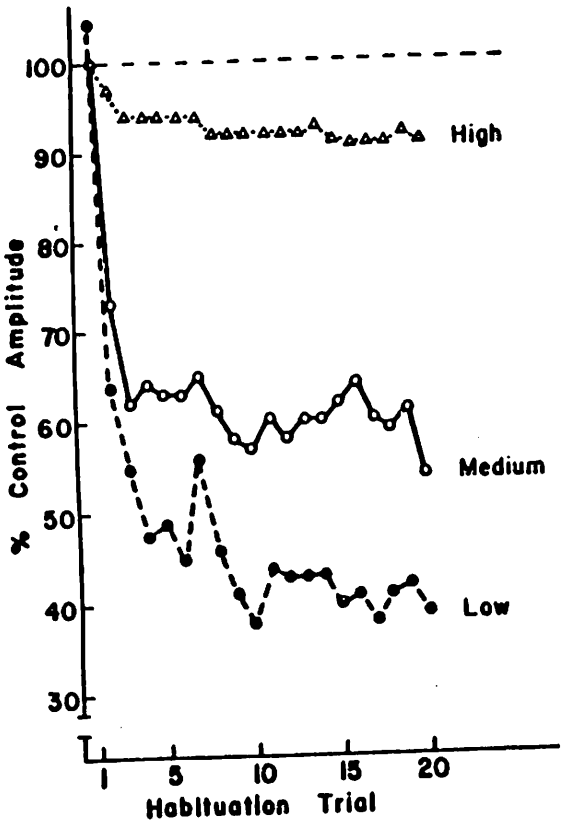


Figure 2.1 Habituation of the ventral root response to lateral column stimulation (Farel, et. al., 1973). Relative response is plotted for stimuli of various intensities. Reprinted with permission of the author and publisher.

The cat flexion reflex displays another form of intensity effect, as shown in Figure 2.2. In this form, at the lowest stimulus intensity, relative response falls rapidly with the first few blocks of stimuli and reaches a low asymptotic level. At a moderate intensity, sensitization gives rise to an initial increase over the control level, followed by a decrease to an intermediate level in later trials. The peak value at the highest intensity is greater, occurs later, and is followed by a fall to an asymptote that may be above the control level.

The curves of Figure 2.2 are more complicated than the simple description given above, however. At the highest intensity, a dip in the curve occurs before the peak begins to build. Groves and his colleagues (Groves, Lee and Thompson, 1969; Thompson, et. al., 1973) suggest that sensitization in the spinal cat consists of two components with different time courses and different reactions to high intensity stimuli. A fast component that dominates responses in early trials actually decreases with increased stimulus intensity, while a slower component that predominates in later trials increases in size with increasing intensity. Other authors have also suggested that sensitization in general consists of more than a single component (Graham, 1973; Goodman and Weinberger, 1973). Mixtures of such effects can be used to fit the habituation curves generated by experiments on particular preparations.

A third kind of intensity effect is found in the ventral root response to dorsal root stimulation in the isolated frog spinal cord

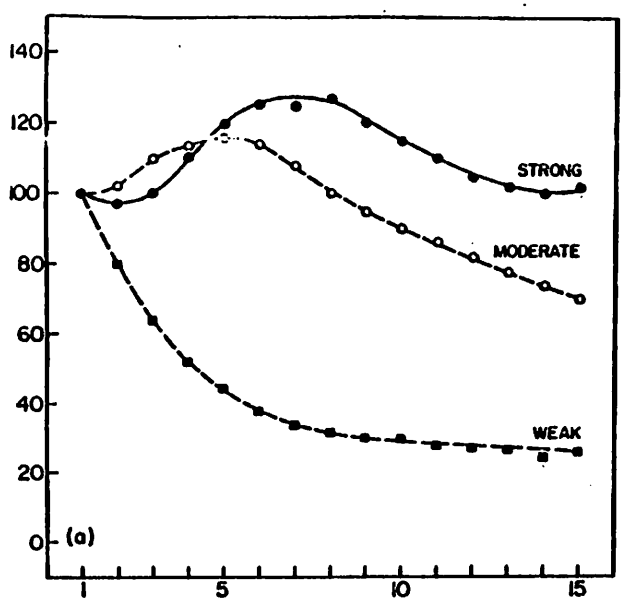


Figure 2.2 Curves of relative habituation of the flexion response of the cat generated by stimuli of different intensities (Thompson, et. al., 1973). Intensities as marked on graph. Abscissa in percent of control, ordinate in blocks of trials. Reprinted with permission of the author and publisher.

(Farel and Thompson, 1972), as illustrated in Figure 2.3, a and b. It may be seen that absolute decrement of response depends directly on stimulus intensity, while for higher intensities, relative habituation is essentially independent of intensity.

It is apparent, then, that the intensity effect is not simple, but may take different forms in different preparations. In a monosynaptic system such as the lateral column of the frog spinal cord, the cause of a given type of intensity effect must be sought in synaptic morphology and biochemistry. Network effects are likely to contribute to the stimulus intensity characteristics of more complex systems, however. It is to such systems that the current modeling effort is addressed. In what follows, it will be shown how curves similar to those of Figure 2.2 can be generated in a simple two-process network. I will consider only systems with unitary sensitization components. Results of investigations of these systems could be used to help guide more complicated modeling efforts. It will be shown, for example, that in a two-component circuit, the position of the peak value of relative response shifts significantly to later trials as stimulus intensity increases. This effect must contribute to the peak shifting seen in systems having multi-component sensitization. A part of the network will generate curves of the form shown in Figure 2.3b, operating according to the dictates of the current TPT (Thompson, et. al., 1973). Curves of the form shown in Figure 1.6 can also be obtained easily with the proper parameter settings. Thus, the model circuit will be shown capable of realizing these three main forms of the intensity effect.

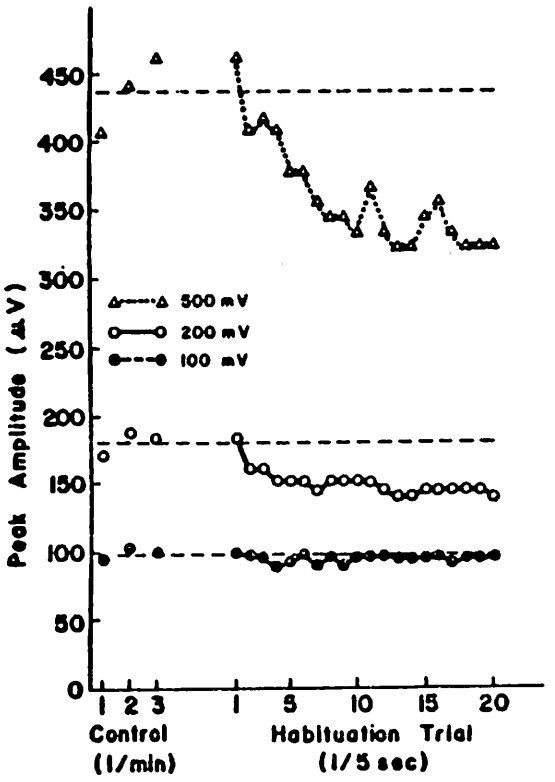


Figure 2.3a Absolute responses during habituation of the ventral-root response to dorsal root stimulation of the isolated frog spinal cord (Farel and Thompson, 1972). Curves are given for three intensities of stimulation. Note that the absolute decrement increases as response intensity increases. Reprinted with permission of the author and publisher.

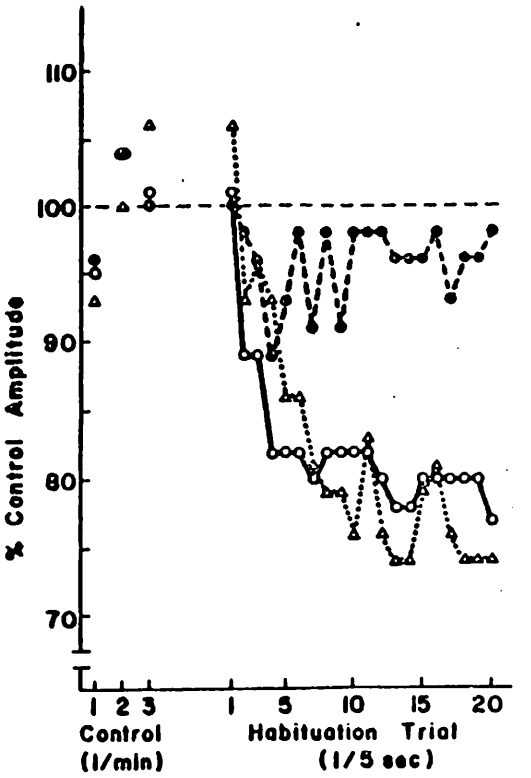


Figure 2.3b Curves of relative response decrement computed from the absolute response curves of 2.3a. Lack of habituation at the lowest intensity is due to stimulation of a nonhabituating monosynaptic pathway. Reprinted with permission of the author and publisher.

The circuit modeling to follow, then, concentrates on the properties of decrement and facilitation and their spontaneous recovery, and on the intensity effect as discussed above. The stimulus generalization effect will be discussed later as a simple extension of the basic two-process circuit. Likewise, the below-zero and long-term effects will be added as extensions of the basic model. This model thus provides an understandable framework within which these characteristics of habituation can be easily related to circuit and synaptic operation.

Qualitative Models of the Two-Process Theory

A simplified two-process circuit. As mentioned in Chapter I, Groves and Thompson (1970) have presented an habituation circuit based on the TPT. Their circuit realizes the TPT using homosynaptic depression and sensitization, as illustrated in Figure 1.13. The circuit consists of several cells that make up separate stimulus-response and state systems. Circuit response is the sum of activity in these two systems, as required by the TPT. Cell and circuit dynamics (as exemplified by the cycle of cells in the state system) act to shape the circuit's response. Response plasticity is a function of the plasticities of the many decrementing and incrementing synapses and of their interactions through connections between the cells. Groves and Thompson do not explicitly specify the forms of dependence of decrement and increment on pre- and postsynaptic activity. Complicated modification rules could add significantly to the complexity of operation of this circuit.

While computer simulation is well suited for dealing with the complexities of this circuit, the results of such simulation might not be readily interpretable. Diverse factors could easily combine to yield effects that would not be easy to explain. Likewise, many modes of circuit operation might be possible, and some could be missed in exploration of model parameter settings. Accordingly, I have chosen to investigate a network based on the one of Figure 1.13, but simplified in two ways.

First, I will not consider the detailed operation of the SR and state pathways. Each pathway will instead be lumped into a single element. The elements of the following models, then, represent processes rather than the cells that realize the processes. Following the dictates of the TPT, the outputs of the decremental and incremental processes will sum on an output element that generates the circuit's response. The response of each process will vary with stimulation in a manner suggested by the physiological results discussed earlier. Modification rules and circuit parameters will be found that enable the simplified process-level circuit to realize the desired habituation characteristics.

The second simplification involves the exclusion of the effects of circuit and cell dynamics. The response of each process will be assumed to be a function of its current input only. The time scale considered--that of synaptic change--is assumed to be great enough that the circuit dynamics can be ignored. Under this assumption, the system is active only during stimulation, and responds to all

stimuli with a stereotyped discharge. Relaxation of this assumption is discussed following presentation of a more detailed population-level model of habituation in Chapter III.

In this chapter, the investigation of two-process habituation networks proceeds in two steps. First, a simple circuit based on the network proposed by Groves and Thompson (1970) is analysed and simulated. Numerical values are found for the parameters of the simplified system so that it displays the desired intensity and frequency effects. Secondly, mechanisms are added to the basic model to realize the habituation characteristics not initially modeled. In the following sections, I first describe the TPT in greater detail, and present the process-level model and results of its simulation.

Details and first models of Thompson's two-process theory. The basic assumptions underlying the TPT, as given by Thompson and coworkers (1973) are:

1. A stimulus both elicits a response and alters the state of the organism. The stimulus-response (SR) path is considered the most direct route through the organism from stimulus to response. The state of the organism refers to its general level of arousal as discussed in Chapter I.

2. Repetition of a stimulus results in an "inferred decremental process" in the SR path, called habituation. This central decrement is the basis of the empirically observed response decrement.

3. Repetition of a stimulus results in an "inferred incremental process" in the state system, called sensitization, that forms the basis of dishabituation.

4. Decrement and increment develop independently of one another but combine to yield the final response.

5. A stimulus that is presented following habituation to another stimulus elicits a decreased response to the extent that it activates elements in the SR path that were activated and habituated by the first stimulus. This "common elements" hypothesis applies as well to the state system, where the response to one stimulus is augmented to the extent that it activates elements that are sensitized following previous stimulation.

This amounts to a restatement of the TPT as given originally by Groves and Thompson (1970).

The first four characteristics of central decrement and sensitization are embodied in the network of Figure 2.4, which is based on the more complex network of Groves and Thompson (1970). Decremental and incremental processes, marked D and S, respectively, in the figure, sum at a final common point to produce the output. If the outputs of D and S change properly with time, number of stim-

uli, and stimulus intensity, then the overall output will display the desired habituation characteristics. A network composed of a number of such circuits as basic elements can employ the common-elements scheme to realize generalization effects. Such a network is illustrated in Figure 2.5. There, the D-process boxes are cross-coupled such that activation of an SR channel causes decrements in nearby channels. This approach, which employs coupled response channels to realize a general habituation network, is similar to those of Horn (1967) and Segundo and Bell (1970).

Note that this network models habituation entirely on a process level, dealing only with the ways the decremental and incremental processes might interact to yield the characteristics of empirical response habituation. So far, only the common elements notion suggests a mechanism involved in the operation of the underlying neural circuitry. It is necessary now to fill in the operational details of the D and S boxes.

Further details. The inferred central processes of decrement and sensitization must be given properties that enable them to realize the defining characteristics of habituation. Thompson and co-workers (1973) give central habituation and sensitization the properties detailed below.

a. During habituation training, habituation proceeds exponentially to its asymptote, while sensitization first increases then decreases.

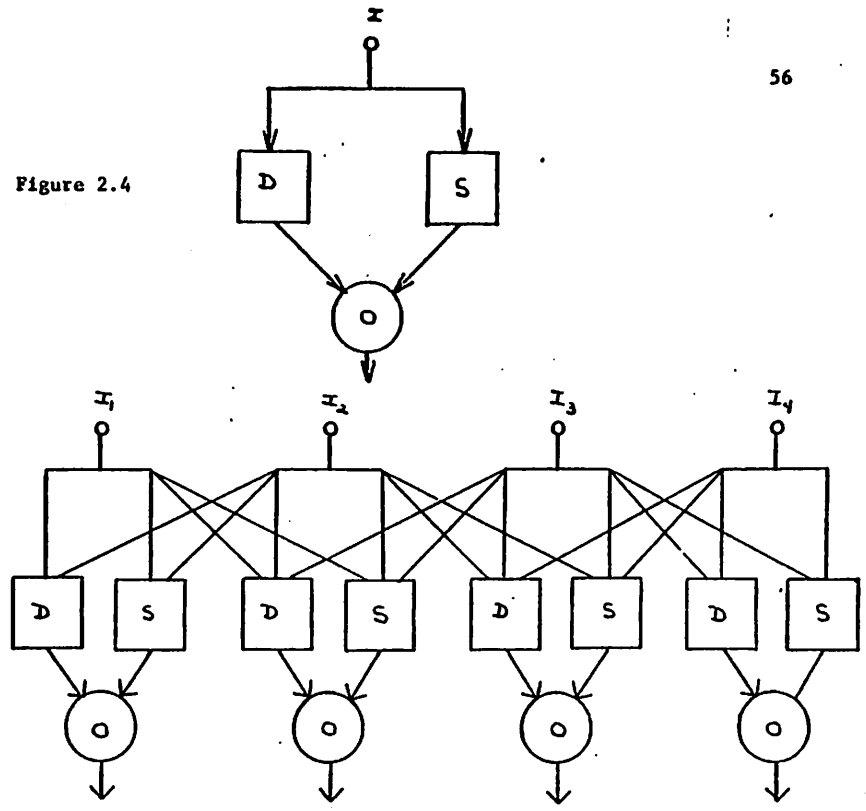


Figure 2.5

Figure 2.4 Lumped two-process model. Boxes represent plastic processes, D denotes decrementing process, S denotes sensitizing process. Box outputs add at produce overall circuit output.

Figure 2.5 Generalization of habituation with lumped model. Boxes are cross-coupled such that stimulation of one causes decrement and sensitization in others.

b. Upon cessation of stimulation, both habituation and sensitization spontaneously recover.

c. The degree of relative habituation is weakly and inversely related to stimulus intensity. Note that this means that absolute decrement due to habituation is directly related to intensity.

d. The degree of relative habituation is strongly and directly related to stimulus frequency.

e. The degree of sensitization is directly related to stimulus intensity and to stimulus frequency at higher intensities.

f. Repeated series of habituation training result in progressively more, or more rapid habituation.

g. Repeated presentations of a sensitizing stimulus result in progressively less sensitization.

These characteristics are sufficient for the sum of central decrement and sensitization to display the frequency, intensity, and long-term properties described in the operational definition of habituation. Characteristics (a) and (b) are directly related to the observed time course of habituation in a number of preparations that display response sensitization. These points refer to points (1) and (3) of the operational definition. The notion that sensitization first increases then decreases is somewhat controversial. Graham (1973) has shown that sensitization that only increases with successive stimulation may be combined with a decremental effect to give curves that first increase, then decrease. This behavior will be verified in the models below, and the role of decreases in sensitization will also be explored.

Points (c), (d) and (e) of the extended definition deal with the effects of stimulus intensity and frequency on habituation, and relate to (4) and (6) of the operational definition. It will be verified later that two frequency ranges may be defined in terms of the recovery properties of the synapses in question. In a range of moderate frequencies, little recovery from habituation occurs between stimuli. In that range, habituation is independent of frequency and depends only on the total number of stimuli. Thompson makes the following simplification of (d), which I have interpreted as above in terms of the recovery time of the decrement:

d'. Within a range of moderate frequencies, relative habituation is independent of stimulus frequency, and depends only on the number of stimuli presented.

For lower frequencies, recovery occurs between stimuli, and relative habituation depends more strongly on frequency.

Finally, the dependence of habituation on stimulus intensity is simplified in a restatement of (c):

c'. Relative habituation is independent of stimulus intensity.

As discussed previously, the ventral root response to electrical stimulation of the dorsal root of the frog spinal cord exhibits intensity-independent relative habituation (Farel and Thompson, 1972). In point (c'), Thompson extends this to a property of habituation in general. It will be shown in the models below that the intensity effect of point (6) of the operational definition of habituation can easily be realized by the sum of habituation and sensitization operating under characteristics (c') and (e) as defined here.

A Quantitative Model of the TPT

In this section I describe a process-level model based on the networks of Figures 1.13 and 2.4 designed to realize the stimulus frequency and intensity effects described earlier. I will show how the detailed characteristics of the previous section may be used to fill in the D and S boxes of Figures 2.4 and 2.5 in a way amenable to analysis and computer simulation.

The simplest way to represent the operation of the D, S and output boxes is to assume that the output of each is a weighted sum of its inputs. This operation is in fact too simple, and will be shown later to lead to a conflict with a requirement of a more detailed statement of Thompson's theory concerning intensity generalization. This assumption is worth pursuing, however, as a first approach to

habituation modeling. In this case, the operation of the unit of Figure 2.4 is given by

$$D = W_0 I$$
$$S = W_1 D$$
$$O = W_2 D + W_3 S$$

where I is the intensity of the input to the unit, and W_j are the coupling weights between processes, $j = 0, 1, 2, 3$. Figure 2.4 is redrawn in these terms in Figure 2.6. There, the direct path from D to O represents the SR channel. The state system is represented by the circle labeled S, and the state system's effect on the output is represented by the total transmission from D to O through S. In this model, the output of each node may be considered to be the level of activity of an associated set of cells, and the coupling weights may be taken as the total synaptic strength between cell populations. In what follows, I will refer to the nodes as cells, for the sake of simplicity.

If the output of this network is to change with successive stimulus presentations, the coupling weights themselves must change with use. The ways the weights change must be consistent with the principles put forth in the previous section. Both exponential decay toward an asymptote and spontaneous recovery can be modeled by representing each weight with a first-order differential equation of the form:

$$\tau_i \dot{W}_i = W_{i0} - W_i + a_i J_i,\ W_i(0) = W_{i0},\ i = 0, 1, 2, 3,$$

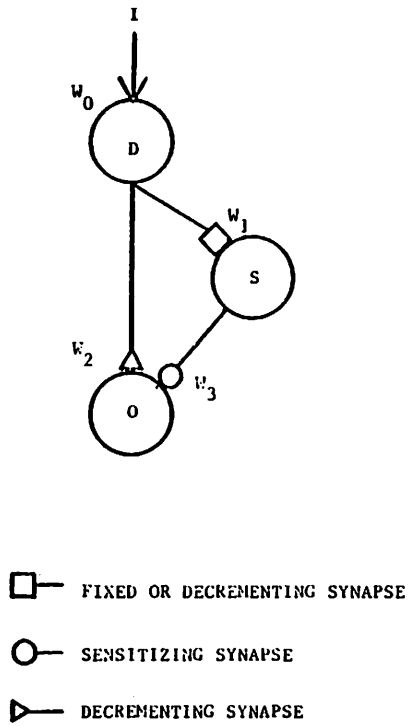


Figure 2.6 Lumped model presented in terms of cells and coupling weights. Synapse modification properties are as indicated on the figure.

where

τ_i is the weight's modification time constant,

W_{i0} is the resting and initial weight value,

J_i is the modification input, and

a_i is the weight's modification gain.

Each weight, W_i has a resting value, W_{i0}, toward which it returns following displacement by the modification stimulus, J. If a_i and J_i are positive, presentation of the modification stimulus will increase the weight value. If their product is negative, the weight will be decreased by the modification input. These two effects simulate sensitization and habituation, respectively.

Differential equations can now be associated with the weights in Figure 2.6. For the sake of simplicity, the weight connecting the afferent, I, with the D-cell will be considered fixed. Any plasticity it might display can just as easily be embedded in the other weights. Central habituation is to occur in the SR path, so W_2 must be decreased by the output of D. Sensitization is to occur in the state system, and is to increase then decrease across trials. The transmission from input to output through the state system is simply the product of W_1 and W_3. This product will increase and then decrease if W_1 decreases slowly while W_3 increases rapidly, or vice versa. The former case will be considered here. Following the notion that habituation is likely to be due primarily to presynaptic effects, each weight will be modified by presynaptic activity only.

These considerations lead to the following equations for the weight values:

$$\tau_1 W_1 = W_{10} - W_1 - a_1 f_1(D)$$
$$\tau_2 W_2 = W_{20} - W_2 - a_2 f_2(D)$$
$$\tau_3 W_3 = W_{30} - W_3 + a_3 f_3(S)$$

where

$W_i(0) = W_{i0}$, $i = 1, 2, 3$, and

$f_i(\cdot)$ determines the way each weight is changed by its modification input. Each f_i is positive for all values of its argument, and each a_i is positive.

It must now be shown that this system fulfills the detailed assumptions on the operation of central habituation and sensitization given in the previous section and that it realized the desired habituation characteristics. As mentioned above, the use of weights governed by differential equations ensures that central habituation will fall toward an asymptote and recover when left unstimulated. Similarly, central sensitization can be made to increase then decrease across trials, as will be seen below. Sensitization will also recover following cessation of stimulation. The model mirrors these basic properties of habituation and sensitization, then.

If relative habituation is to be independent of stimulus intensity, an odd assumption is necessary concerning the way W_2 is modified by the D-cell's activity. To see this, suppose that the

output depends only on the SR path, so that only the properties of central habituation are reflected in the response habituation. The control-level output (the output before habituation training) is

$$O_c = W_{20} I,$$

assuming W_0 equal to one, so that $D = I$. After habituation to asymptote, W_2 has been decreased according to the differential equation above to the value

$$W_2 = W_{20} - a_2 I$$

where it is assumed that $f_2(D) = D$ for the sake of illustration. The output after habituation is given by

$$O_H = (W_{20} - a_2 I) I,$$

and the absolute amount of habituation is

$$A = O_c - O_H = a_2 I^2.$$

As desired, this depends on the intensity of the stimulus. The relative habituation, however, is

$$R = A/O_c = \frac{a_2 I^2}{W_{20} I} = \frac{a_2}{W_{20}} I,$$

which is not independent of I. A similar result will be obtained for any function f_2 whose value increases with increasing I.

In order to obtain a relative decrement that is independent of stimulus intensity, the decrement of W_2 must itself be independent

of intensity. Then

$$O_H = (W_{20} - a_2)\, I,$$

$$A = a_2\, I,$$

and

$$R = \frac{a_2}{W_0}$$

Intensity-independent relative habituation can be achieved in this model only when the synaptic decrement is independent of intensity. As will be discussed later, a further statement of Thompson's theory requires a synaptic decrement that depends on stimulus intensity. The simple model considered here will give way at that point to a more detailed population model. The current model may still be used in an initial exploration of the characteristics of decrement and sensitization required by the definition of the previous section.

Central habituation is independent of stimulus frequency in this model when no recovery occurs between stimuli. In that case, the degree of habituation depends only on the number of stimuli presented. This independence occurs when the time constant of W_2 is considerably greater than the interstimulus interval. Similarly, central sensitization depends on stimulus frequency within a particular range when the time constant of W_3 is on the order of the interstimulus interval in that range. The model can therefore realize the desired frequency characteristics if the time constants are properly adjusted.

Intuitive arguments show that this simple network, with coupling weights that vary according to first-order differential equations, can be made to realize the desired frequency and intensity characteristics. Detailed numerical evaluation of this system is carried out below to investigate the intensity and frequency effects.

Computer Simulation of Intensity and Frequency Effects

Introduction. The model described above has nine parameters that must be specified in order to generate habituation curves of the sort shown in Figure 2.2. The weight time constants have been loosely specified in terms of their relation to interstimulus interval in some range, but the parameters have not been determined beyond that. In order to generate qualitatively correct relative habituation curves at different stimulus intensities, I consider two cases. In the first case, the stimulus is assumed to be applied continuously, rather than periodically. The effects of recovery between stimuli are not included in this case, then, corresponding to the situation discussed above in which moderate stimulus frequencies are employed. Further, the synaptic modification functions employed in this case are linear functions or constants. This situation is simple enough to be solved analytically, leading to an expression for relative habituation as a function of time. A computer program called HABIT was designed to evaluate this expression and to generate the resulting curves of relative habituation. The operation of the circuit of Figure 2.6 was explored in this simplified case using HABIT.

Next, a more complicated computer program called UATION was implemented to carry out numerical integration of the synaptic differential equations. This program enabled me to explore the frequency effects using non-constant stimuli. The effects of nonlinear modification functions were also investigated. Using results from HABIT, parameter ranges were easily found within which the desired habituation characteristics could be realized. I describe below the results of simulations with these two programs.

Simulation, Part 1: HABIT. The HABIT program was designed to explore the ways the network of Figure 2.6 can duplicate the intensity effect. Under the assumptions of continuous stimulation and linear or constant weight modification, the synaptic equations in the last section may be solved analytically, yielding:

$$W_1(t) = W_{10} - a_1(1 - e^{-t/\tau_1})$$

$$W_2(t) = W_{20} - a_2(1 - e^{-t/\tau_2})$$

$$W_3(t) = W_{30} + a_3 I(W_{10} - a_1)\ (1 - e^{-t/\tau_3}) + a_1 a_3 I\ \frac{\tau_1}{\tau_1 - \tau_3}(e^{-t/\tau_1} - e^{-t/\tau_3})$$

These equations are derived in Appendix A. The HABIT program allows exploration of the possible parameter settings to determine ranges of parameters that give the desired habituation curves.

Curves of relative habituation are to be generated that are similar to those of Figure 2.2 illustrating the variation of relative habituation with time and with stimulus intensity. In the case of continuous stimulation, the curves are functions of time rather

than of the number of discrete stimuli. Qualitatively, the curves generated by the program with low intensity stimuli should be negative exponentials of the number of stimuli. As stimulus intensity increases, relative habituation should decrease, and sensitization--responses greater than the control level--should appear. At the highest intensities, the response may show only sensitization, but the sensitization should increase and then decrease with time. The model will have duplicated the effects of stimulus intensity on relative habituation with the generation of such curves.

Two strategies may be used in the system of Figure 2.6 to generate curves that increase then decrease similar to the habituation curves of Figure 2.2. In Strategy I, the D to S weight is held constant, while the D to O weight decreases and the S to O weight increases. In this strategy, then, central sensitization only increases across trials. If W_3 increases more rapidly than W_2 falls, the output will first increase then decrease. This strategy works best when the D to O path dominates the output before habituation. For example, if the initial transmission through the SR path is equal to that from D to O through S, then decreased transmission in the SR path will lead to a maximum of 50% habituation. In this strategy, then, large percentages of relative habituation are obtained at low stimulus intensities when the initial value of W_2 is considerably greater than the product of W_1 and W_3. At higher intensities, the increase in W_3 leads to reduced drop in overall output, as desired. This strategy is illustrated in detail below.

In Strategy II, W_2 is held fixed while W_1 decreases and W_3 increases with time. As mentioned above, if W_1 decreases more slowly than W_3 increases, the overall output will first increase then decrease. At a low intensity, the drop in W_1 will cause a significant drop in overall output when the path through the state system dominates the output. Hence, Strategy II works best to produce pronounced relative habituation at a low intensity when W_2 is less than the product of W_1 and W_3. At a higher stimulus intensity, the increase in W_3 causes an initial increase in response, but a significant drop in W_1 leads to decreased outputs at later times, as shown below.

The curves shown in Figures 2.7 through 2.12 illustrate these strategies. Each figure shows relative habituation curves generated by evaluating the equations above with the HABIT program. In each case, the abscissa is marked in terms of percentage of the unhabituated output and the ordinate is given in arbitrary units of time. Following the condition derived in the previous section concerning intensity-independent relative habituation, weights W_1 and W_2 decrease independently of stimulus intensity. Weight W_3 is modified by a linear function of the output of cell S, reflecting the fact that sensitization is to increase with increasing stimulus intensity. Table 2.1 gives the values of parameters used to generate the figures referred to below. Note that the weight time constants are expressed in terms of the same units of time used to prepare the figures that follow.

TABLE 2.1

Parameter Values Used With HABIT

Figure #	W_1	W_2	W_3	τ_1	τ_2	τ_3	a_1	a_2	a_3
2.7	1	1	.1	12	9	3	0	.9	.15
2.8	1	3	.1	12	9	3	0	2.9	.4
2.9	1	1	.1	12	12	4	0	.9	.15
2.10	1	0	1	9	6	6	.8	0	.5
2.11	1	1	.1	12	9	3	.2	.9	.15
2.12	1	1	1	12	6	3	.5	.9	.35

Figure 2.7 illustrates Strategy I. There, the SR path dominates at low intensities, W_1 is fixed, and W_2 changes more slowly than W_3. At the lowest intensity, the curve is a negative exponential with no peak. At moderate intensities, an initial peak is followed by a drop to a value below the control level. At high intensities, the peak gives rise to a level above the control level. Note that the position of the maximum value shifts with stimulus intensity, to occur later at higher intensities. This is a general property of these curves, as described in Appendix A. A similar peak shift occurs in the data from the cat spinal cord, as discussed earlier. While some of the shift is likely to be due to the combined effects of different weight time constants, the property illustrated here must also be a factor.

Note, too, that sensitization only increases across trials here. One assumption of the TPT is that sensitization first increases then decreases across trials (Groves and Thompson, 1970). It has been suggested by Graham (1973) that the decrease is unnecessary for the production of curves like those of Figure 2.2. The result of Figure 2.7 shows this to be the case. It will be seen later, however, that decreasing sensitization can be employed in a parsimonious explanation of the habituation of dishabituation.

When the SR path is a more dominant factor in the output, greater relative habituation is generated with low intensity inputs at the expense of the size of the peak at higher intensities. This effect is illustrated in Figure 2.8, where the value of W_2 is greater than

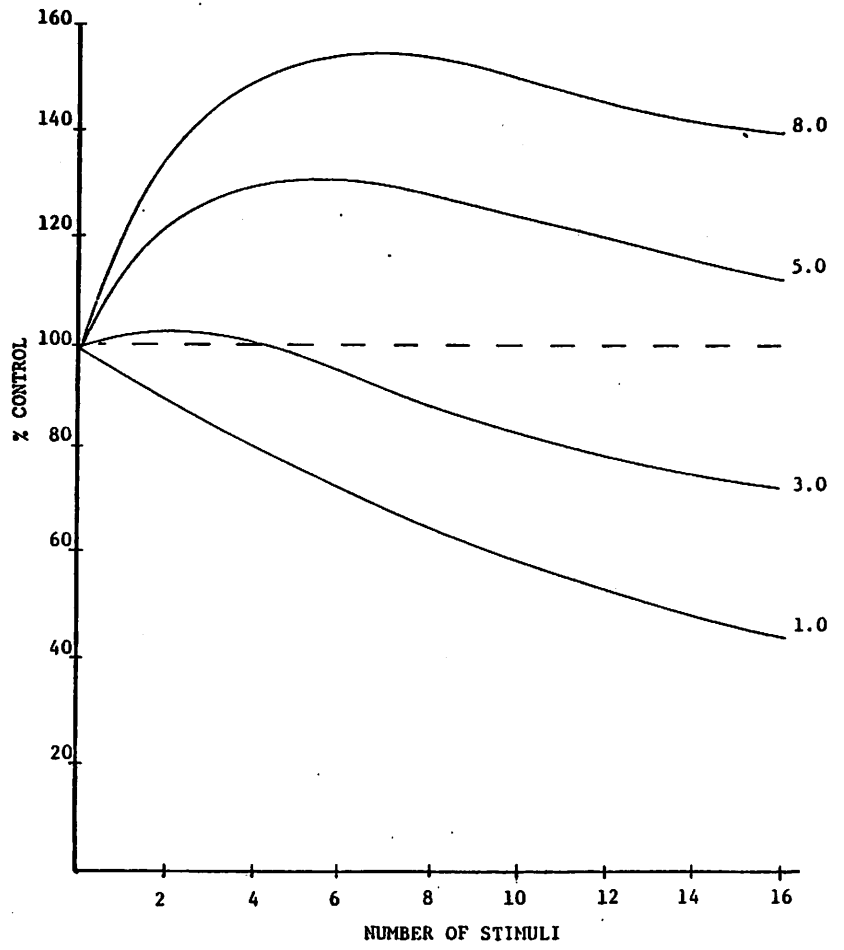


Figure 2.7 Curves of relative response generated with HABIT. Stimulus intensity marked on each curve. Associated parameters are given in Table 2.1.

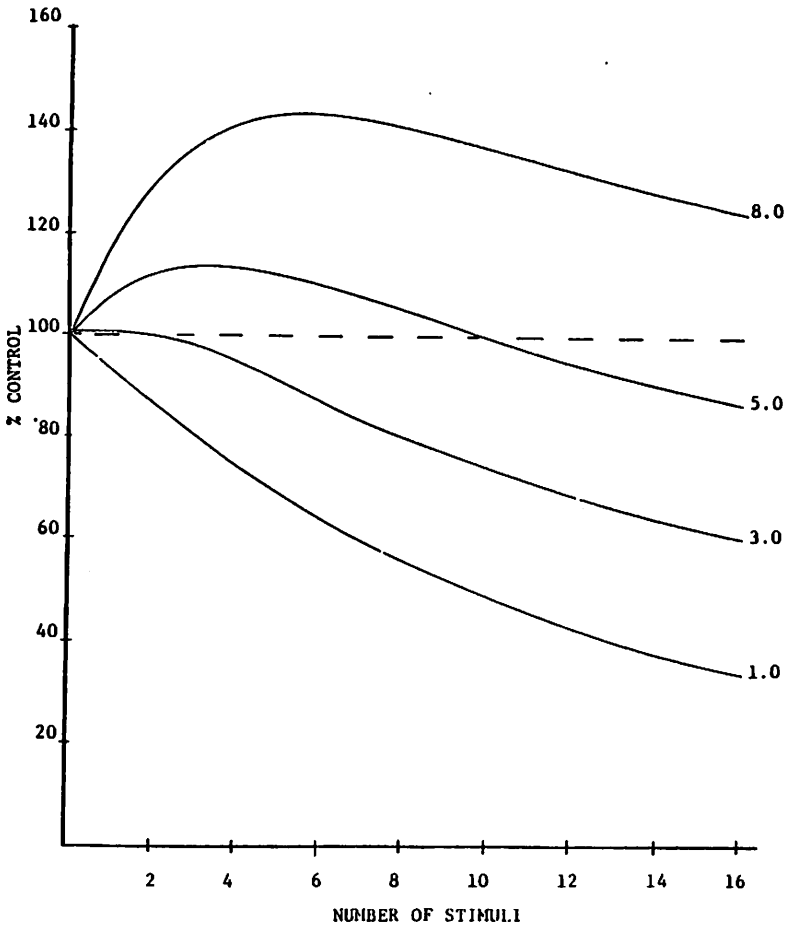


Figure 2.8 Curves of relative response generated with HABIT. Stimulus intensity marked on each curve. Associated parameters are given in Table 2.1.

that used to generate Figure 2.7. The modification gains have also been altered such that the final value of W_3 is greater. A still larger W_3 modification gain would result in curves more like those of Figure 2.7. The curves of Figure 2.7 and 2.8 illustrate the dependence of the model's response on the absolute values of the synaptic coupling weights. Curves of the proper shapes can be obtained with any value for W_2; only the relative values matter in determining relative response levels. In what follows, W_2 will be set to unity.

Figure 2.9 shows that weights with larger time constants result in curves that change more slowly, as would be expected. In particular, the peaks at higher intensities are somewhat broadened and occur at later times. It was mentioned earlier that the later peak in the high intensity curve of Figure 2.2 might be due to a sensitization component with a large time constant. Figure 2.9 verifies that such a component gives rise to a later peak.

Strategy II is illustrated in Figure 2.10 for the case that the transmission through the direct SR path is identically zero. A large W_3 modification gain is employed to obtain greater peak levels at high stimulus intensities. To get large decrements at lower intensities, W_1 must be made to decrease considerably, leading to asymptotic output levels that differ little with different stimulus intensities. This compression of asymptotes is a disadvantage of Strategy II, as habituation data generally show greater final differences across intensity. This result makes it more likely that

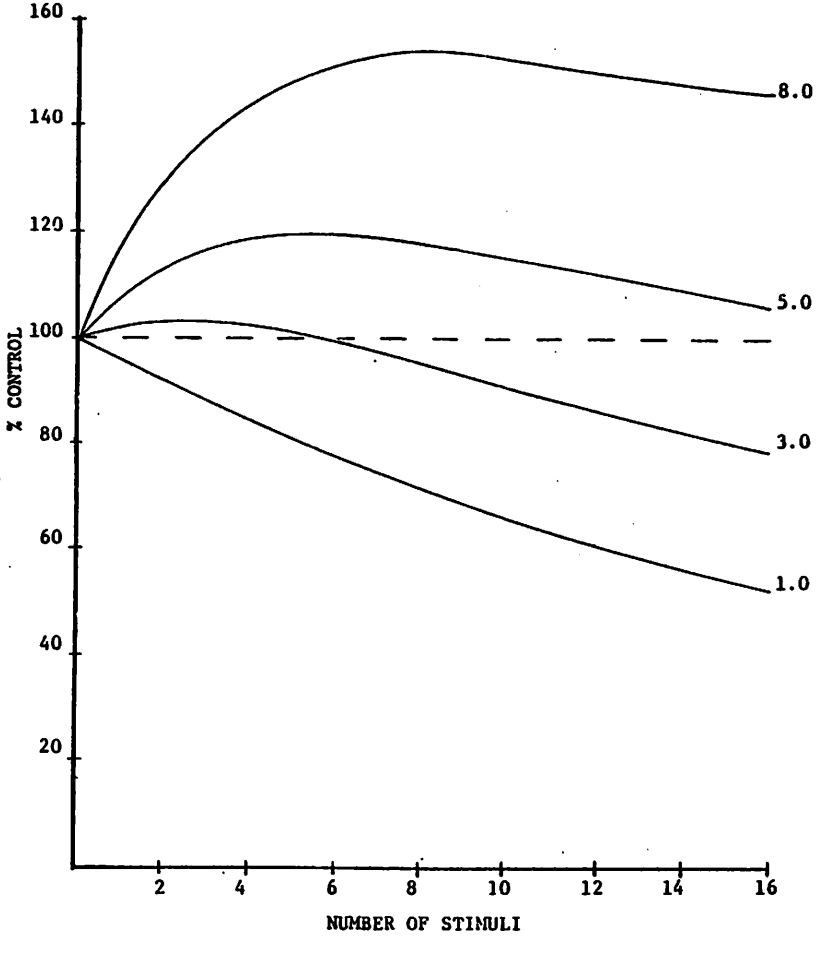


Figure 2.9 Curves of relative response generated with HABIT. Stimulus intensity marked on each curve. Associated parameters are given in Table 2.1.

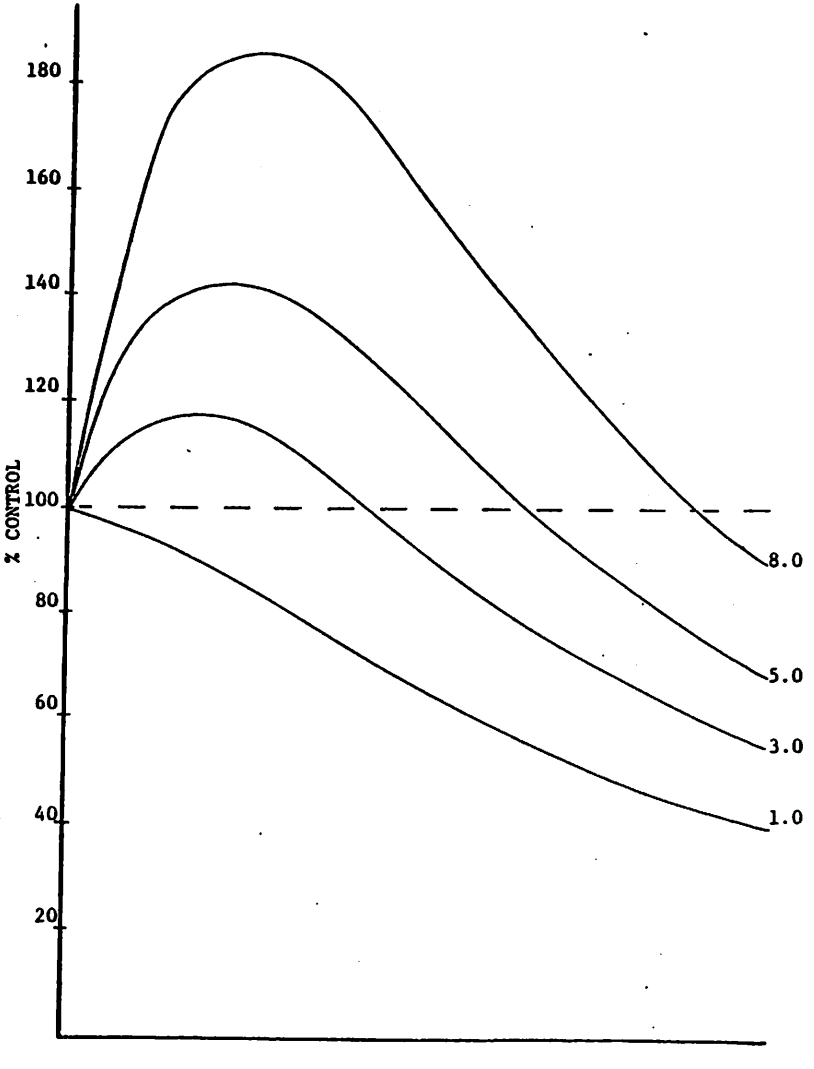


Figure 2.10 Curves of relative response generated with HABIT. Stimulus intensity marked on each curve. Associated parameters are given in Table 2.1.

the two processes develop in parallel as Groves and Thompson (1970) suggest. The serial combination of Strategy II gives rise to curves of the wrong form.

The two strategies may be mixed to good advantage, however, as illustrated below. If Strategy II is added to a system that employs Strategy I by allowing a small decrement in W_1 with a long time constant, the tails of the relative habituation curves are pulled down. In that case, greater peak values may be obtained with a larger W_3 modification gain, while asymptotic levels are kept closer to the control level. This is illustrated in Figure 2.11, where W_1 variation is added to the system of Figure 2.7. Note that the curves all drop to lower final levels, but that the final levels are not as compressed across intensities as those of Figure 2.10, where the drop in W_1 was significant. In a similar fashion, Strategy I may be added to a Strategy II system to allow greater decreases at low intensities due to the decrease in W_2. This is shown in Figure 2.12 for the case in which the SR and state paths have the same initial strengths. Once again, the asymptotic levels do not differ as greatly across intensity.

Either strategy may be used, then, in the simple network of Figure 2.6 to generate curves of the proper shapes, depending on which pathway through the system dominates the response. When the direct SR path dominates, Strategy I is needed to allow the response to fall to low levels with repeated stimulation. Strategy II is required when the state system is foremost in generating the re-

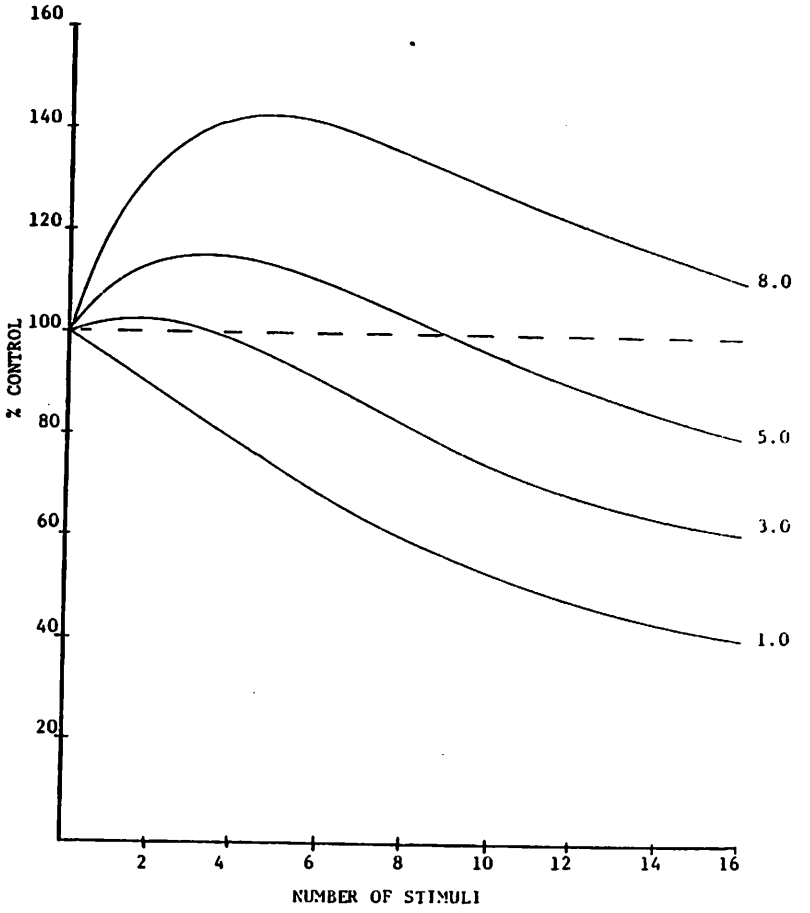


Figure 2.11 Curves of relative response generated with HABIT. Stimulus intensity marked on each curve. Associated parameters are given in Table 2.1.

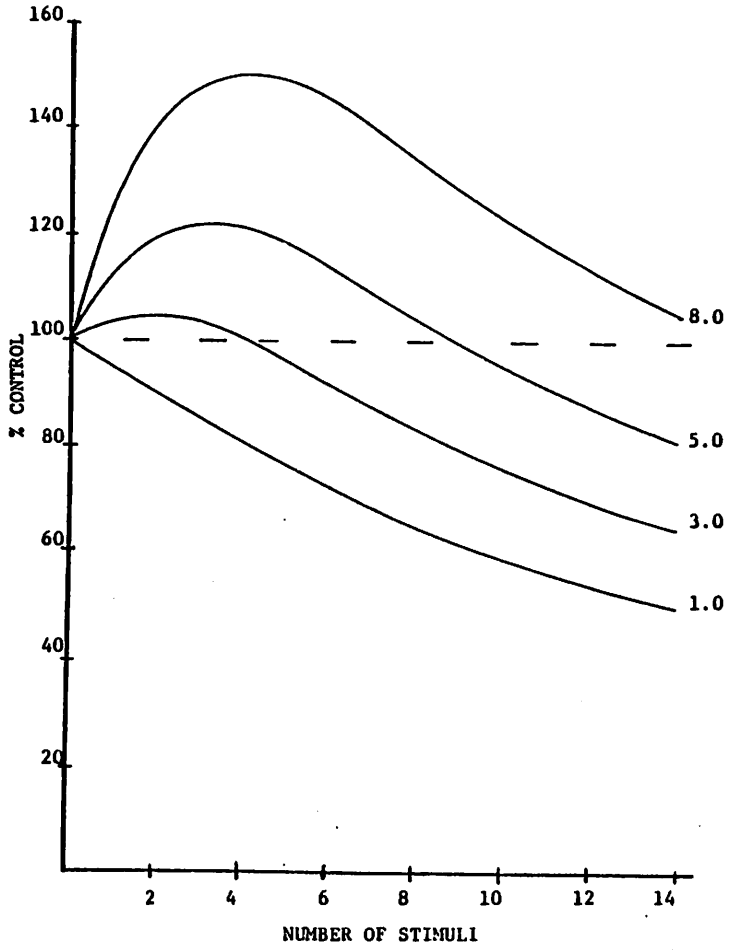


Figure 2.12 Curves of relative response generated with HABIT. Stimulus intensity marked on each curve. Associated parameters are given in Table 2.1.

sponse. The simulations illustrated here have shown that the two strategies can also be mixed to allow variation in all three of the weights.

The important results to come out of the investigations of the HABIT system concern habituation strategies and relevant parameter ranges. The circuit of Figure 2.6 is simple enough that two strategies could be readily found for which the desired curves are generated. These strategies show how the SR and the state pathways can be used in habituation. Simulations have shown what parameter values are necessary within each strategy in order to generate the proper relative habituation curves. In Strategy I, the incrementing weight, W_3, must rise more rapidly than the decrementing weight falls; so $\tau_3 < \tau_2$. In Strategy II, also, $\tau_3 < \tau_1$. In order to get curves that are near asymptote after about ten stimuli--in this case, after about ten units of time--the time constant of the decreasing synapse must have a value equal to three or four units of time. Processes with time constants having these relations should be sought experimentally.

The weight values, too, are fixed by the model. Strategy I applies when the unhabituated transmission through the SR channel is greater than that through the state channel. In this case, the model indicates that intensity curves of the proper sort can be generated using sensitization that only increases across trials. When the state path has a significant effect on the control-level response, relative habituation is less at low stimulus intensities. In this case, Strategy II may be necessary to pull the response down in later

trials. The model thus illustrates when decreasing sensitization is of use in fitting relative response curves. The assumption of decreasing sensitization that is part of the TPT can be checked in a given preparation if the state and SR channels can be independently stimulated to determine the contribution of each to the output.

Two further uses of this modeling can be suggested here. First, the simple model presented above can be extended by relaxing the assumptions made initially concerning multiple components of decrement or sensitization. Using the knowledge gained through explorations with HABIT, the full circuit of Groves and Thompson (1970) could profitably be explored. It should be possible, for example, to determine what minimal anatomy--that is, how many serial or parallel decremental and incremental components and how many cross-couplings between state and SR systems--is needed to fit a set of relative habituation curves with a prescribed level of accuracy. This minimal circuit will then give some indication of what circuit to expect in the preparation in question. It may be found that some species require much simpler model circuits than others, clearly indicating differences in their underlying response circuitry.

The notion of curve fitting to data from particular preparations indicates the second use of the model. An on-line computer system could use experimental data as it is taken and could generate the set of model parameters that best fit that data. These sets of parameters could then be used to "track" a preparation during the course of an experiment to tell whether its responses were changing materially. This could make it easier to tell when a preparation had become--or was becoming--nonphysiological. Individual preparations could be more easily compared in terms of their model parameters, as well. Through comparisons of model parameters, the use of different drugs or surgical procedures could be more easily compared in terms of their effects on the systems that underlie the response. While modeling has been used in this way in Freeman's work on cortical average evoked potentials (Freeman, 1975), it has seldom been used elsewhere. The models discussed here offer the experimenter in habituation the use of this simple but powerful tool.

Simulation, Part II: UATION. The system considered above is based on the assumptions that the synaptic modification functions are all linear and that continuous stimulation (or stimulation with interstimulus intervals considerably less than the smallest time constant of the system) is employed. These assumptions enabled a straightforward solution of the synaptic dynamic equations to be found and programmed. The response of the linear habituation network to pulsed inputs can also be calculated, as shown in Appendix A. The model investigated above could therefore be extended to handle noncontinuous stimulation. It is not always possible, however, to derive analytical solutions for the system response when nonlinear weight modification functions of any complexity are employed. In order to continue investigation of the properties of the habituation network, I implemented the UATION program for numerical integration of the synaptic modification equations.

The UATION program is described in detail in Appendix B. This program allows a variety of simple habituation networks to be simulated. As in HABIT, cell dynamics are not considered here, so the output of each cell is a function only of its current input. Synaptic dynamics can be more complicated now due to the numerical integration procedure. A pulsed input can also be simulated with UATION, so frequency effects based on recovery of the synaptic weights may be studied.

In each of the simulations discussed below, the network was started with all synaptic weights set to their resting values. A pulsed input of constant intensity, duration and period was presented. During the times the stimulus was nonzero, the network's output was calculated. The synaptic values were calculated using their differential equations throughout the duration of the simulation. This process is illustrated in Figure 2.13. The entire simulation corresponds to one habituation session. Each period of time during which the input is nonzero is denoted a stimulus presentation.

In order to construct habituation curves, the network's response to a stimulus presentation must be defined. Since the output changes during each presentation as the weights change, the initial or maximum response levels do not adequately capture the operation of the system. In what follows, the response at each stimulus presentation is defined as the total output generated by the network during that presentation. This amounts to counting the total number of spikes emitted by a neuron following presentation of a stimulus, or taking

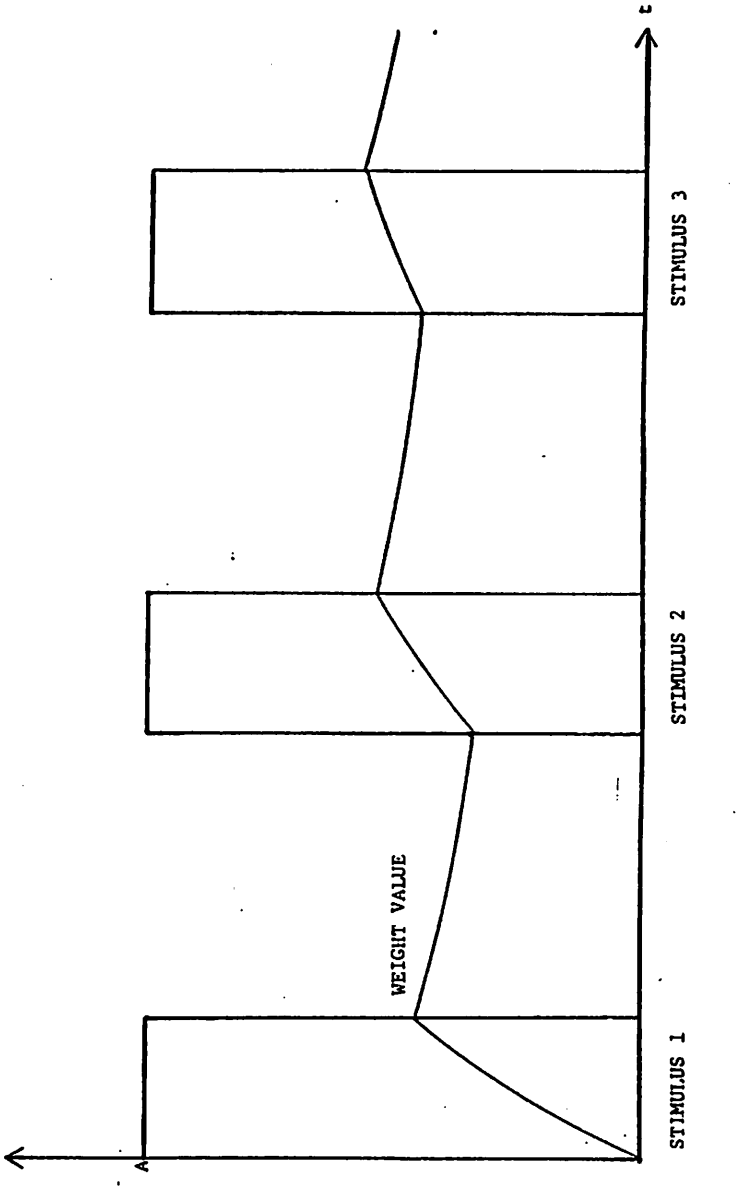


Figure 2.13 Stimulus conditions and representative resulting weight variation in the UATION program. The weight increases when the stimulus is on, recovers between stimuli.

the area under the potential evoked by an input. The system's control response must also be defined. The control response level is to correspond to the response given by an unhabituated system, as discussed in the introduction. This level may be defined in two ways in the simulation. In the first way--called the frozen control paradigm--the response is calculated with all weights held at their initial (resting) levels. This control level is truly that given by an unhabituated system. In the second scheme, the control level is that given in response to the first habituation stimulus, and includes the effects of that stimulus on the synaptic weights. Similar results can be gained using either paradigm, though parameter values differ in each case. In the simulations discussed below, the frozen control scheme was used. The parameters used in these simulations are shown in Table 2.2.

Figure 2.14 shows the result of a simulation by UATION of the same system investigated with HABIT. All weights employ linear modification functions, though in this case a pulsed input with period and duration indicated in Table 2.2 is employed. Figure 2.14 was generated by a Strategy I system corresponding to that of Figure 2.7 simulated by HABIT. The similarity between the two figures confirms the use of UATION in investigating these habituation systems.

The main reason for constructing UATION was to be able to explore nonlinear weight modification functions. Thompson (personal communication) notes that the transmission through the state path should be modified as a sigmoidal function of stimulus intensity.

TABLE 2.2

Parameter Values Used With UATION

Figure #	W_1	W_2	W_3	τ_1	τ_2	τ_3	a_1	a_2	a_3	SIGMOID+ PARAMETERS C	SIGMOID+ PARAMETERS D	STIM.* DURATION	STIM.* PERIOD
2.14	1	1	.1	12	9	4	0	-6.5	1.5			.6	2
2.16	1	1	.1	12	9	4	0	-6.5	12.0	44	.84	.6	2
2.17	1	1	.1	12	9	4	0	-6.5	17.0	376	1.31	.6	2
2.18	1	1	.1	12	9	4	-.5	-6.5	2.0			.6	2
2.19	1	1	.1	12	9	4	-.5	-6.5	17.0	376	1.31	.6	2

+ - Where applicable; see text.

* - Stimulus period and duration expressed in terms of same time units as weight time constants.

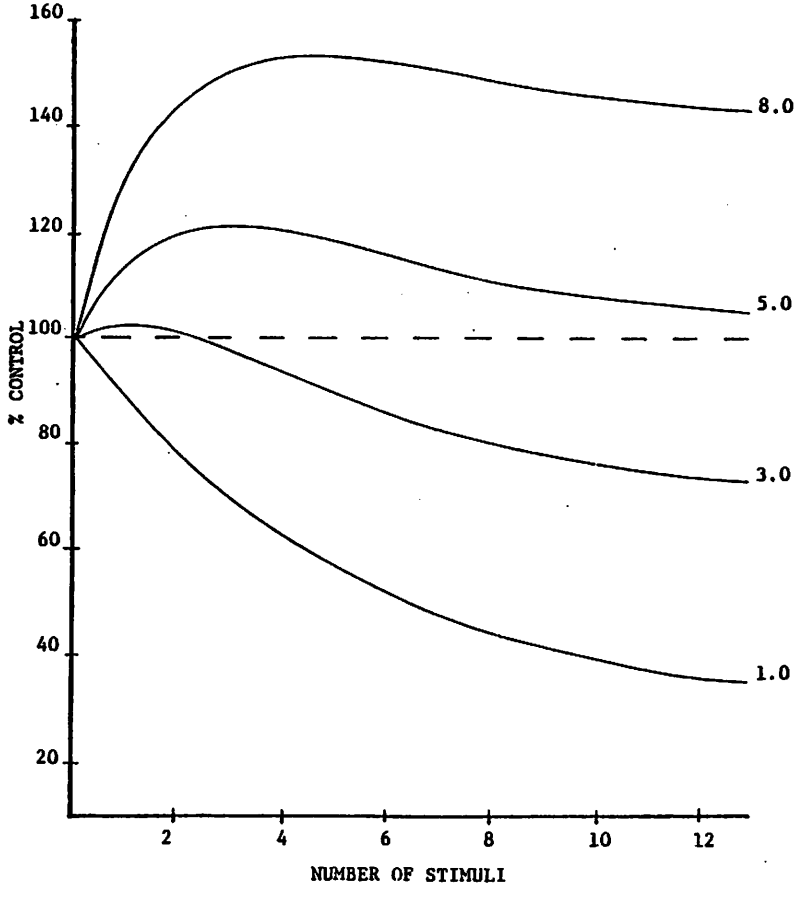


Figure 2.14 Relative habituation curves generated by UATION in Strategy I system.

This function is illustrated in Figure 2.15, and given analytically by

$$f(S) = \frac{1}{1 + ce^{-ds}}$$

This function may be used to alter the value of W_3, making that weight's differential equation:

$$T_3 \dot{W}_3 = W_{30} - W_3 + a_3 \cdot \frac{1}{1 + ce^{-ds}}$$

Figure 2.16 shows the result obtained in a Strategy I system when the differential equation above for W_3 is used. At the lowest stimulus intensity, the effective modification gain is very small, and the decrease in W_2 dominates the response (compare with Figure 2.14 above). At an intermediate intensity, the overall modification gain is larger, leading to reduced final drop in the curve. At the highest intensity, there is a greater relative drop in response than in the linear case due to the saturation of the sigmoid curve. The sigmoid's main effects, then, are to increase the relative decrements at low and high intensities and to decrease the decrement at moderate intensities when compared with the linear case.

In Figure 2.17, a sharper sigmoid is used together with an increase in the modification gain of W_3. Because the sigmoid value is very small at the lowest intensity, the response curve there falls rapidly to a low value. The responses at moderate and high intensities are enhanced, however, and the increased W_3 gain leads to more prominent peaks. A sharper sigmoid, then, can be used to generate

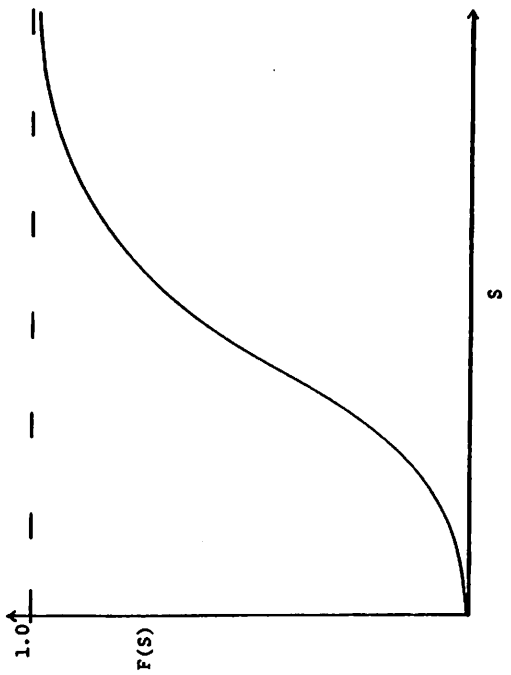


Figure 2.15 Sigmoid function employed in sensitizing weight modification.

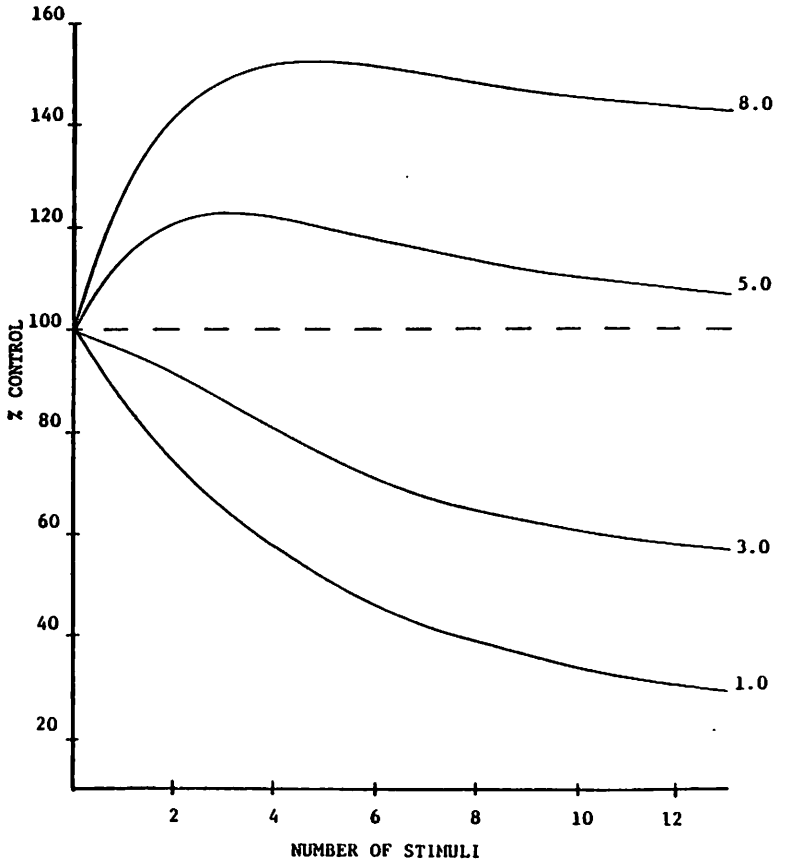


Figure 2.16 Relative habituation curves generated with sigmoid function employed as modification gain of sensitizing weight.

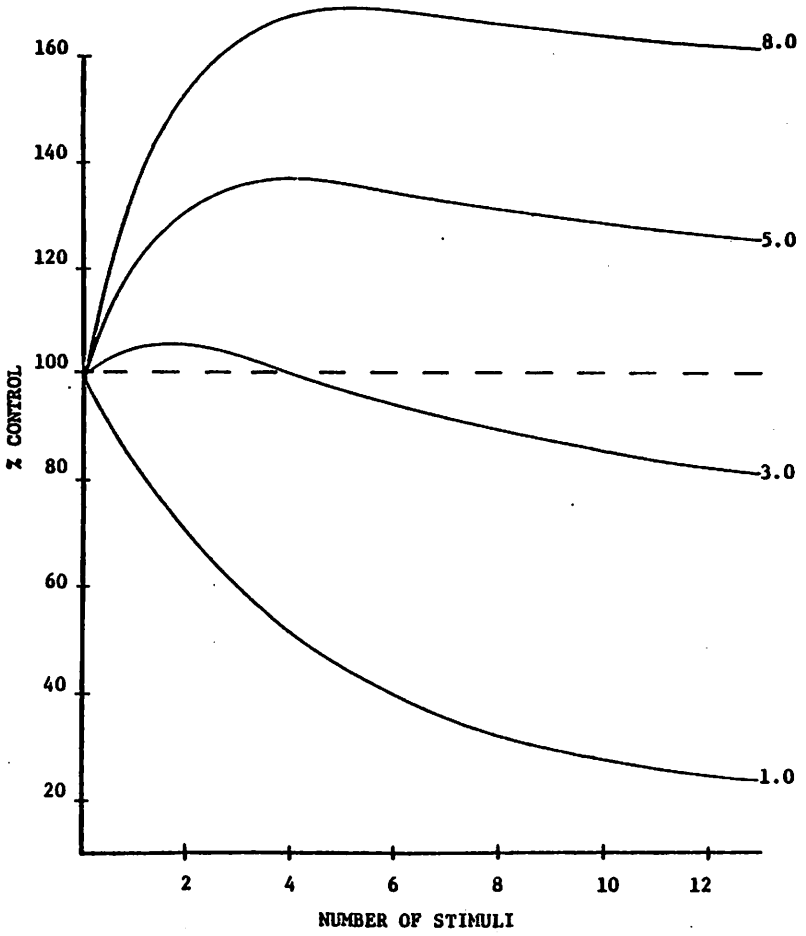


Figure 2.17 As 2.16, but with sharper sigmoid curve and increased W_3 modification gain.

considerable response depression at low intensities, and greater peaks at higher intensities.

Strategies may be mixed here, as in the systems simulated using HABIT. As in the HABIT systems, addition of a small variation in W_1 to a Strategy I system gives greater relative decrement in later trials, while allowing greater peak responses in early trials. This effect is illustrated in Figure 2.18, for comparison with Figure 2.14. In this case, the final response levels differ somewhat less across intensities due to the decrease in W_1. Figure 2.19 shows the effect of mixing Strategy I and Strategy II when a sharp sigmoid is employed in the modification of W_3. The curves for moderate intensities are severely depressed by the drop in W_1 as reflected through the sigmoid (compare with Figure 2.17).

These results show that the simple network of Figure 2.6 can give rise to the desired effects of stimulus intensity when non-constant stimulation is employed. The two programs discussed above give some indication of the ways the proper curves can be generated and what parameter values give rise to those curves. The model can also demonstrate the property of recovery from habituation and the effects of dishabituation, as discussed below.

To test recovery and dishabituation, the system is first habituated to a low intensity stimulus, and then a novel (usually strong) stimulus is applied in the rhythm of the habituating stimulus train. The habituating stimulus is then reapplied as a test input to deter-

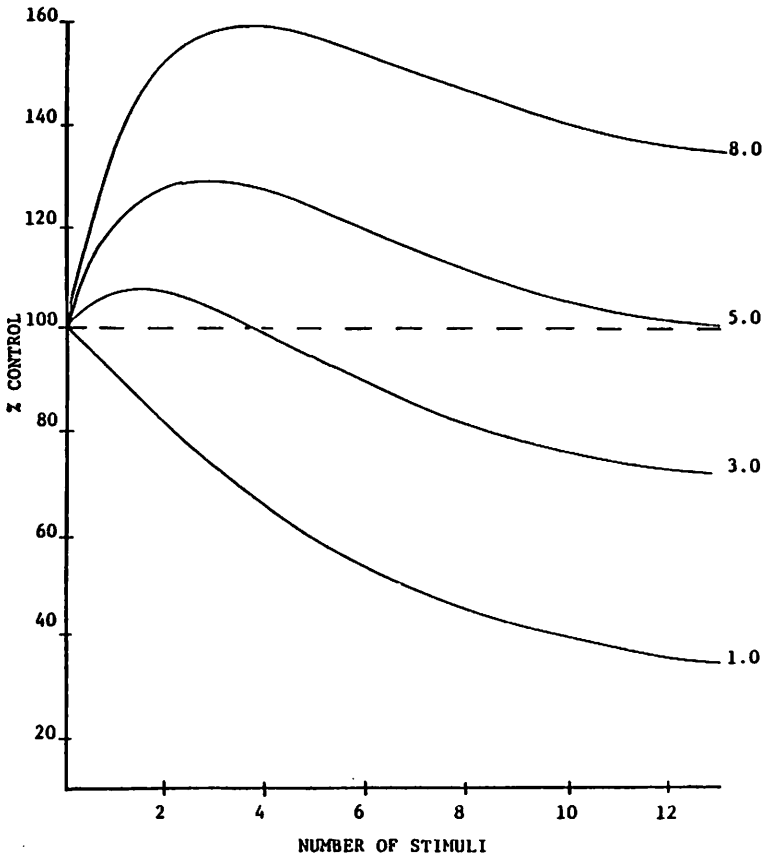


Figure 2.18 Effect of W_1 depression in Strategy I system. Parameters as in Figure 2.14, but with increased W_3 modification gain.

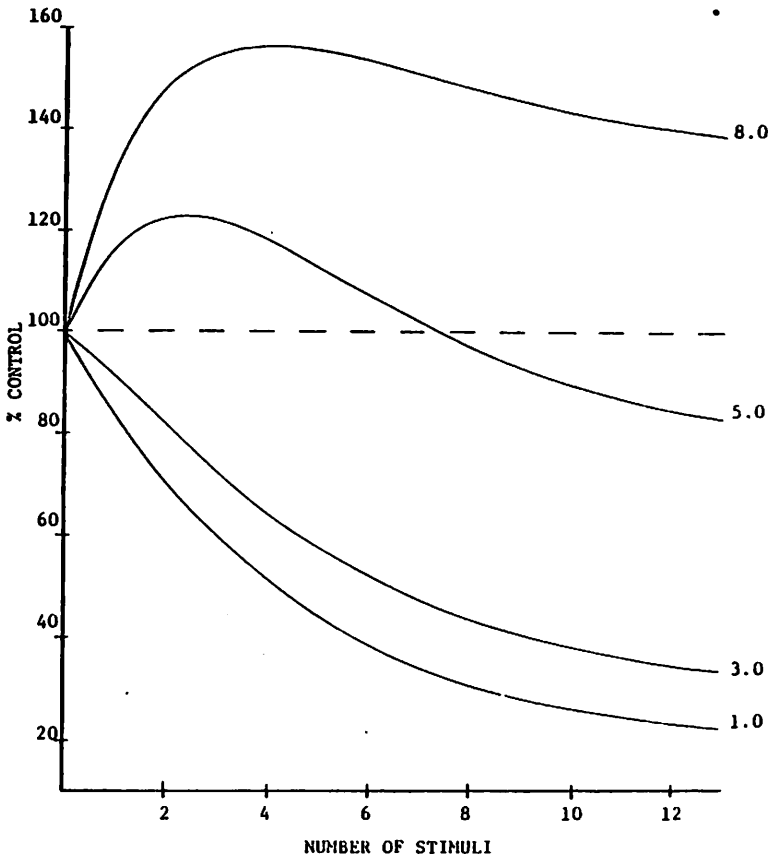


Figure 2.19 Effect of W_1 depression on system of Figure 2.17.

mine how the response has been altered by the dishabituating stimulus. Following brief retesting with the habituating stimulus, the system is allowed to rest for a time and is then retested to determine how much it has recovered from habituation training. Figure 1.2 shows the result of performing this experiment on the cat spinal cord.

As discussed previously, dishabituation in Thompson's theory is due to heightened transmission through the state pathway following presentation of a novel stimulus. I will consider here the case in which dishabituation appears following the presentation of a strong stimulus. The cases in which dishabituation arises from presentation of a weaker stimulus or a stimulus of a different type will be touched on in the section on generalization effects. Figures 2.20 through 2.22 illustrate the results of performing dishabituation experiments on the systems of Figures 2.14, 2.16 and 2.18, respectively. In each case, a curve results that is similar to that of Figure 1.2. Note that, due to saturation, the percentage of increase following dishabituation is less when the sigmoid weight modification is employed (Figure 2.21). The increase is reduced still further in the mixed-strategy system, following the decrease of W_1 (Figure 2.22). As will be described later, this effect can form the basis of the phenomenon of habituation of dishabituation.

Results of recovery tests performed on each of the systems considered above appear on the figures illustrating dishabituation. In each case, recovery to approximately 90% of the control level occurs

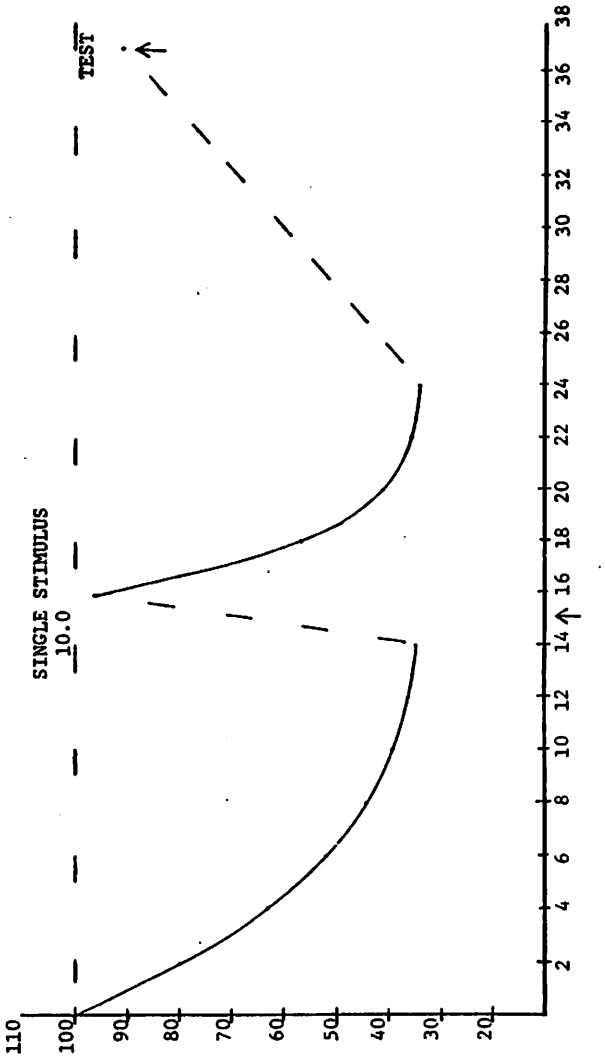


Figure 2.20 Dishabituation and recovery in system of Figure 2.14.

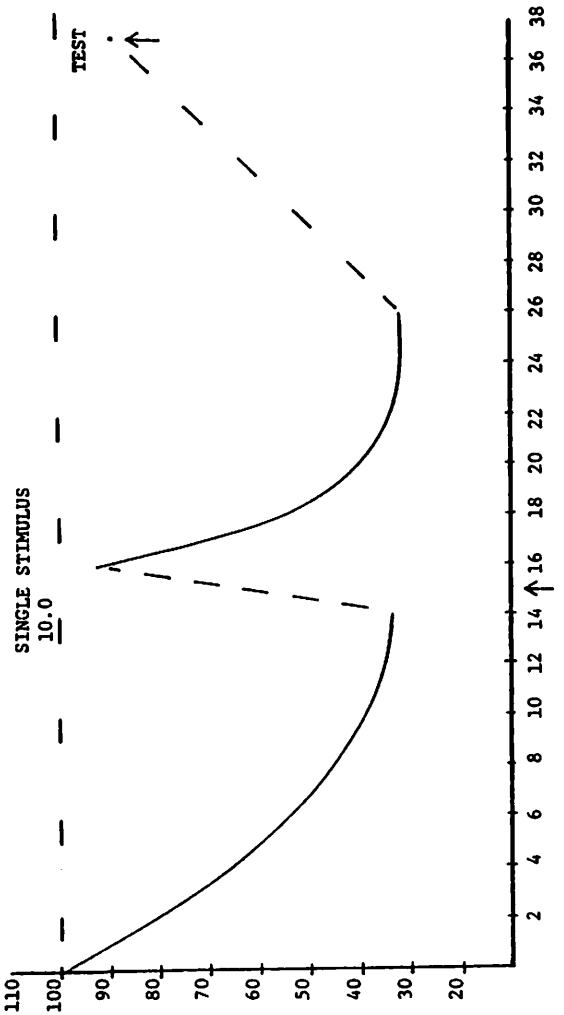


Figure 2.21 Dishabituation and recovery in system of Figure 2.16.

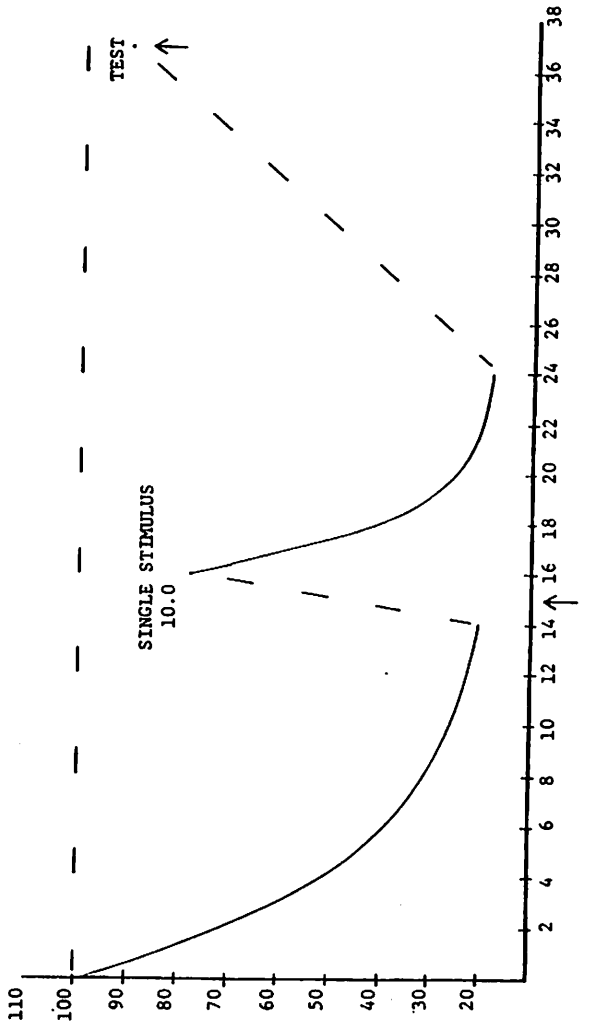


Figure 2.22 Dishabituation and recovery in system of Figure 2.18.

after about ten stimulus periods. This time course of recovery reflects the fact mentioned earlier that in order to obtain habituation curves that vary in the proper way with the number of stimuli, time constants of the decreasing weights must equal several periods of the stimulus.

Note that these figures were not derived from experiments that fairly tested the recovery properties of the system. In a fair test, the circuit would be habituated and dishabituated, and a single test stimulus (identical with the habituation stimulus) would be applied at some time following dishabituation. A number of circuits (or the same circuit at different times) can be subjected to this treatment using different recovery-time intervals between dishabituation and testing, and the results brought together in a single graph illustrating the recovery properties of the system. In the figures above, the system was given several habituating stimuli following dishabituation. The result of this train of stimuli was to keep the decreasing weight at asymptote while the value of the sensitized weight fell. Therefore, the curves of Figures 2.20 through 2.22 illustrate the decay of sensitization, rather than the full recovery properties of the system.

Figure 2.23 shows the result of the full recovery test procedure described above. System parameters are given in the figure. Following habituation to asymptote with twenty low-intensity stimuli, a strong dishabituating stimulus is given. A single test stimulus identical with the habituating stimulus is then applied at various inter-

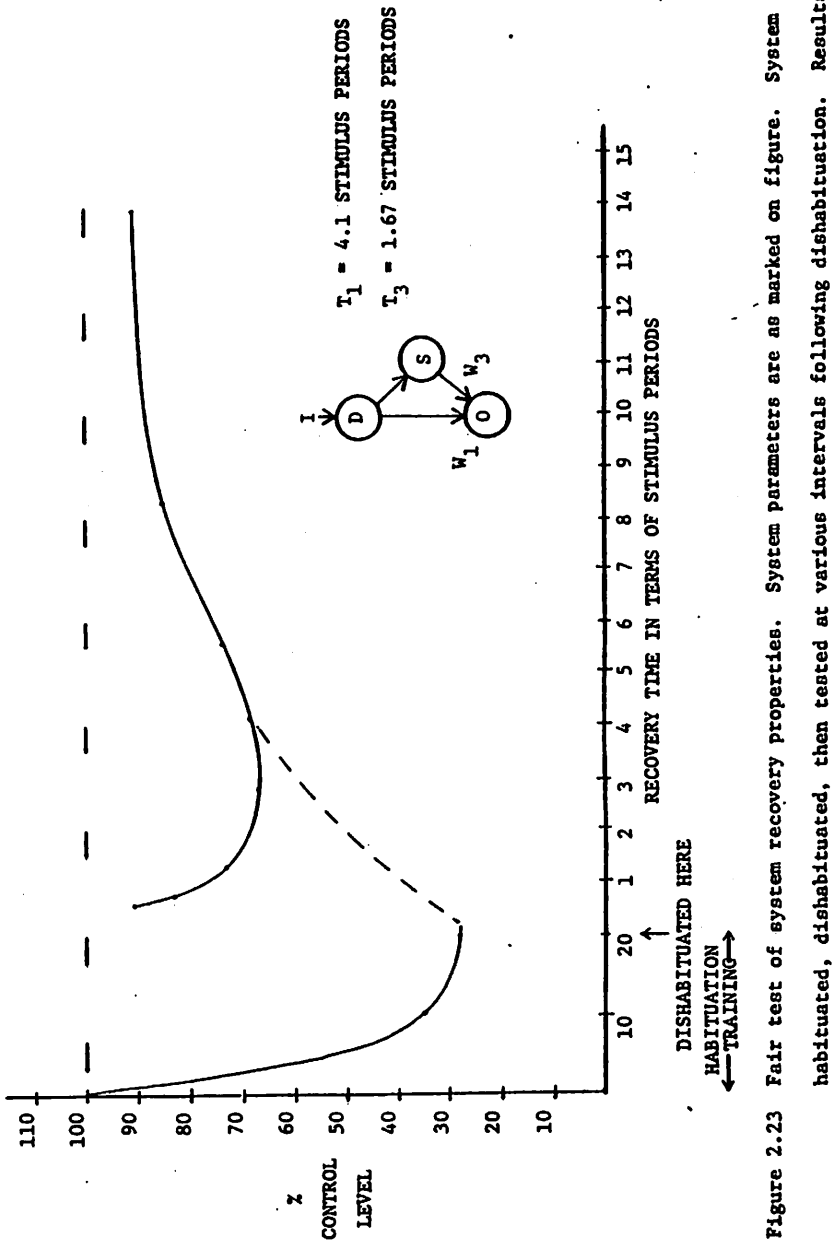


Figure 2.23 Fair test of system recovery properties. System parameters are as marked on figure. System habituated, dishabituated, then tested at various intervals following dishabituation. Results

vals following dishabituation, with the results shown on the graph. The response increases shortly after dishabituation due to the increase of the sensitized weight, then decreases as sensitization decays. At the same time, the habituated (decreased) weight recovers as indicated by the dotted line, eventually bringing the output up to about 90% of control level. When the full recovery properties of the system are tested, then, sensitization appears as before, but against the backdrop of recovery of habituation. This is quite similar to the result obtained by Spencer and coworkers (Spencer, et. al., 1966, their Figure 6). Note that if the time constant of the habituating weight were made very large with respect to that of the sensitizing weight, then little recovery from habituation would have occurred while sensitization was apparent. In that case, Figure 2.23 would have looked more like Figures 2.20 through 2.22.

In this model, recovery between stimuli accounts for the effects of stimulus frequency. Since the weights are modeled with differential equations, the final level of habituation depends strongly on interstimulus interval in a range of intervals approximately equal to the weight time constants. This effect is explored in Appendix A, where results are given for a linear system subjected to a pulsed input. Stimuli with periods in the range of the system time constants or greater give rise to a frequency effect in which shorter period stimuli yield greater habituation than longer period stimuli. Over a range of frequencies with interstimulus intervals that are short with respect to the weight time constants, however, little recovery occurs between stimuli, and habituation is approximately independent of stimulus frequency. Thus both the frequency effect alluded to in the operational definition of habituation and that given in Thompson's more detailed specification of habituation can be realized by the model.

The effect of stimulus frequency is shown in Figure 2.24 for the system of Figure 2.14. Stimulus period is indicated on each curve in terms of the time constant of the decrementing weight. A low intensity stimulus is employed so that only the frequency properties of the decrement are displayed by the output. It may be seen that the higher frequencies lead to more pronounced habituation. For this range of frequencies, then, the systems illustrated display characteristic (4) of the operational definition of habituation. At still higher stimulus frequencies, there is little recovery between stimulus presentations and the response is approximately independent of frequency.

Point (e) of the more detailed specification of the properties of central decrement and sensitization states that sensitization is directly dependent on stimulus frequency at higher intensities. In the model presented here, sensitization has a smaller time constant than central decrement (Compare the time constants of W_3 and W_2 in Tables 2.1 and 2.2.). Thus at stimulus rates for which decrement is weakly dependent on frequency, sensitization is more strongly dependent, and the model is able to satisfy points (d') and (e) simultaneously.

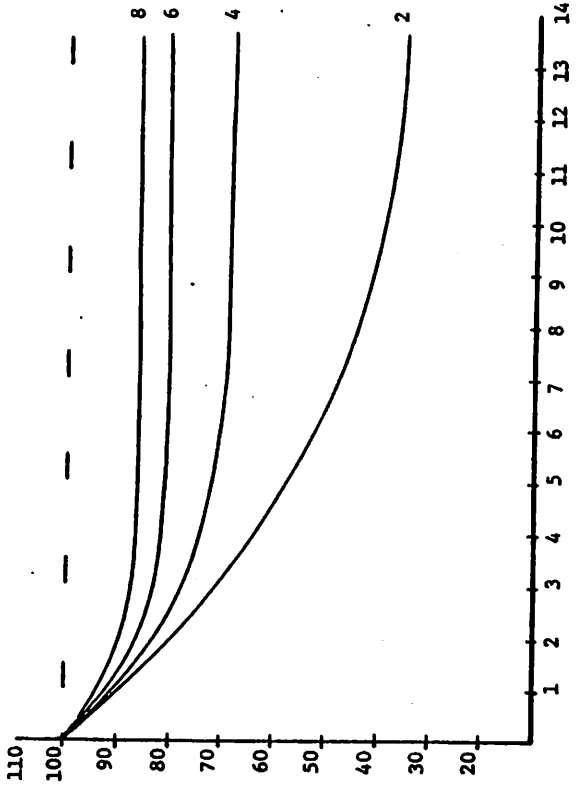


Figure 2.24 Effect of stimulus frequency on the system of Figure 3.16. A low-intensity stimulus was employed with different interstimulus intervals as marked on the curves.

The UATION program, then, has verified the results of the HABIT program and extended those results to the case of a nonlinear weight modification function and to noncontinuous stimuli. It was seen how the sigmoid-shaped modification function for sensitization altered the habituation curves and allowed greater peak values and asymptotes closer to the control level. Simple dishabituation and recovery tests were attempted and were found to follow experimental data. Finally, the effects of stimulus frequency were seen to depend on the weight time constants relative to the interstimulus interval. Both frequency dependence and independence could be obtained with the proper parameter settings.

These two sets of computer simulations have given detailed support to the intuitive arguments presented earlier that the basic habituation network operating according to Thompson's definitions of central decrement and sensitization can realize the desired frequency and intensity effects. In so doing, parameter ranges have been found within which the model circuit behaves properly. As discussed in Chapter I, these ranges of values can be compared with measurable parameters of physiological systems to help verify the TPT. Sets of parameters that yield the best fit between experimentally generated habituation curves and those generated by the model can also serve as the basis for comparing two different preparations. In this way, the model serves to relate experimental data more directly to the underlying mechanisms of habituation.

One further use of the computer simulations comes in the ability to tailor synaptic modification functions to fit habituation data. The sigmoid curve employed above in driving the sensitizing weight has three parameters that determine its shape. I have shown how the shape of the sigmoid affects the habituation curves themselves. In fitting data, the sigmoid parameters would be adjusted along with weight values and time constants. The computer program in fact allows the use of more complex functions; any such function would also be adjusted to fit the response to the desired level of accuracy. This allows the possibility of finding the best function or family of functions to fit the data. The shape of that function might help to tell the experimenter more about the circuit or modification rule underlying the data. Again, the function shape and parameter values also offer quantitative bases for comparing data from different preparations, or data taken from the same preparation at different times.

Further Habituation Characteristics

The modeling of the previous sections has dealt with the characteristics of recovery and the frequency and intensity effects of habituation. In this section I show how the other characteristics of habituation can be realized within the framework of the present circuit model. It will be shown how the circuit can be extended to realize the properties of generalization of habituation and dishabituation. The synaptic modification rules themselves will be altered in simple ways to obtain the below-zero and long-term effects. Together, these changes allow the quantitative expression of all of the major habituation characteristics within a single model.

Part I: Generalization of habituation and dishabituation. So far, I have dealt with the effects of a single stimulus applied to a single habituation unit. Experiments with natural systems show that the habituation produced by one stimulus may generalize to other stimuli. Likewise, the response to one stimulus may be dishabituated following presentation of a different stimulus. In Thompson's theory, generalization effects are explained through the use of the common elements concept introduced earlier. The habituation unit of Figure 2.6 can easily be extended to model generalization effects according to the common elements theory. This extension will be discussed following presentation of experimental data on generalization effects, and discussion of other circuit models of generalization.

Generalization of habituation and dishabituation are illustrated in Figure 2.25, a and b. In Figure 2.25a the response to stimulus 1 is habituated, as indicated. During the course of this habituation, test stimuli are given to channel 2. Habituation is said to generalize from stimulus 1 to stimulus 2 to the extent that the test responses to stimulus 2 are decreased by habituation to stimulus 1. In Figure 2.25b, a strong stimulus applied to channel 2 results in dishabituation of a previously habituated response to stimulus 1. Again, dishabituation exhibits generalization to the extent that the response to stimulus 1 is affected by presentation of stimulus 2.

Figure 2.25a Generalization of habituation. Habituation stimulus is applied to input 1, as indicated by the solid bar on the graph for input 1. Test stimuli, indicated by individual impulses, are given at the same time to input 2. It is assumed here that these test stimuli are presented at such a rate that they have no modifying effect on the system. The output due to stimulus 1 declines. The output due to stimulus 2 also declines as the result of generalization of habituation.

Figure 2.25b Generalization of dishabituation. Habituation stimulus is applied to input 1, as in Figure 2.25a. Midway through the habituation training, a strong stimulus is given to input 2, as indicated by the bar on the stimulus 2 graph. The output due to stimulus 1 is temporarily increased, illustrating generalization of dishabituation.

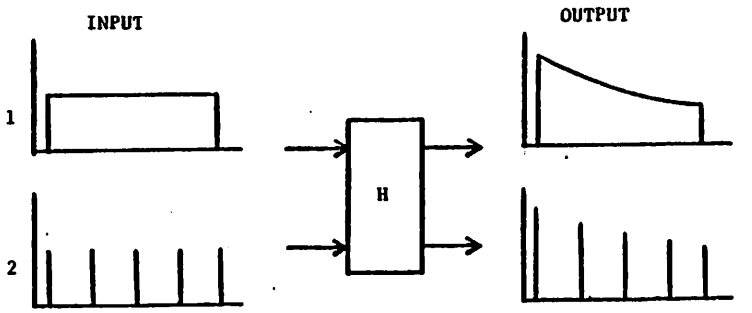


Figure 2.25a

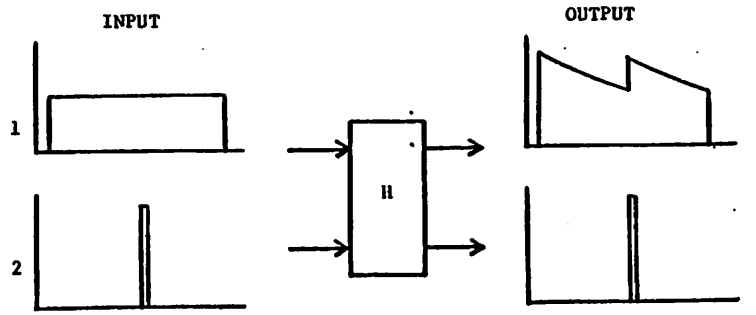


Figure 2.25b

Such generalization effects have been demonstrated in a variety of preparations. Horn (1974) shows generalization of habituation in a rabbit tectal unit. In the isolated frog spinal cord, the response recorded in the ventral root due to electrical stimulation of the dorsal root exhibits habituation that generalizes across segments of the cord (Farel and Thompson, 1972). Similar generalization of habituation has been found in the response of the cat spinal cord to stimulation of two separate branches of the same afferent nerve (Wickelgren, 1967, a, a; Thompson and Spencer, 1966). Examples of generalization of habituation to auditory stimuli of different frequencies are found in cat cochlear nucleus response (Buchwald and Humphrey, 1973), in human GSR (Graham, 1973) and in the orienting response as measured by the duration of alpha rhythm depression following presentation of a tone (Sokolov, 1961). In each case, a curve similar to that in Figure 2.26 is found. Following habituation to stimulus S_o, the response to similar stimuli is depressed. Note that generalization exhibits a pronounced gradient, such that stimuli more different from the habituating stimulus 1 are less affected. Pakula and Sokolov (1973) cite an example of habituation in snails that generalizes within a single modality (light), but does not generalize across modalities.

It should be noted that in some systems, no generalization of habituation has been found. These systems usually involve a monosynaptic pathway from stimulus to response. This is the case in the gill withdrawal response of Aplysia (Kandel, et. al., 1970), in which sensory cells make direct, habituating contacts with motor

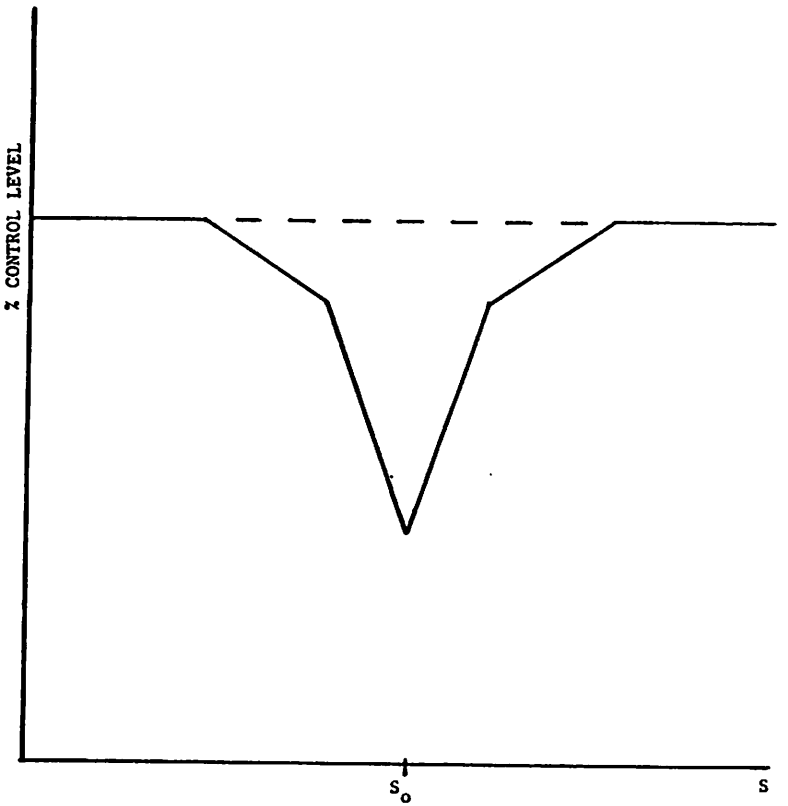


Figure 2.26 Gradient of habituation generalization. Ordinate gives values of a continuous stimulus parameter, abscissa shows absolute response levels before (dotted line) and after (solid line) habituation to stimulus S_o. After habituation, responses to stimuli near S_o are depressed. Those sufficiently far removed are unaffected. (After Sokolov, 1961).

neurons. Habituation of the response elicited by tactile stimulation of a portion of the gill served by one sensory neuron does not generalize to areas served by different sensory cells. Similarly, the monosynaptic path from the lateral column to ventral root of frog spinal cord does not exhibit generalization of habituation (Farel, et. al., 1973). Presumably, in each of these systems, there is no overlap between elements activated by different stimuli, hence there is no transfer of habituation (Farel, et. al., 1973).

Generalization of dishabituation also occurs in a variety of preparations. Horn (1974) shows a unit in locust tritocerebrum that is dishabituated by a shock applied to an afferent nerve. Buchwald and Humphrey (1973) indicate that the cat cochlear nucleus response habituated by an 800 HZ stimulus could be dishabituated by presentation of a 300 HZ tone. Farel and Thompson (1972) find dishabituation across segments of the frog spinal cord, in the dorsal root to ventral root response. Similar dishabituation is seen in the response recorded in ventral roots of the cat spinal cord to stimulation of different afferent nerves (Wickelgren, 1967 b; Thompson and Spencer, 1966). In Aplysia (Kandel, et. al., 1970), the habituated gill withdrawal response returns following a strong stimulus applied to the animal's head.

As discussed previously, habituation may be realized in neural circuits by a build-up of inhibition or by a depression of synaptic transmission. These two mechanisms lead to different circuit realizations of the generalization effects. Figure 2.27a shows a circuit

due to Wickelgren (1967 b) that generalizes habituation by means of inhibition. In this circuit, the inhibitory path through interneuron 1 is potentiated by repeated stimulation. The potentiation could be due to a decrease of the cell's threshold or to an increased effectiveness of its efferent synapses. In either case, the response to stimulus 2 will be decreased below the control level following habituation to stimulus 1. Wickelgren (1967 b) shows that this circuit may be cast in terms of presynaptic inhibition, postsynaptic inhibition acting at the somas of cells 1 and 2, or postsynaptic inhibition acting at selected cites on the dendritic trees of cells 1 and 2.

The common elements concept is illustrated by the circuit of Figure 2.27b. There, habituation is the result of decrement of the synapses from cells 1, 2, and 3 to the output cell. Cell 3 is common to the pathways from both stimuli to the output. Following habituation to stimulus 1, the response to stimulus 2 is below its control level due to the decrease in cell 3's efferent synapse. A variant of this circuit is given by Horn (1970, 1967), who shows that generalization of habituation may be due to activation of overlapping sets of afferents.

Following my previous analysis of systems that involve synaptic decrement as the basis of habituation, I will consider generalization models based on this mechanism. Computer models based on other mechanisms--such as the inhibitory circuits of Wickelgren (1967 b)--could also be constructed. As discussed in Chapter I, such modeling would allow detailed comparison of the workings of different habituation

Figure 2.27 Circuits illustrating generalization of habituation.

a Generalizaiton via inhibition (after Wickelgren, 1967 b). Circles represent cells, solid dots are excitatory synapses, short vertical lines are inhibitory synapses. Inhibitory synapses grow stonger with use. Generalization occurs due to the fact that each afferent fires the inhibitory cell (cell 3) that affects each transmission cell (cells 1 and 2).

b Generalization via the common element concept. Synapses from transmission cells (cells 1, 2 and 3) to output cell (cell 4) decrease with use; other synapses are nonplastic. Stimulation of either afferent excites cell 3 causing its synapse on cell 4 to decrease. That decrease is common to both afferents, then, effecting generalization.

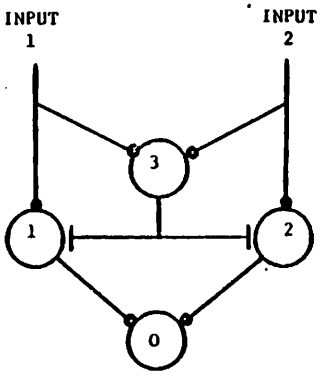


Figure 2.27a

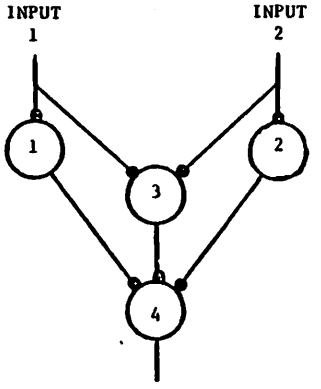


Figure 2.27b

mechanisms. First note that the lack of generalization in some monosynaptic systems is easily explained with the common elements notion, as mentioned above. Gradients of generalization can also be explained in these models through the use of gradients of overlap between nearby stimulus-response channels, as illustrated in Figure 2.28. There, the weights of synapses on cells in rank 2 from a given cell in rank 1 decrease with the distance from that cell. For example, the connection from cell 1 to cell (e) is stronger than that from cell 1 to cell (h). Suppose that only the synapses from cells of rank 2 to the output cell, R, decrease with use. Following habituation to stimulus 1, the synapses to the output cell from cells (c) through (g) are decreased. A test stimulus applied to cell 2 will yield a considerably reduced output due to the overlap of paths from cells 1 and 2 to the output. Cell 3 shares fewer paths with cell 1, however, so the response due to stimulus 3 will be less affected by habituation to stimulus 1.

The basic unit of Figure 2.6 may be used in a similar way in circuits that exhibit generalizaiton of habituation and dishabituation. Figure 2.29 shows one such circuit. This network employs an input layer each of whose cells excites several habituation units as in Figure 2.28. Here, stimulation of input cell 3 leads to depletion of the paths from input 1 to cell O_1, leading to generalizaiton of habituation. Generalization of dishabituation is also realized here. For example, stimulation of cell D_1 by input cell 3 results in activation of cell S_1, with a consequent increase in the synapse from S_1 to O_1. Application of a strong stimulus to input cell 3 will there-

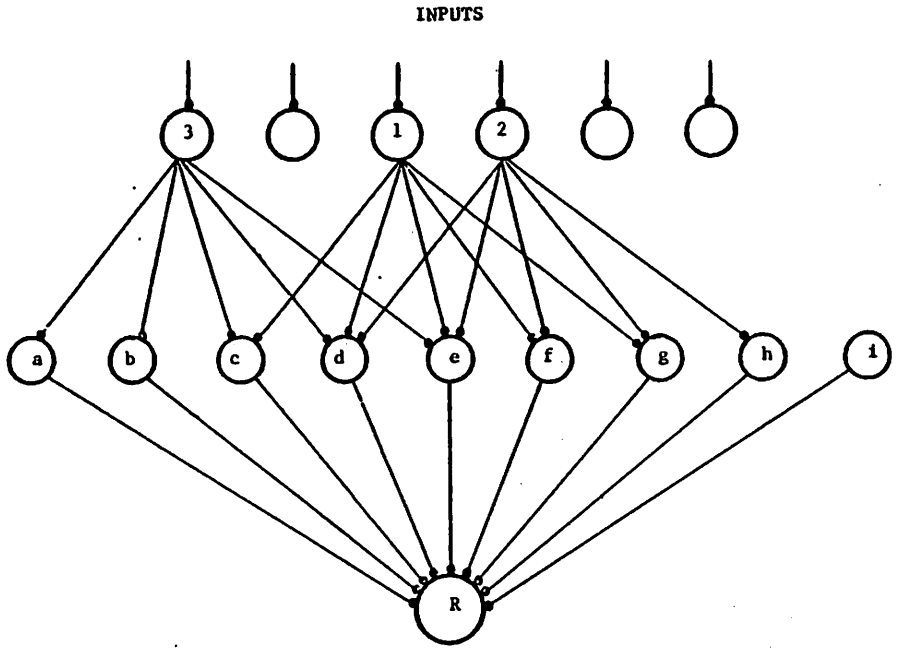


Figure 2.28 Circuit illustrating gradient of habituation generalization. Synapses from cells of rank 2 to output cell (cell R) decrease with use. Habituation generalizes as the result of overlap of stimulated synapses (after Horn, 1967).

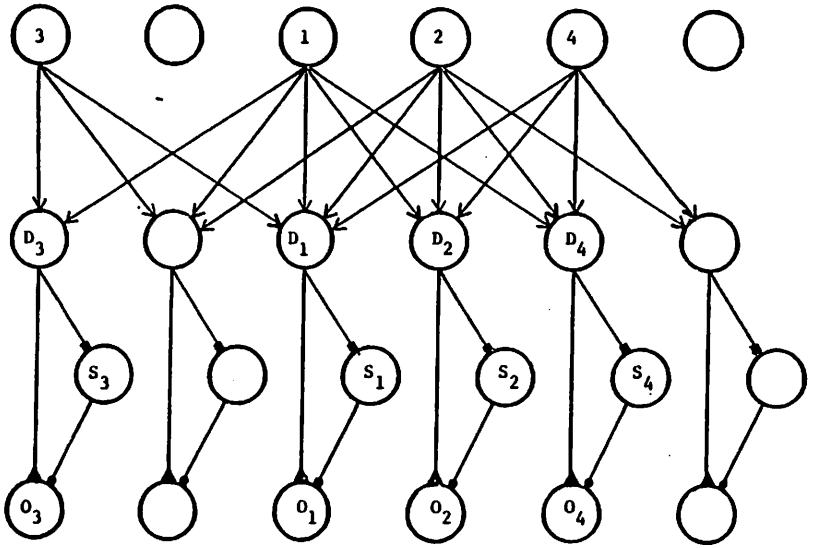


Figure 2.29 Generalization circuit based on the unit of Figure 2.6 and the scheme of Figure 2.28. Synapses are as denoted in Figure 2.6. Synapses marked with arrows are nonplastic.

fore restore the response of O_1 to input 1 after habituation to input 1. This form of dishabituation will itself habituate if the synapse from D_1 to S_1 is an habituating synapse. Repeated dishabituation will cause that synapse to deplete and reduce the effect of input 3 on S_1, as was illustrated in Figure 2.22.

A further effect may be gained through cross-coupling the D and S layers as shown in Figure 2.30. Now each D-cell excites neighboring S-cells and output cells, and each S-cell excites neighboring output cells. Activation of an input causes both increases and decreases in the paths from the input to its own output. In the case of input cells with considerable path overlap, transmission decreases may predominate, leading to generalization of habituation with a gradient depending on the overlap as before. For inputs with less overlap, however, the connections shown in Figure 2.30 result primarily in transmission increases. Activation of input 3, for example, leads to an increase in the connection from S_2 to O_4, and no decrease in any path from input cell 4 to O_4. Following habituation to input 3, then, the response to input 4 will be increased rather than decreased. Such a phenomenon is reported by Graham (1973) in studies of the habituation of human GSR. Note that this effect relies on the fact that the final common path between input 3 and input 4 is a sensitizing synapse from S_2 to O_4. If the basic unit were such that the synapses from layer D to layer S were sensitizing and those from S to O were decrementing, this effect would not occur. The basic unit was structured as it is partly for this reason.

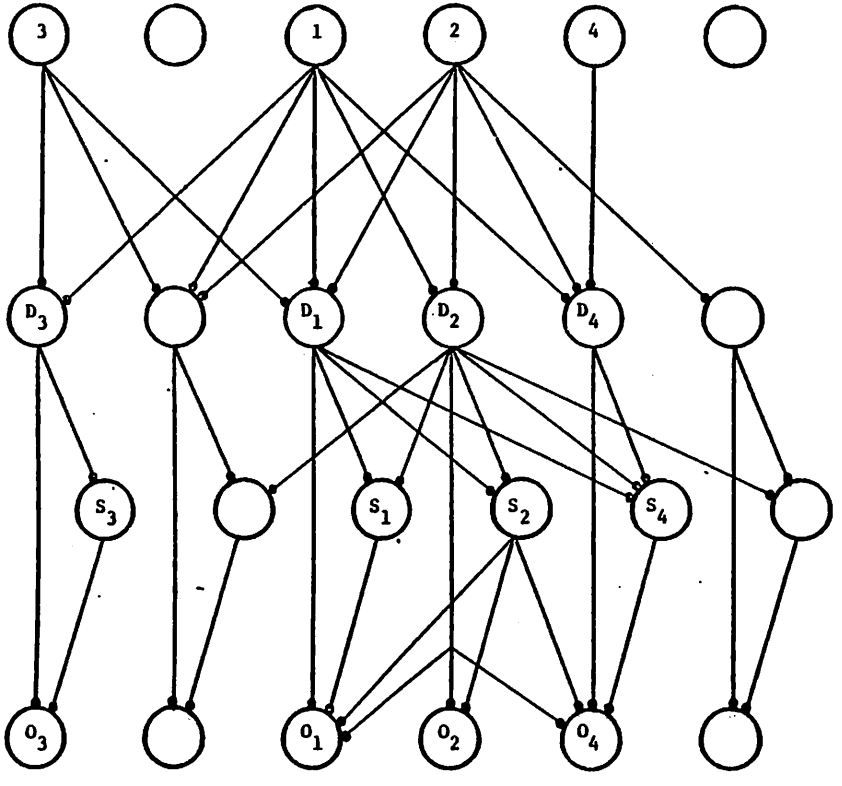


Figure 2.30 Further generalization effects due to cross-coupling of basic units. Synapses are denoted with dots here, but modification properties are assumed to be the same as those of Figure 2.29 with the exception that synapses from D-rank to cells of S-rank decrease with use. Full set of connections is shown only for unit 2, all other units are coupled in the same fashion.

It is clear then that this model, using the basic unit of Figure 2.6 can duplicate some of the data on generalization of habituation and dishabituation. The data on dishabituation of a response due to presentation of the habituating stimulus at a decreased intensity can be explained easily within this framework (Groves and Thompson, 1970). If it is assumed that a change in stimulus intensity causes a shift in the active stimulus channels, then the model above applies directly, where different input cells code different stimulus intensityies. A similar mechanism has been proposed by Sokolov (1975) as the basis of intensity effects. As noted previously, in the TPT the "neural model" of the stimulus--including its intensity--is the set of synapses depleted by its repeated presentation, and different stimuli are compared to this model in terms of the overlap between the sets of synapses they activate and the set of depleted synapses.

The basic habituation unit in the generalization network of Figure 2.30 operates under the rule derived for intensity-independent relative habituation; viz., depleting synapses are decreased independent of stimulus intensity. Thompson and co-workers (1973) indicate that in order to deal with the full range of intensity-generalization data, a synapse is needed whose value decreases as a function of stimulus intensity. As discussed previously, such a rule cannot be employed in the basic unit of Figure 2.6 if relative habituation is to be independent of intensity. A more detailed basic model is required, then, to fit Thompson's theory. That model will be presented following a brief discussion fo two final habituation effects.

Part II: The below-zero effect. In some preparations, recovery from habituation is slowed if presentation of the habituating stimulus is continued after the asymptotic response level has been reached. This is known as the below-zero effect (BZE). When it occurs, the BZE is generally pronounced enough to be detected without recourse to statistical methods. The effect does not occur in all preparations, however, so it should be discussed in the light of data from a particular preparation. Following a discussion of the ways the BZE appears in physiological systems, I will sketch some general approaches to modeling the BZE that may be applied in particular cases.

The below-zero effect is displayed by physiological systems in three ways. First, a response may fail entirely during habituation training and recover more slowly following prolonged presentation of the habituating stimulus. This is the classical below-zero effect, as first described by Prosser and Hunter (1936) in the startle response of intact rats. In the second case, the response reaches a non-zero level and remains at that level with further stimulus presentations (Farel and Thompson, 1972; Farel, et. al., 1973). Whatever change is involved in the BZE in this case does not alter the asymptotic level itself. Finally, in the third form of the BZE, the response level slowly declines with continued habituation training, and increases more slowly upon recovery (Farel, et. al., 1973). That one system (the ventral root response to electrical stimulation of the frog lateral column) displays both the second and third forms of the BZE makes it seem a subsidiary effect that may depend heavily on the particular preparation studied. Further, some systems, such

as the gill withdrawal response of Aplysia (Kandel, et. al., 1970) do not display the BZE at all.

The BZE can nonetheless be added to the model of Figure 2.6 in the following way. The equation employed above for the depleting synapse is

$$\tau \dot{W} = W_0 - W - a \hat{I},$$

where $\hat{I}$ is determined by

$$\hat{I} = \begin{cases} 1; & \text{stimulus is non-zero} \\ 0; & \text{otherwise} \end{cases}$$

Slower recovery may imply a larger weight time-constant following prolonged stimulation. The BZE may be realized by making the weight time-constant a variable governed by:

$$N \dot{\tau} = \tau_0 - \tau + B \hat{I}$$

where

N is the weight time-constant,

τ_0 is the time-constant initial and resting value,

B is the modification gain of the time constant,

and

$\hat{I}$ is defined as above.

As the synapse is stimulated, then, its time-constant increases. The asymptote to which the synapse is driven, and therefore the asymp-

tote to which the output falls, is independent of the change in the synaptic time-constant. If τ_0 is much less than N, and if B is not too large, then the synapse will fall to its asymptote before the time constant becomes too large, and the behavior of the system during initial stimulation will be essentially as it was with a fixed time-constant. With prolonged stimulation, however, the time-constant will continue to increase and a greater period of recovery will be needed to bring the weight back to its control level. This model is applicable in the case that the response reaches a steady asymptote, whether the asymptote is zero or not.

In the case of a system that shows the BZE with a slowly falling asymptote, two modeling approaches may be taken. In the first approach, the system may be modeled as above with changing weight time-constants, but with the addition of a second SR path. The second path in this model has a smaller initial transmission than the first and a larger time-constant. In initial habituation trials, the first path depletes, accounting for most of the habituation of the response. As habituation training continues, however, the second path slowly depletes, resulting in a small, slow decline in the response level. This second path will recover as slowly. This slow recovery can lead to a pronounced BZE without the use of a variable weight time-constant. The possibility of two paths with different time-constants was mentioned earlier in conjunciton with sensitization and with recovery in Aplysia. The utility of this notion suggests the investigation of models with populations of synapses having different modification characteristics. The beginnings of such modeling are discussed in

the next chapter. For now, it is enough to note that a form of the BZE may be modeled using two paths with different time constants.

The second approach to the BZE in systems with nonsteady asymptotes involves changes in the weight resting level. If W_0 is decreased by stimulation and recovers slowly with respect to the weight time-constant, then an effect similar to that above can be achieved. The form of habituation and recovery expected from these two models is illustrated in Figure 2.31. Again, this is similar to the data on recovery in Aplysia.

This model of the below-zero effect would allow an exploration of the overall characteristics of the effect and its interactions with the other habituation effects. This formulation does not shed much light on the underlying circuit or synaptic mechanisms that may be responsible for the BZE. As indicated above, more detailed models of this effect will be discussed following presentation of the population model of habituation in the next chapter.

Part III: The long-term effect. As described briefly in Chapter I, many systems display a savings of previous habituation training. Successive episodes of habituation followed by recovery result in more pronounced or more rapid habituation. This long-term effect (LTE) of habituation training is demonstrated in Figure 2.32 taken from Farel and Thompson's (1972) work on the ventral root response to stimulation of dorsal root of frog spinal cord. It may be seen that following habituation and recovery, further habituation training

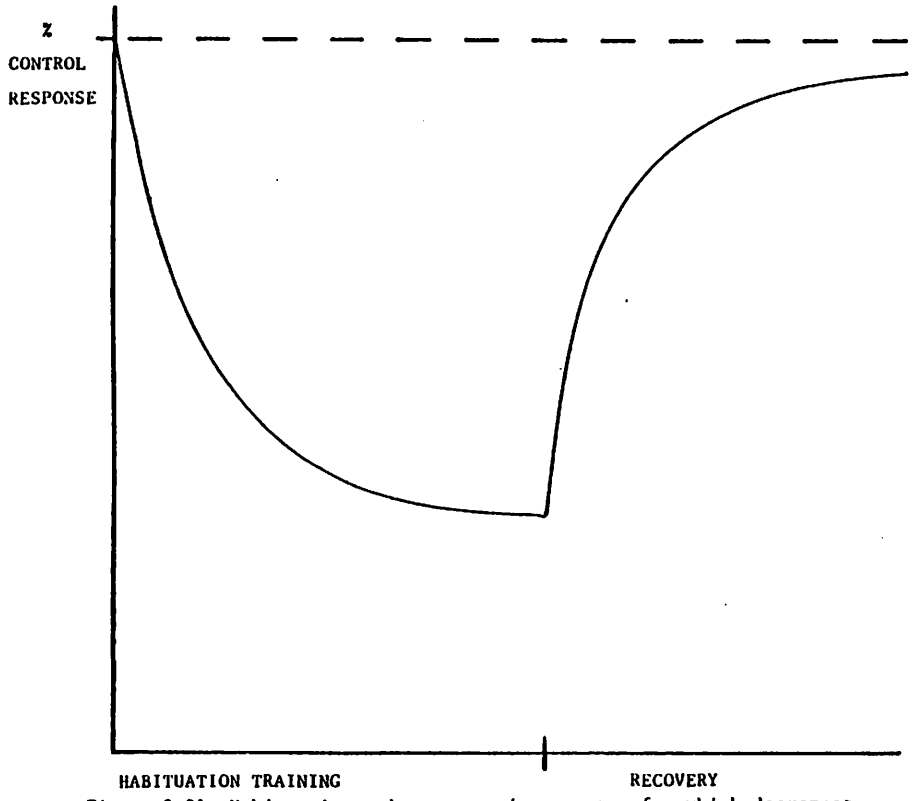


Figure 2.31 Habituation and recovery in a system for which decrement displays two time constants. Abscissa gives response level, ordinate gives time during habituation training and recovery. Fast component leads to rapid drop of response during habituation and rapid recovery to a level below 100%. Slower component gives slow drop in asumptote during habituation and slow recovery toward 100% response level.

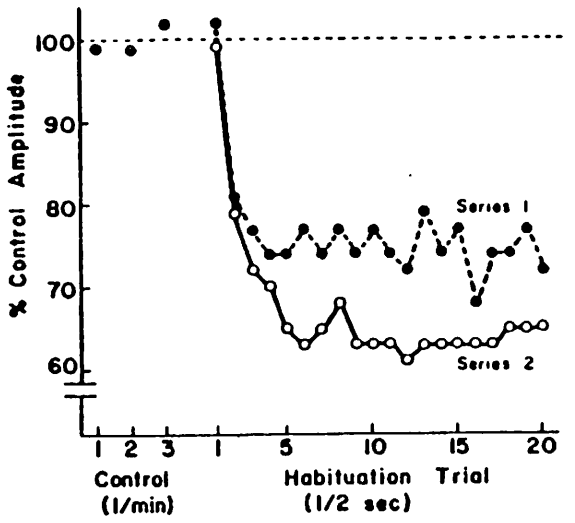


Figure 2.32 The long-term effect in the ventral root response to dorsal root stimulation in the isolated frog spinal cord (Farel and Thompson, 1972). Two habituation training sessions were employed, with a period of recovery between them. Habituation is greater in the second session than in the first. Reprinted with permission of the author and publisher.

leads to a lower response asymptote. A similar effect is shown in Farel's (1973) work on the lateral column to ventral root response in the frog spinal cord. The LTE is shown in the same form in the Aplysia gill withdrawal response (Carew, et. al., 1972), the cat flexion reflex (Thompson and Spencer, 1966), the EMG produced in the thigh muscle of the frog in response to electrical stimulation of the foot (Farel, 1971), and a wide variety of invertebrate preparations (see Wyers, et. al., 1973, and references therein). In each case, the response falls more rapidly to a lower asymptotic level.

Evidently, then, though transmission through the system has recovered, the propensity to decrement is enhanced by repeated habituation training. That this effect is not due simply to the total number of stimuli given is illustrated in the elegant experiments of Carew and co-workers (Carew, et. al., 1972) in Aplysia. In these experiments, one group of animals was given repeated sessions of habituation followed by recovery, while another group was given the same total number of stimuli in a single session. Figure 2.33 shows that the massed training led to less habituation and more rapid recovery upon retesting than the spaced training. As mentioned in Chapter I this result is demonstrated in enough preparations that Petrinovich (1973) has suggested that it be adopted as a tenth defining characteristic of habituation.

As with the BZE, I will show how the long-term effect may be added to the basic model discussed previously. As with the BZE, the addition will say little directly about the circuit or synaptic

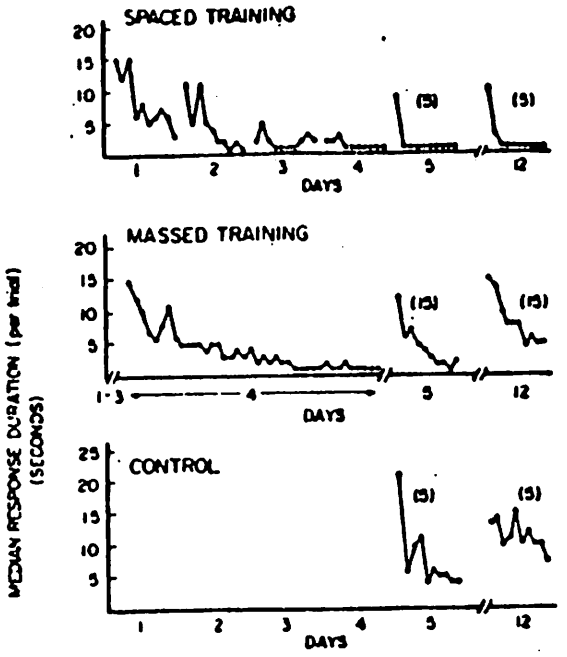


Figure 2.33 Massed verses spaced-trials effect in Aplysia (Carew, et. al., 1972). Gill-withdrawal response of Aplysia is tested under two conditions. In (a), forty stimuli are given over the course of three training sessions. In (b), forty stimuli are given in a single session. Habituation tested after recovery is more pronounced in the former case. Reprinted with permission of the author and publisher. Copyright 1972 by the American Association for the Advancement of Science.

mechanisms that may be involved in expression of the LTE, beyond the suggestion of some functional forms that the synaptic variation may follow. More detailed modeling of LTE must await the population model to be presented in the next chapter.

In the process-level model defined previously, the system's asymptotic response is determined by the asymptotic levels of the decrementing and incrementing channels. Each channel's asymptote is in turn determined by the modification gain of the associated synapse. The modification gain of the decreasing weight may itself be a variable governed by a differential equation; viz.,

$$\tau_w \dot{W} = W_0 - W - a \hat{I}$$
$$\tau_a \dot{a} = a_0 - a + V \hat{I}$$

where parameters are defined as before. If the time constant of the modification gain is much greater than that of the synaptic weight, then successive habituation sessions will leave the weight modification gain increased, even though the weight itself is allowed to recover between sessions. With each session, the gain is greater and the asymptotic weight and output levels are lower.

Though this model realizes a form of the LTE, it does not handle the effects of massed trials properly. In fact, the modification gain is increased as a function of the total number of stimuli, rather than as a function of the total number of habituation sessions. This situation can be remedied in two ways. First, a short-term fatigue effect can act on the modification gain equation to decrease its drive as stimulation is continued. Thus:

$$\tau_v \dot{V} = V_0 - V - c \hat{I}$$

If τ_v is about equal to τ_w, the change in the modification gain of the synaptic weight occurs mainly in the first few stimulus presentations in each session. As the weight itself recovers from habituation, so does the ability to increase its modification gain. Massed trials will be less effective than spaced trials, then, in producing the LTE in this system, owing to the decrease in V.

This model of LTE employs three differential equations driven entirely by presynaptic activity. Another approach is motivated by the observation that a system that is subjected to spaced-trials habituation is more active than one given massed trials. The cells in the spaced-trials experiment fire more over the course of the testing due to the periods of recovery between habituation sessions. In the massed-trials experiment, habituation occurs rapidly and cells fire less for the same number of stimuli. This observation suggests that the LTE may be due partly to postsynaptic activity of the cell driven by the habituating synapses.

Consider, for example,

$$\tau_a \dot{a} = a_0 - a V O$$

where O is the firing rate of the output cell. In the model considered here, (Figure 3.8)

$$O = W D$$

ignoring the sensitization terms, so

$$\tau_a \dot{a} = a_0 - a + V W D$$

The modification gain, and therefore the synaptic weight itself, depends on a combination of pre- and postsynaptic activity. Though this situation is remeniscent of the Hebb synapse (Hebb, 1949), the effect here is negative, rather than positive, feedback. As a increases, it forces W to decrease and reduces its own rate of increase. If τ_a is large with respect to τ_w, the LTE and massed-trials effect will again be exhibited by the model.

These two models for the long-term effect can be differentiated experimentally if a way can be found to depress the activity of the post-synaptic cell. If the cell is unable to fire, the second model says that long-term habituation will not build up; post-synaptic activity makes no difference in the first model. One means to depress the post-synaptic cell midht be to stimulate a pathway known to mediate a tonic inhibition of that cell. The cell's activity might also be affected by drugs that block synaptic transmission or that desensitize post-synaptic receptors.

The long-term and below-zero effects can therefore be built into the current model in various ways. Doing so would allow a "grand model" to be constructed in which the effects of stimulus intensity, frequency, and duration and the number of habituation sessions could

be explored. Such an analytical model would allow quantitative exploration of the habituation characteristics and of the ways they interact to produce the overall observed response decrement. The model is presented on a process level, but suggests underlying circuit mechanisms in its basic anatomy and in the forms of the synaptic modification rules employed. In the next chapter, I present a more detailed and realistic model of the circuit and synaptic effects that might underlie response decrement.

CHAPTER III

A POPULATION MODEL OF HABITUATION

The process-level model detailed in the last chapter is able to duplicate a number of the characteristics of habituation, under the assumption that synaptic decrement is independent of stimulus intensity. However, as mentioned in Chapter II, Thompson's full theory requires a synapse that decreases as a function of intensity. In this chapter, I present a more detailed habituation model that can realize intensity-independent habituation with intensity-dependent synaptic change. This model represents a closer look at the action of the SR channel in the lumped model of Chapter I. The single cell of Chapter II is now replaced with a rank of cells; the single fiber now becomes a fiber bundle. This model is still a lumped model, however, in that the fibers show no spatial preference in their contacts with the cells. The population model exploits the effects of distributions of cell and afferent fiber thresholds to realize habituation that varies inversely or not at all with stimulus intensity. The operation of the model is expressed in terms of the statistics of populations of cells and afferent fibers. Results obtained here may therefore be more easily related to circuit mechanisms operating in particular physiological systems than those of the process-level model. Following presentation of the population model, extensions to circuit-level models of the below-zero and long-term effects will be discussed.

The anatomy of the population model is shown in Figure 3.1. There, a bundle of afferent fibers makes contact at random with a

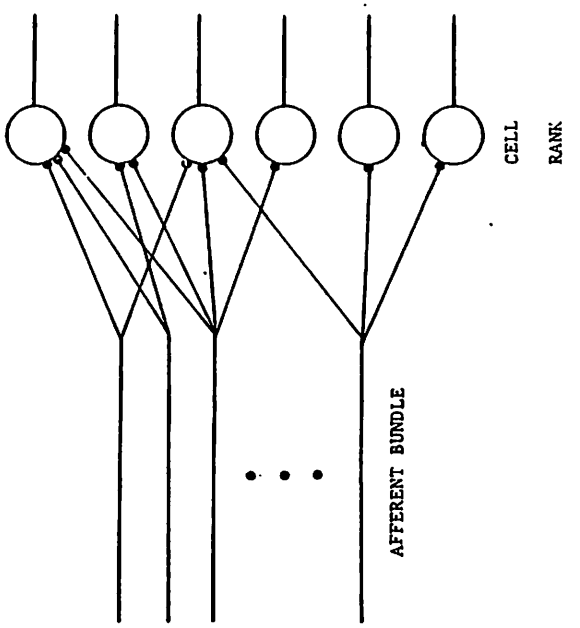


Figure 3.1 Anatomy of the population model. A bundle of afferents makes random contact with a population of cells. Synaptic values decrease with use.

rank of cells. The number of afferents excited by a system input depends on the intensity of stimulation. This activation of afferents results in postsynaptic potentials that tend to excite the cells. A given cell fires if its total excitation exceeds its threshold. The firing rate of a cell is defined as the positive difference between its total excitatory input and its threshold. The output of the system due to a given stimulus can be described in terms of the average number of active cells or the average firing rate of those cells. To realize habituation, the values of the synapses decrease with presynaptic activity.

In what follows, I first derive formulas for the probability of cell firing and for the average cell firing rate based on the distributions of the cell and efferent fiber thresholds, and on the distribution of the number of afferents each cell receives. Results of computer evaluations of these formulas and curves of relative habituation based on them are presented next. Finally, extensions of the model are discussed.

System Response Calculations

In calculating the system response, I first consider the distribution of input fibers to a given cell. I assume that each fiber makes contact with a given cell with fixed probability. Because of this random connection rule, the number of fibers making contact with a given cell is a binomially distributed random variable (Marr, 1969; Rall, 1955).

Let

$P_N(n)$ = probability that a given cell has n input fibers

The random connection rule above gives

$$P_N(n) = \binom{N_s}{n} p_c^n \ (1 - p_c)^{N_s - n}$$

where

N_s is the total number of input fibers, and

p_c is the probability of connection from a given fiber to the cell.

Afferent fibers are activated by an input as a function of stimulus intensity. It is necessary to calculate the distribution of activity in the afferent bundle as a function of intensity, then. Each afferent fiber is assumed to have a threshold with respect to a given stimulus. Thresholds vary over the population of fibers, so the activity in the afferent bundle must be expressed in terms of its threshold distribution. For example, if the stimulus is current generated by a pair of electrodes placed in or near the bundle, nearby fibers will be activated at lower stimulus intensities than fibers farther away from the electrodes. Note, also, that a fiber activated at a given intensity is also activated by higher intensity stimuli.

Let $P_{SA}(I)$ denote the fiber threshold density function induced by a particular stimulus, expressed in units of stimulus intensity. A given fiber is activated only if the stimulus intensity is greater

than its threshold. At a fixed intensity, say I^*, the probability that a given fiber is activated is the probability that its threshold is less than I^*. That is,

$$\Pr[\text{fiber is activated at intensity } I^*] = \Pr[\text{threshold} \leq P_{SA}\ I^*) = F_{SA}(I^*),$$

where $F_{SA}(I)$ is the cumulative distribution function associated with $P_{SA}(I)$. This relation between fiber activity and intensity is illustrated in Figure 3.2b for the case of the flat fiber threshold density shown in Figure 3.2a. Note that this probability is the same for all fibers, so the activity over the entire fiber bundle is a binomial random variable with

$$\Pr[\text{m fibers active at intensity I}] = P_A(m, N_s) = \binom{N_s}{m} F_{SA}(I)^m\ (1 - F_{SA}(I))^{N_s - m}$$

This distribution of active fibers is reflected in the distribution of active fibers making contact with each cell. If a cell has a total of j input fibers, then

$$\Pr[\text{cell has m active fibers}] = \binom{j}{m} F_{SA}(I)^m\ (1 - F_{SA}(I))^{j - m}$$

as above. The probability that a given cell has m active fibers, then, is

$$P_I(m, I) = \sum_{j=m}^{N_s} P_N(j) \binom{j}{m} F_{SA}(I)^m\ (1 - F_{SA}(I))^{j - m}$$

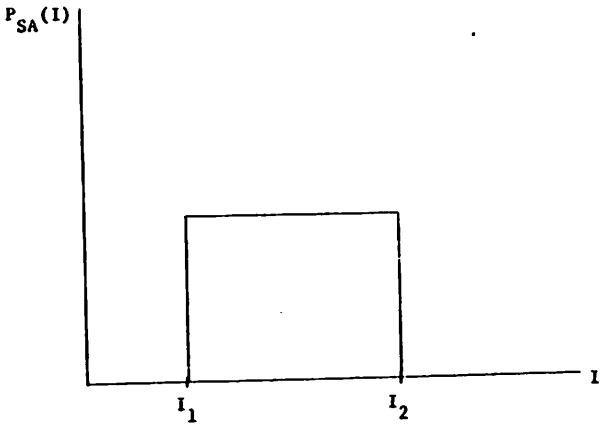


Figure 3.2a Fiber threshold distribution. Ordinate in units of stimulus intensity.

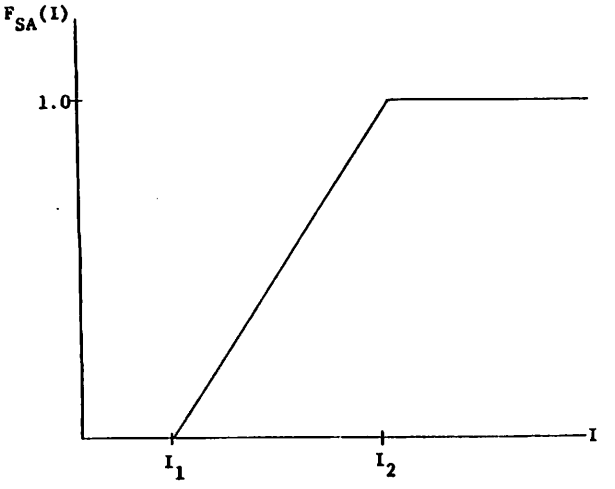


Figure 3.2b Resulting fiber activation probability distribution function.

This form takes into account all the ways that m active fibers contact cells with different numbers of input fibers.

The activation of a number of fibers on a given cell yields a postsynaptic potential depending on the strength of the activated synapses. It is generally accepted that activation of a chemical synapse by a single action potential releases a number of quanta of transmitter substance, each of which leads to a small depolarization of the postsynaptic membrane (Auerbach, 1972). The number of quanta released is a random variable, generally considered to have a Poisson distribution. In what follows, I suppress the statistical nature of this process and deal only with the mean amount of transmitter released by a synapse upon stimulation.

Each cell is assumed to have a threshold chosen from a probability distribution. The cell fires if its postsynaptic potential is greater than or equal to its threshold, and is silent otherwise. Likewise, the cell's firing rate is assumed to be the difference between its postsynaptic potential and its threshold when this difference is positive, and is zero otherwise. Let the amount of depolarization produced by activation of a given synapse (the synaptic weight) be expressed in terms of threshold units. I again ignore the statistical nature of the synapses, and suppose that all synapses have the same weight. A cell with m active input fibers experiences a postsynaptic depolarization equal to mW, where W is the common synaptic weight, and where linear summation of postsynaptic potentials is assumed.

A cell with threshold equal to A fires when

$$mW \geq A$$

So,

$$\Pr[\text{synaptic drive} \geq \text{cell threshold}] = P_E(A, W, I) = \sum_{m=0}^{N_s} P_I(m, I)\, U(mW \geq A),$$

where

$$U(x) = \begin{cases} 1; & \text{if } x \text{ is true} \\ 0; & \text{otherwise} \end{cases}$$

Note that this formula assumes a standard activation of the synapses, independent of stimulus intensity and frequency. This assumption and the assumption of linear addition of postsynaptic potentials will be discussed later.

Now

$$mW \geq A$$

when

$$m \geq [A/W]$$

where $[x]$ = least integer greater than or equal to x.

Let

$$R = [A/W],$$

then

$$P_E(A, W, I) = \sum_{m=R}^{N_S} P_I(m, I), \text{ and}$$

$$P_F(W, I) = \Pr[\text{cell fires}] = \Pr[\text{synaptic drive} \geq \text{threshold}]$$
$$= \int_0^\infty P_A(x)\, P_E(x, W, I)\, dx$$

where it is assumed that the cell thresholds are continuously distributed with density function $P_A(x)$. This integral can be approximated by

$$P_F(W, I) \approx \sum_{j=m}^{M} P_A(j \cdot \Delta A)\, P_E(j\Delta A, W, I)\, \Delta A$$

for some minimum and maximum thresholds given by $m \cdot \Delta A$ and $m \cdot \Delta A$, respectively, and summation increment ΔA.

This formula gives the probability of firing of a given cell in terms of stimulus intensity and synaptic weight. Since the cells fire independently of one another, the total output activity is a random variable that is again binomially distributed, giving

$$\Pr[k \text{ cells fire}] = \binom{N_c}{k} P_F(W, I)^k\, (1 - P_F(W, I))^{N_c - k}$$

where N_c = total number of cells.
In particular, the average number of active cells is just $N_c \cdot P_F(W, I)$, which depends directly on the probability of cell firing.

A similar calculation gives the expected output rate of a given cell. Let

$$T(x) = \begin{cases} x; & x > 0 \\ 0; & x \leq 0 \end{cases}$$

Then the output rate of a cell with synaptic drive mW and threshold A is T (mW - A). The expected output rate of a cell with threshold A is, then,

$$F_E(A, W, I) = \sum_{m=o}^{N_S} \Pr[\text{cell has m active fibers}] \cdot T(mW - A)$$
$$= \sum_{m=o}^{N_S} P_I(m, I) \cdot T(mW - A)$$

The expected output rate taken over all cell thresholds is

$$F_F(W, I) = \sum_{j=m}^{M} P_A(j \cdot \Delta A)\, F_E(j \cdot \Delta A, W, I)\, \Delta A,$$

where a sum is used to approximate the integral of the threshold density function, as before.

These formulas express the probability of cell firing and the expected output rate as functions of stimulus intensity and synaptic weight. Computer-generated curves of these quantaties are shown in Figure 3.3 for fixed synaptic weight. The probability of firing is a sigmoid function of intensity, while the output rate function grows quadratically. Similar results for the cell firing probability were derived by Rall (1955) who considered the action of this system. Rall did not deal with output rate, however. Though Rall did investigate facilitation, he did not work with synaptic depression, as is dealt with here.

In order to model habituation in this system, a rule is needed for the way the synaptic weight iecreases with repeated stimulation. As in the process-level model of Chapter II, the synaptic weight will

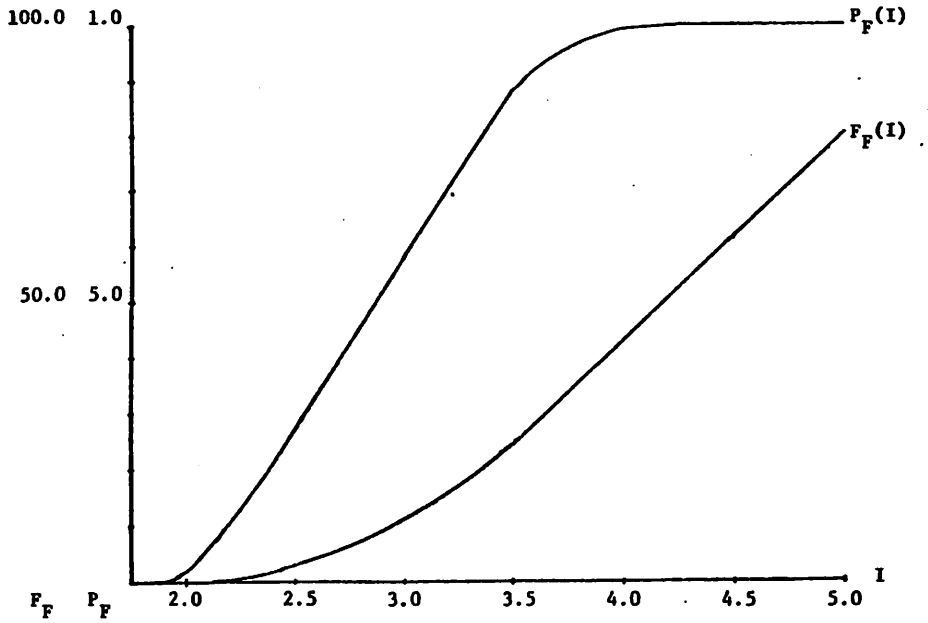


Figure 3.3 Curves of probability of cell firing and average output rate as functions of stimulus intensity.

be assumed to vary according to a first-order differential equation in order to capture the properties of exponential decrease and spontaneous recovery. Thus, the synaptic equation is:

$$\dot{W} = W_0 - W - a\,I^+$$

where

I^+ is the synaptic modification input,

and all other parameters are defined as in the process-level model discussed in Chapter II.

The synaptic modification input may be a function of stimulus intensity or may be intensity-independent. Both cases are explored in simulations detailed below. In this model, since all synapses are assumed to be identical, all of the synapses activated by a given input can be modeled with a single differential equation. Relaxation of the assumption of identical synapses will be discussed later.

System Simulation

The formulas for probability of cell firing and expected output rate can easily be evaluated, given values for all parameters. A computer program was written to perform this evaluation and to construct curves of relative habituation as stimulation proceeded. As in the UATION program of Chapter II, stimulation was a constant intensity pulse with predefined duration and period. Also as before, system response was defined as the cell firing probability or average output rate at the start of each stimulus presentation or as the

total rate or firing probability during the time the stimulus was non-zero. These two measures of system response yielded similar results. In the curves presented below, the latter measure is employed. The control level of response was defined as the response to the first stimulus. Since all synapses are assumed to be identical, a single diffential equation was used to model changes in the common synaptic weight. Results are presented for the cases in which I^+ is constant and in which it depends on stimulus intensity. Simulation parameters are as listed in Table 3.1, and on the figures themselves.

Case 1: I^+ constant. Figure 3.4 shows curves of relative habituation of firing rate generated with constant synaptic modification drive and flat fiber and cell threshold density functions. These curves display the inverse-intensity effect: as stimulus intensity increases, relative habituation decreases. This effect is due here to the nonlinear nature of the population output function. At a low intensity, small changes in synaptic drive lead to the silencing of relatively many cells. At higher intensities, more cells are above their thresholds, and fewer drop out as the synaptic weight falls. The change in overall output rate is approximately linear in the synaptic decrement at higher stimulus intensities, since the T function is linear in W.

In this case, the synaptic modification drive is independent of stimulus strength, so relative habituation is also independent of input strength at intensities great enough to activate almost all cells. The curves of Figure 3.4 are the result of this interaction

TABLE 3.1

Parameters Used in Population Model Simulation

$N_s = 200$

$P_c = .75$

$W = 1.0$

$= 4.0$

Stimulus duration = .03

Stimulus period = .2

$F_{SA}(I) = 1/4\ (I - 1),\ 1 \leq I \leq 5$

Threshold values and synaptic modification gains indicated on figures.

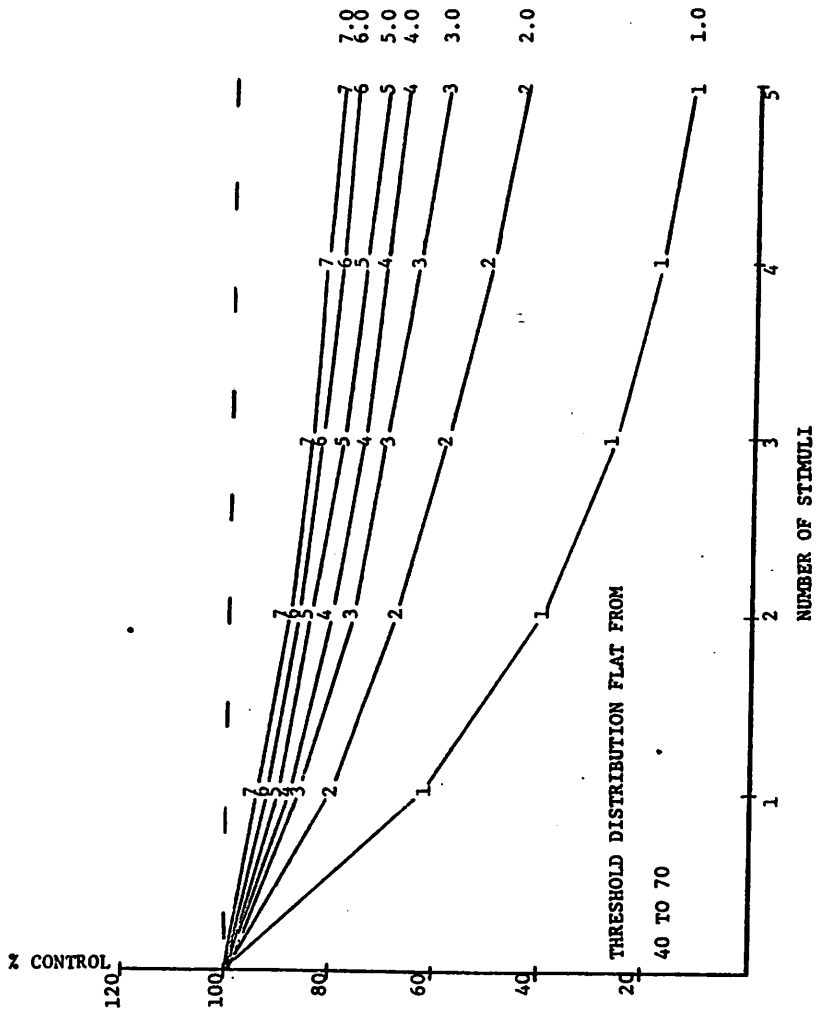


Figure 3.4 Relative habituation with constant synaptic modification drive. Curves of relative habituation of output rate at various intensities are shown. Intensity values as indicated on the curves; the same values are employed on all following figures.

between synaptic decrement and cell thresholds. Note that if all cell thresholds were zero, this case would reduce to the process model simulated previously, in which cell thresholds were not taken into account. At higher intensities, where the effects of cell thresholds are lost due to activation of all the cells, the two models yield the same results.

The intensity effect illustrated here can be sculpted by changes in the cell threshold density function, as shown in Figures 3.5, 3.6, and 3.7. Figure 3.5 was generated using a cell threshold density function skewed toward higher thresholds, as indicated. Due to the predominance of higher thresholds, decrements are somewhat greater in Figure 3.5 than in Figure 3.4. The curves of Figure 3.6 were generated using a cell threshold density function skewed toward lower thresholds, resulting in less relative decrement at all stimulus intensities. In Figure 3.7, a Gaussian threshold density function with the indicated parameters was used. Relative decrement is greater at low intensities due to the lack of low-threshold cells. At higher intensities, however, the cells are all activated, and relative habituation is approximately intensity-independent as before.

Case 2: I^+ a function of stimulus intensity. In this case, synaptic decrement depends directly on stimulus intensity. It might be expected that the effect of an increasing synaptic decrement could offset the threshold effect, to produce intensity-independent relative habituation in some range of intensities. That is, at a low stimulus intensity the output depends strongly on the synaptic decrement, but

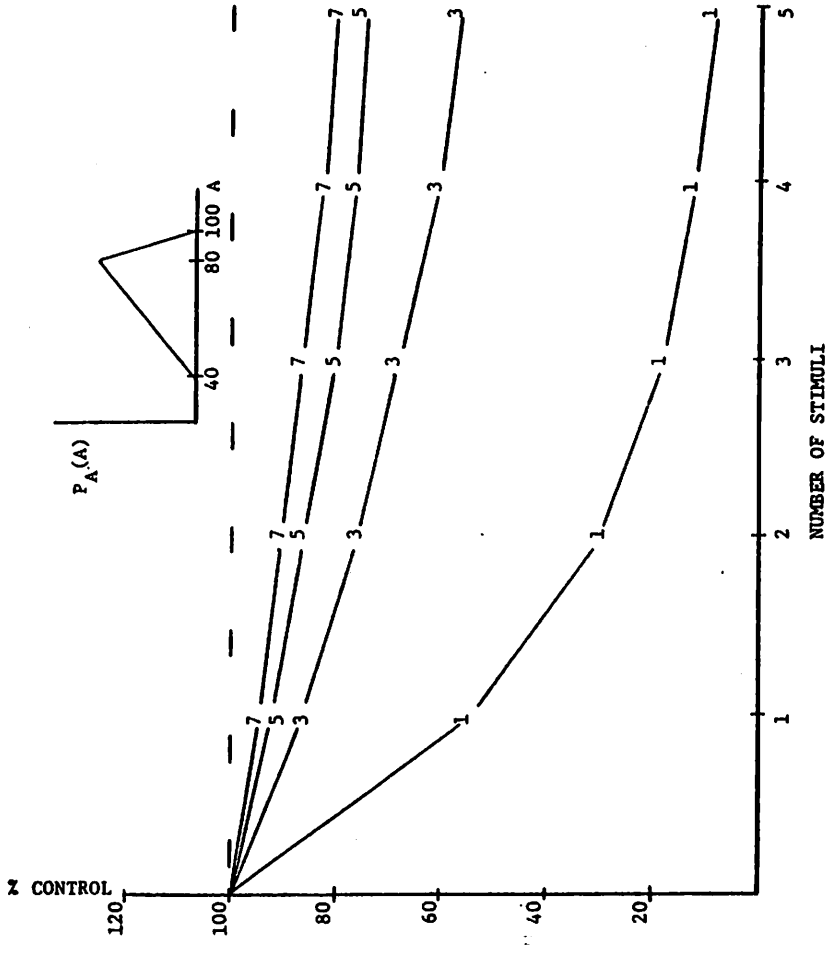


Figure 3.5 Curves as 3.4 but with skewed cell threshold distribution, as indicated in the figure.

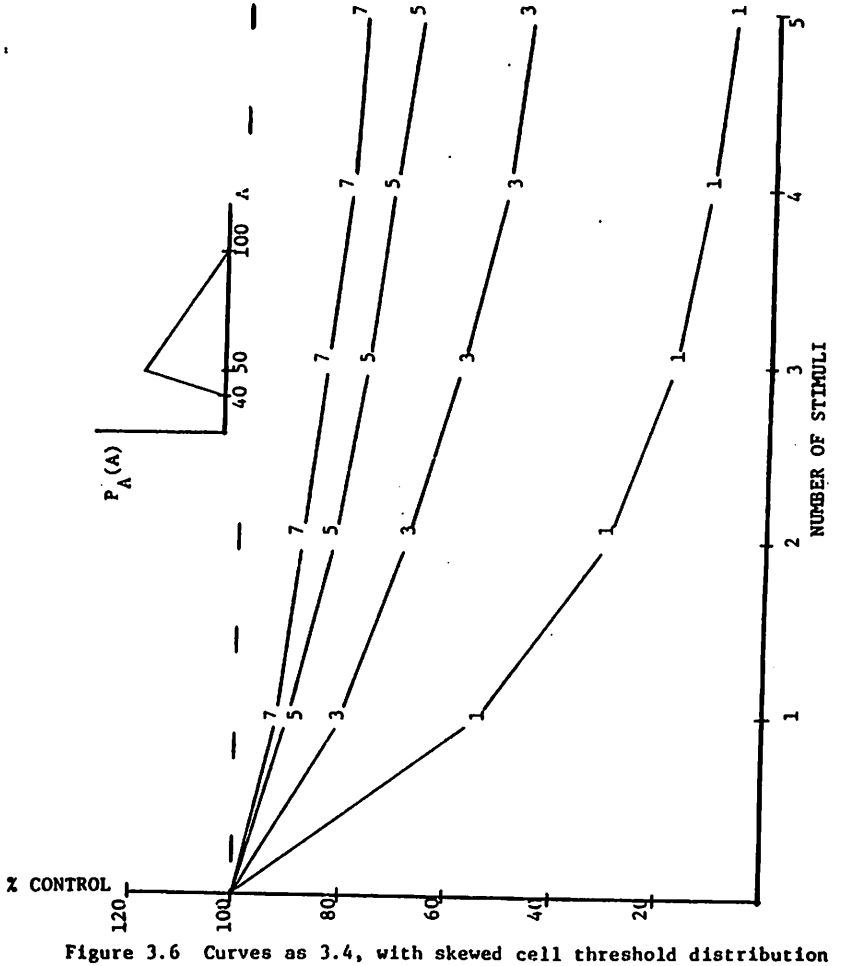


Figure 3.6 Curves as 3.4, with skewed cell threshold distribution indicated in the figure.

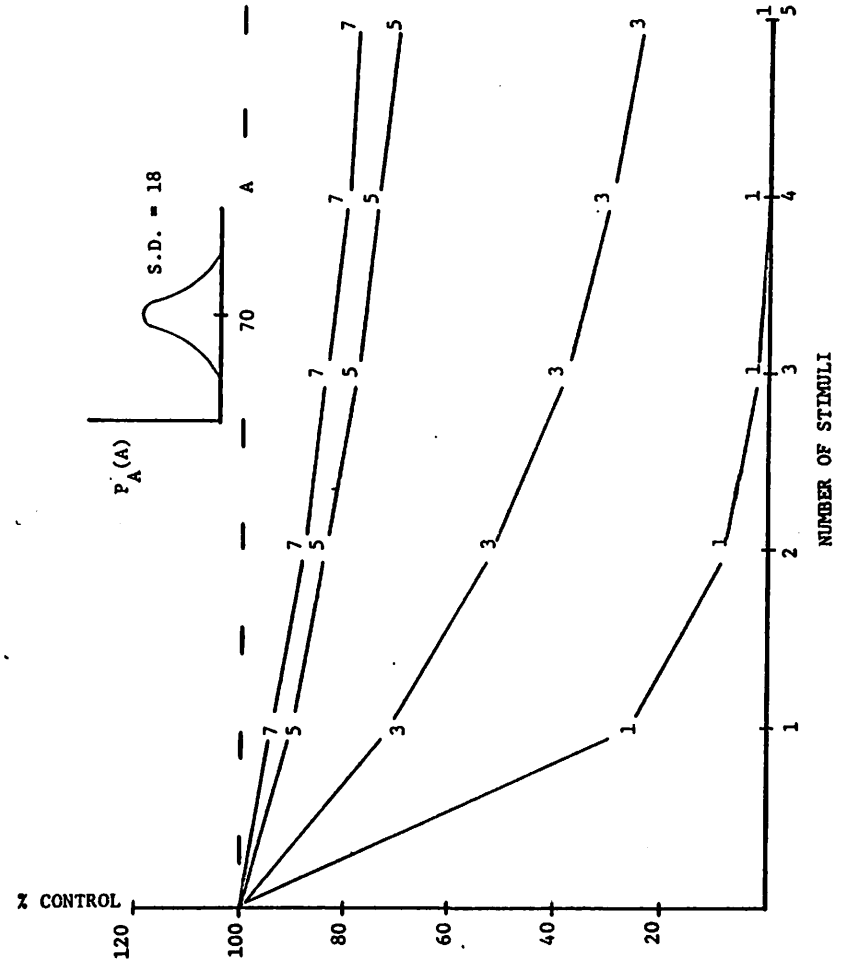


Figure 3.7 Curves as 3.4, generated using Gaussian cell threshold distribution with parameters indicated on the figure.

the decrement is small. At a moderate intensity, the decrement is greater but the output rate depends less strongly on the synaptic level. At the highest intensity, the cells are all active, and the threshold effects are absent. There, relative habituation should actually increase with increasing stimulus intensity, since, as above, output rate is linearly dependent on synaptic weight at high intensities.

A clear example of this behavior is shown in Figure 3.8, generated by a system with flat fiber and cell threshold density functions. At the lowest intensities, the threshold effect dominates, and relative habituation decreases with increasing stimulus strength. At moderate intensities, however, relative habituation is approximately independent of intensity, as indicated by the fact that curves marked 3, 4, and 5 (curves of moderate intensities) are covered by the curves marked 6 and 7 (curves of higher intensities). At the highest intensity (curve 7) the effect begins to reverse as all the cells become active, and relative habituation becomes directly dependent on stimulus intensity.

With increased synaptic modification gain, this overall effect is less pronounced, and an inverse-intensity effect appears. In Figure 3.9, a stronger synaptic modification drive leads to greater spread between curves. At higher intensities, all the cells are activated, and the inverse-intensity effect reverses. In systems with intensity dependent synaptic modification approximately intensity-independent relative habituation can be produced if synaptic

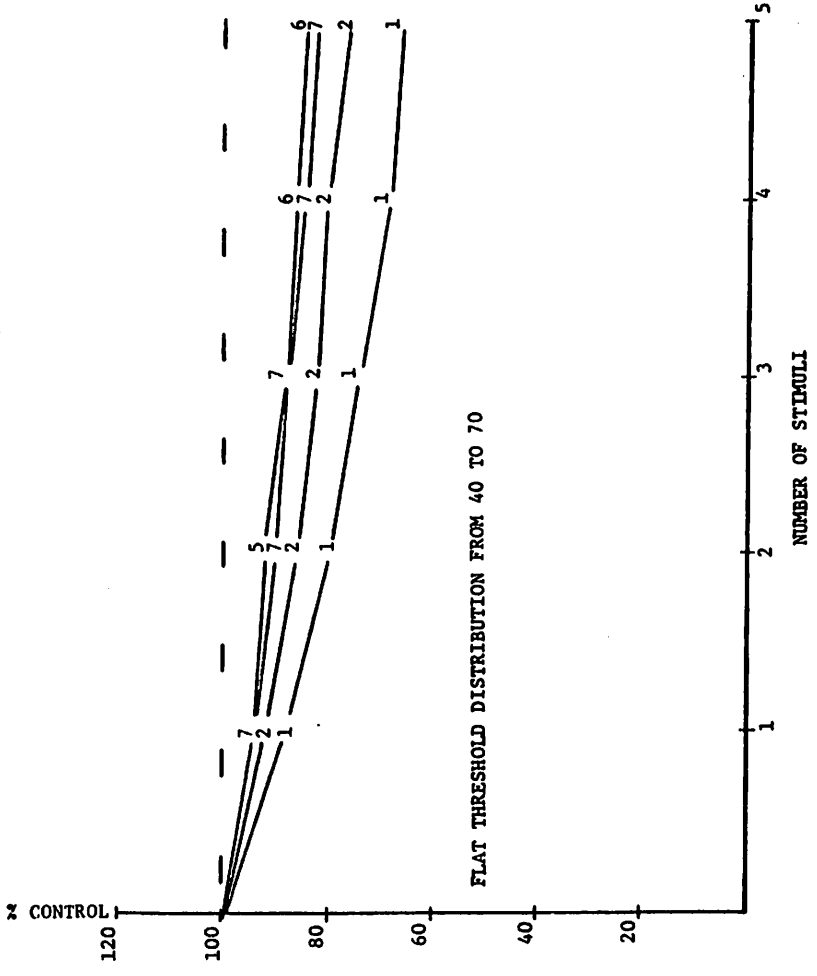


Figure 3.8 Curves of relative habituation generated using synaptic modification dependent on intensity of stimulus.

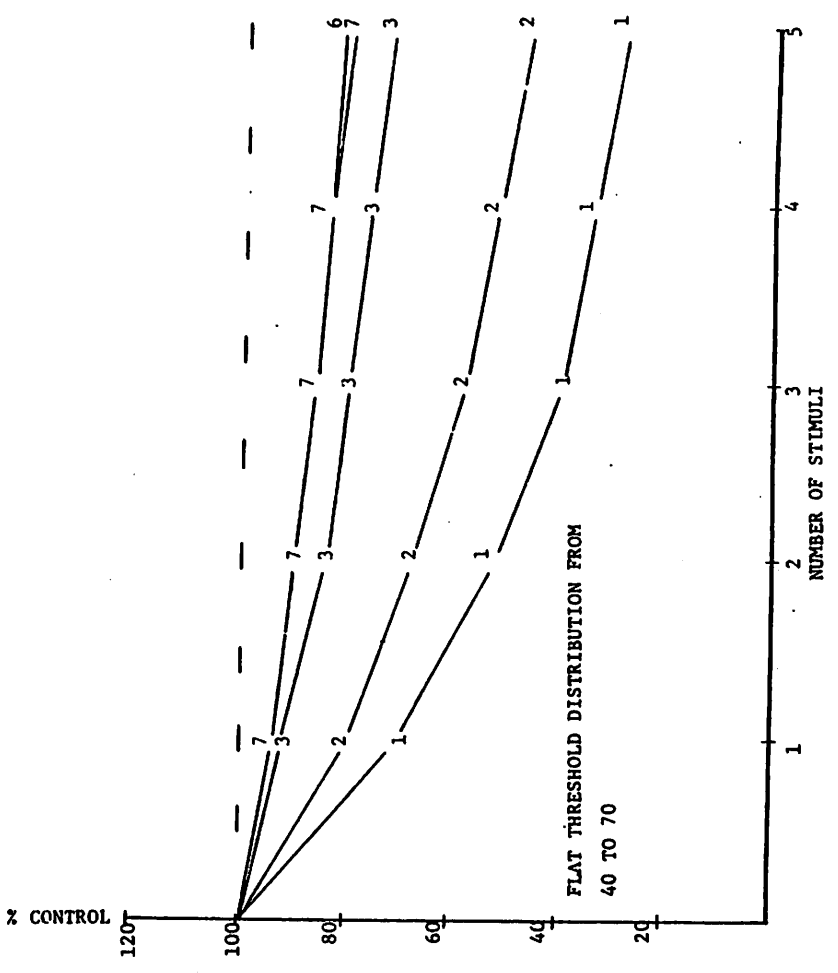


Figure 3.9 Curves as in Figure 3.8, synaptic modification gain increased.

modification drive is low and low-threshold cells predominate. As thresholds and synaptic drive increase, relative habituation becomes intensity-dependent.

Discussion

Computer evaluation of equations characterizing the system of Figure 3.1 operating with two types of synaptic modification shows that, in each case, both an inverse-intensity effect and approximately intensity-independent relative habituation can be realized. In the case of constant synaptic drive, relative habituation is inversely proportional to intensity at stimulus levels below which all cells are activated. At higher levels, where all cells are above threshold, relative habituation is independent of stimulus intensity, as in earlier models that involved no cell thresholds. When synaptic modification depends on stimulus intensity, system parameters may be set to yield either approximately intensity-independent relative habituation, or relative habituation that varies inversely with intensity. In these systems, however, relative habituation varies directly with stimulus intensity at levels that are great enough to activate all the cells.

S.- I. Amari (unpublished results) has recast the system in a form amenable to mathematical analysis under the assumption that the variance of the number of afferents to each cell is zero. His analysis confirms the results given here for high intensity stimuli; namely, that intensity independence occurs with constant synaptic modification drive and that intensity-dependen. drive leads to in-

tensity dependent habituation. Amari's analysis and the computer simulation give different results for low-intensity stimuli, however. While the simulations show an inverse-intensity effect, analysis indicates that no such effect should occur. This conflict points to the possibility that the variance in the number of afferents--taken into account in these simulations, but ignored in the analysis--is the cause of the inverse-intensity effect. In fact, if the number of afferents to each cell in the simulations is made large enough that their variance is a smaller factor in overall circuit operation, the inverse-intensity effect disappears, as Amari's analysis says it should.

The two forms of behavior of relative habituation with respect to intensity occur in a number of physiological systems. That habituation is inversely dependent on stimulus intensity is a part of the generally accepted operational definition of habituation. Earlier simulations illustrated the way this behavior could be gotten from a system containing both decremental and incremental effects. The system studied here shows how a simpler anatomy employing cell thresholds achieves the same result. In the light of Thompson's recent elaboration of his two-process theory of habituation (Thompson, et. al., 1973), it is especially interesting that approximately intensity-independent relative habituation can be achieved in a system employing a synapse whose value decreases with increasing stimulus intensity. In my earlier model, intensity-independent habituation was realized only through the use of synapses whose values decreased independent of stimulus intensity. The model presented here is thus

closer to current habituation theory.

The two cases of synaptic modification considered here correspond to two different physiological situations. Examination of these two cases will help to show how the model can be applied to a given physiological system. In this model, each synapse is affected only by the signal on its own fiber. In order to determine the way a synapse reacts to different stimulus intensities, it is necessary to see how intensity might be encoded on a single afferent fiber. If it is assumed that a fiber propagates only standard action potentials, then only interspike interval and burst duration are available for coding stimulus intensity.

It may be that neither of these effects codes intensity on single fibers, so that intensity information appears only in the number of lines activated by the stimulus. Physiologically, this corresponds to the situation in which an afferent bundle is directly stimulated electrically. Regardless of stimulus intensity, only a single action potential is generated on each activated line. The corresponding case in the model is that of intensity-independent synaptic modification (I^+ a constant).

An example of this situation is found in Zucker's exploration of the mechanism of habituation of the crayfish escape response (Zucker, 1972; Zucker, et. al., 1971; Krasne, 1973). Tactile stimulation of hairs on the abdomen of the crayfish elicits a powerful tail flip that carries the animal away from the source of stimulation.

The tail flip is mediated by the action of a pair of nerve fibers activated by the abdominal sensory cells directly and through interneurons. It has been shown that habituation of the escape response is due to decreased transmission between the interneurons and the sensory fibers. The model presented here was based on this circuit configuration. Zucker (1972) employed an electrical stimulus to excite the afferent bundle directly, corresponding to the case discussed above.

Intensity-dependent synaptic modification in the model corresponds to the case that stimulus intensity is coded on each afferent line in the form of burst frequency or duration. The afferent bundle, then, must be composed of the axons of a set of cells that are excited by the stimulus. This case is applicable to physiological systems involving polysynaptic chain of cells. Longer burst discharges induced by higher stimulus intensities yield more action potentials per stimulus, and lead to greater synaptic decrements. Simulations employing constant synaptic decrement and stimuli with durations that increased with intensity show that stimulus duration may be balanced against the threshold effect to yield results similar to those obtained with intensity-dependent synaptic modification. The model's intensity-dependent synaptic modification, then, may represent the effect of intensity-dependent burst duration.

Synaptic decrement has been shown to be frequency dependent in a number of preparations (Brunet and Kennedy, 1970; Groves, et. al., 1973; Wickelgren, 1967). At low frequencies, decrement is a direct

function of frequency. At higher rates, however, sensitization can occur, and synaptic efficacies may increase rather than decrease. Intensity-dependent synaptic decrement in the model can be considered to correspond to the effect of increasing burst frequency with increasing stimulus intensity only if it is assumed that burst frequencies remain within the range of synaptic decrement. This case can be applied, then, to systems in which the bursting properties of cells driving the decrementing synapses are properly related to the frequency sensitivities of those synapses.

A number of assumptions were made in deriving this model. Relaxation of these assumptions leads to extensions of the model to handle more of the properties of habituation. For the sake of simplicity, the synapses were modeled as a population of identical elements. This assumption allowed the use of a single differential equation to simulate the entire set of synapses activated by a stimulus. Extensions may be made in two ways to deal with a population of synapses having a range of parameters. First, the synapses could have different resting values. This change leads to a complication of the equation for the probability of cell firing. Further complications might then involve correlations between fiber threshold and synaptic weight. For example, it may be that low-threshold fibers are more likely to end in more powerful synapses. In the case of constant synaptic modification drive, this would lead to reduced relative habituation at low stimulus intensities. The synapses may also have different modification parameters and the parameter values might be correlated with synaptic strength or fiber threshold. For

example, if synapses on higher-threshold fibers were more difficult to modify, the inverse relation of relative habituation to stimulus intensity would be augmented.

Two related circuit models of the below-zero effect can be based on the current model using populations of synapses with different time constants. First, consider the case of a system whose response falls to zero upon sufficient stimulation. An example of such a system is the rat startle response (Prosser and Hunter, 1936). Suppose that the model described above holds, so that a single differential equation can represent the common value of a population of identical synapses. If the synaptic modification gain is great enough, the synaptic value will fall below the point that any cell can be activated, and the system response will go to zero. Cessation of stimulation at the point that the response goes to zero will lead to response recovery at a rate determined by the synaptic time constant. Continued stimulation may drive the synaptic level lower, causing the system response to remain zero for a time following cessation of stimulation. Once the response becomes non-zero, it will recover with a time course independent of the length of time it remained zero; i.e., independent of the duration of the habituation session. This simple model does not capture the property that systems that exhibit the below-zero effect recover more slowly following prolonged stimulation.

A modification of this model is suggested by the mechanism added to the process-level model to handle the BZE. Changing the time con-

stant of synaptic recovery in that model can correspond in the population model to a depletion of synapses with longer time constants. Suppose then that the population model involves a set synapses with a distribution of time constants, rather than the set of identical synapses with a single time constant considered above. Cell thresholds may be set such that depletion of a significant portion of the synapses sends the system response to zero. During initial habituation trials, the synapses with smaller time constants deplete significantly, while those with long time constants change less. Recovery from that point in habituation training will follow essentially the time course of the faster synapses. Habituation training continued after the response has fallen to zero will result in depletion of the synapses with larger time constants. Recovery following prolonged habituation will follow the time course of the slower synapses, since the immediate recovery of the fast synapses will not be enough to activate the cells. In this case, then, in which system response falls to zero, the below-zero effect may arise as a result of a distribution of synaptic time constants.

In Farel and Thompson's (1972) work on frog spinal cord, an example is given of a system that displays the BZE with a non-zero asymptote. This behavior may most easily be modeled using the BZE system considered above together with a set of nonplastic synapses, as illustrated in Figure 3.10. The pathway to the output through the interneurons behaves as above, displaying the BZE with an asymptotic response of zero. The direct pathway from input to output establishes a non-zero level towards which the response level falls as

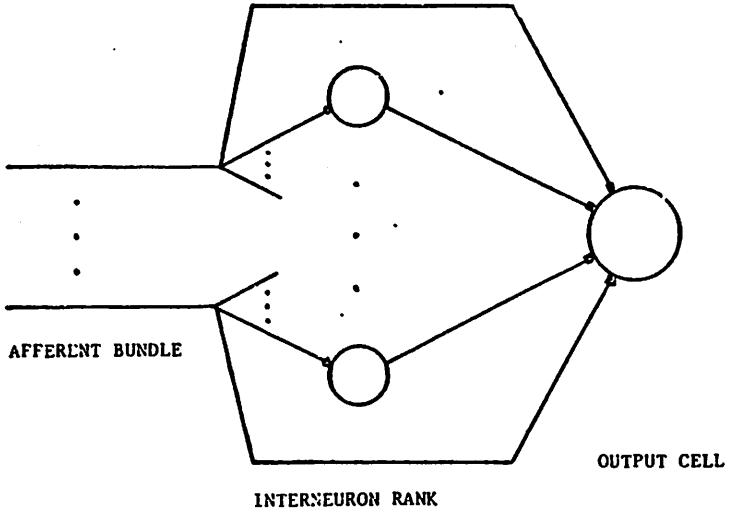


Figure 3.10 Circuit showing below-zero effects. Afferent bundle makes contact with interneuron rank as in Figure 3.1. Afferents and interneuron outputs contact output cell, R. Only synapses from afferents to interneurons are plastic.

the interneurons cease firing. This system, then, exhibits the BZE with a non-zero asymptotic response. It is interesting to note in this regard that Farel and Thompson (1972) mention a nonplastic monosynaptic component in the response they studied, though according to Brookhart and co-workers (1960) this monosynaptic path is very weak. In more complex systems, the monosynaptic pathway shown in Figure 3.10 can be replaced with a nonplastic polysynaptic pathway from stimulus to response, or with a set of nonplastic synapses that make contact with the same interneurons served by the plastic connections.

The population model developed in this section, then, can serve as the basis of an extension to plausible neural circuitry of the process-level BZE model considered previously. This population model of the BZE can be applied where the necessary circuitry exists to support it. Thus the rat startle response (Prosser and Hunter, 1936) and the electrical response of frog spinal cord ventral root to dorsal root stimulation (Farel and Thompson, 1972) may be modeled in terms of the system considered here. Each involves polysynaptic chains of cells in which the systems of Figures 3.1 and 3.10 can easily be embedded. This population model cannot be applied to monosynaptic responses that do not involve cell thresholds (Farel, et. al., 1973). Nor can it be applied to systems that involve one or a very few fibers and cells. In such cases, the BZE is likely to be a purely synaptic affair, and more must be known about the biochemistry of habituation at a single synapse before a detailed explanation of the BZE can be given.

The population model can also be applied to a circuit-level description of the long-term effects, as mentioned in Chapter II. The response of a system that displays the long-term habituation effect falls to a lower asymptote with each habituation session. The BZE model considered above can be modified to show this behavior in a manner similar to that considered in Chapter II. If the synapses of the nonplastic pathway are made plastic, and are governed by one of the two LTE schemes considered in Chapter II, then the output of the model considered above will display both the LTE and BZE due to the different plastic properties of two separate sets of synapses. The LTE, however, are realized in this model with a set of synapses that are themselves assumed to display the LTE. While such synapses appear to exist--for example, the monosynaptic gill-withdrawal response in Aplysia (Carew, et. al., 1972) has been shown to display long-term effects--it is interesting to consider a possible circuit realization of the LTE using simpler synapses.

The two sets of effects--and the two separate pathways--may be integrated if the neurons are given saturation characteristics. In the basic population model, the firing rate of an interneuron is taken to be the positive difference between its total synaptic drive and its threshold. As synaptic drive increases, output rate also increases without bound. Suppose now that the firing rate saturates above some total synaptic drive, so that further increases in postsynaptic potential lead to no further increases of firing rate. Such a characteristic can be realized by a variety of functions (Wigstrom, 1975; Stein, et. al., 1974; Leibovic, 1972; Freeman, 1975).

In the model with saturation, cells driven strongly enough will be saturated, and synaptic decrements will have little or no effect on their output rates. These saturated cells can correspond to the nonplastic pathway in the BZE population model of Figure 3.10. A cell that is less strongly driven (either due to a higher threshold, fewer active afferent fibers, or lower synaptic weights) will be strongly affected by changes in synaptic values. The output rates of such cells will decline, giving rise to the desired response decrement. Those cells whose outputs fall to zero will serve as the basis of the BZE, as in the previous model, due to changes in synapses with long time constants. The outputs of cells that are neither zero nor saturated will continue to fall throughout the course of habituation training. This model realizes the third form of BZE, then, in which a slow decline in output is seen, rather than a rapid fall to a steady asymptote.

The long-term effect in this model system is in fact a sort of "above maximum effect" that employs a non-zero saturation in the same way that the zero "saturation" gives rise to the below-zero effect. If with each habituation session, synapses with very long time constants on each saturated cell are decreased, such a cell may eventually fall out of saturation. Its output rate will fall, possibly to zero, as the faster synapses connected to it decrease in value. As this long-term process continues, the asymptotic output of the overall system will decrease with successive habituation sessions. If the long time constant synapses act only to bias the cells into saturation, then the responses of those cells will again

be saturated when the fast synapses recover. The overall system response will therefore recover between sessions, but fall to a lower asymptote with each session, as desired. Note that full recovery requires that the cells that do not saturate have few slow synapses, so this model may correspond more closely to the one of Figure 4.10 in which LTE and BZE paths are separated.

The below-zero and long-term effects can be modeled on a level closer to plausible neural circuitry, then, employing the population model developed in this section. Two possible realizations of the below-zero effects are given here to complement the classification of the BZE on the process level given in Chapter II. Similarly, the synapses of the population model may operate according to the process-level equations of Chapter II, or a cell saturation effect may be incorporated into the population BZE model to realize the LTE with simpler synapses.

Generalization effects can be modeled in this system by relaxing the assumption that each fiber makes contact with each cell with fixed probability. This relaxation amounts to making a distributed system of the current lumped model. In that case, a single differential equation would be needed to represent the value of each model synapse. A given stimulus would excite a particular portion of the overall system. The excited portion would behave like the lumped population system presented here. Changing the stimulus would result in activating a different set of afferents and synapses. A distributed model would take such changes into account. Such a model would allow

the detailed investigation of the common elements notion and its expression in stimulus generalization (Thompson, 1965; Thompson, et, al., 1973). The population model presented here forms the basis for understanding the operation of the more complex distributed system.

Finally, note that cell and fiber dynamics were not included in this model. It was pointed out that the derivation of the equation for the probability of cell firing assumed simple addition of post-synaptic potentials and involved frequency- and intensity-independent cell and synapse responses. Synapses may not be able to follow high frequencies of stimulation, however, leading to decreased potentials at high frequencies. Similarly, since cell dendrites may be modeled with first-order differential equations, potentials due to high frequency stimulation add to produce a greater total EPSP than that produced by low frequency stimulation. A more complicated model could incorporate the synapse effects through the use of a frequency- or intensity-dependent term in the U-function. Cell dynamics could also be added through use of one of a number of single cell models (Stein, et. al., 1974; Segundo, et. al., 1968). Non-linear summation of post-synaptic potentials could also be included in a more complicated model (Martin, 1955). These complications would help to show how more detailed properties of cell and synaptic responses work to shape the relative habituation curves.

CHAPTER IV

SUMMARY AND CONCLUSIONS

This work began with the desire to construct relevant model habituation networks. These models were to embody the generally recognized characteristics of habituation in simple networks that employed physiologically reasonable mechanisms to realize habituation. The two-process theory of Thompson and his colleagues was chosen as the basis of this modeling for a variety of reasons. Among these reasons were that the two-process theory is complete in its handling of habituation data, detailed in its statement of the operation of its components, and satisfying in its close relation to physiological circuitry. A simplified model based on a more complex circuit due to Groves and Thompson (1972) was formulated and was explored through analysis and computer simulation. It was shown that this simple model could realize several of the properties of habituation, as reviewed below. Other habituation properties could be added to this network in straightforward ways that involved complications of the synaptic modification rules.

A more detailed model of habituation in a single SR channel involving a population of cells with thresholds was proposed and simulated. This model showed how the desired intensity properties of habituation might arise as a result of interactions between synaptic depletion and cell thresholds. This model also opened the door to explanations of other habituation properties on a level closer to physiological details than the previous model. More complex models of those effects were discussed in the chapter dealing

with the population model.

This work has demonstrated that a simple two-process network can realize the properties that seem central to the operation of habituation; viz., the properties of decrement and recovery, and the effects of stimulus intensity. The decrement and recovery properties were imbedded in the model through the use of differential equations to represent synaptic weights. Modification rules of this form readily realize the properties of exponential decay of response during habituation training, and gradual recovery following cessation of stimulation.

The desired effects of stimulus intensity in the model are twofold, according to Thompson and co-workers (1973). The relative decrement due only to central habituation should be independent of stimulus intensity, as exemplified by habituation of one pathway in the isolated frog spinal cord (Farel and Thompson, 1972). This effect was realized in the initial model with a constant synaptic decrement. It was pointed out that this modification rule lead to a decrease in synaptic efficacy that was independent of intensity, and that further modeling required a synapse whose value decreased more as intensity increased. The constant synaptic modification rule was a stopgap measure, then, designed to allow easy exploration of the two-process network at a level once-removed from more complicated modeling considerations. In the population model, the required intensity-dependent synaptic modification could be used to realize the required intensity-independent habituation , properly

sealing the gap that was at first more crudely stopped.

The two sets of simulations using the HABIT and UATION programs showed that curves of reasonable fit to data taken from experiments on the spinal cat (Thompson, et. al., 1973) could be generated by the simplified two-process network. Two strategies were found whereby the circuit generated those curves, and mixtures of the strategies were simulated and discussed. Parameter values were found so that the desired curves could be generated, based on either of the two operating strategies or their mixture. Within the framework of the simple circuit, a nonlinear weight modification function was investigated, and its contribution to the overall response was discussed.

As was mentioned in the introduction, the intensity curves of Figure 2.2 are more complex than those the circuit of Figure 2.6 can generate. Investigations of the simple circuit verify the hypothesis made in the introduction that a two component sensitization process would be enough to generate curves with forms closer to those of Figure 2.2. A fast component whose relative contribution to the overall response decreases with increasing intensity and a slower component with a value that increases with increasing intensity could combine with the decrementing component to generate curves like those of Figure 2.2. The physiological basis of the switching from one form of sensitization to the other might be inhibition that reduces the output of the faster component at high stimulus intensities. In any event, the simpler network of Figure 2.6 provides a base from which more complex models can readily be constructed.

The desired effects of stimulus frequency were also realized through the use of weights that vary according to differential equations. Because of recovery between stimuli, habituation in the model decreases with decreasing stimulus frequency. These are the results required of central habituation in Thompson's theory.

Two qualifications must be made to this conclusion, however. The first concerns the range of frequencies employed, and the second deals with the effect of confounding the frequency and intensity characteristics. Data from a number of preparations (Farel and Thompson, 1972; Groves, et. al., 1969; Farel, et. al., 1973; Thompson and Spencer, 1966) indicate that as frequency of stimulation increases, so does habituation, when stimulus frequency is low or moderate. Low and moderate are defined here with reference to the recovery properties of the response in question. Thus, in the isolated frog spinal cord, recovery is relatively rapid (a time-constant of one or two minutes), and stimulus rates employed in the work cited above are moderate with respect to the recovery time-constant (1/sec. to 1/10 sec.). Groves and co-workers (Groves, et. al., 1969), however, report increasing decrements of response in the flexion reflex of the spinal cat up to frequencies of 16/sec., while the recovery time-constant appears to be on the order of a few minutes (Spencer and Thompson, 1966). Decrement of a single slow process cannot be the cause of this variation at high frequencies.

An explanation for this frequency effect is that the cat flexion reflex involves decremental processes with short and long time-con-

stants. The faster process would recover between single stimuli delivered at low rates, and therefore contribute little or nothing to habituation at low and moderate frequencies. At a higher stimulus frequency, the fast process would be depleted, resulting in greater habituation at low and moderate frequencies. This hypothesis could be checked by observing the recovery from habituation at high stimulus rates. Recovery should go rapidly for a short time, then proceed more slowly. Note that this is reminiscent of recovery in Aplysia (Pinsker, et. al., 1970), where recovery consists of two components. Recovery curves at this level of detail were not given by Groves and his co-workers, so this hypothesis can be checked only by further experimentation. A similar explanation of the frequency effects has been put forth by Graham (1973).

The second qualification to the frequency effect in the model concerns possible interactions between the frequency and intensity effects. At a low stimulus intensity, only the effect of central habituation appears in the response curves. At a higher intensity, however, central sensitization becomes a factor in the response. Sensitization is frequency dependent, as is decrement, and increases with increasing frequency. At a moderate intensity, then, response decrement may actually decrease with increasing stimulus frequency due to the fact that sensitization increases with frequency while the effect of central habituation changes little. Such a result was reported by Griffin (1970) in work on the flexion reflex of the spinal cat. Griffin concluded that the decrement seen in habituation must be due to some process that is triggered by repeated stimulation

but that takes place during the interstimulus interval. With more time to act between stimuli delivered at low rates, such a process would result in greater habituation at those rates.

As indicated above, these results could also be due to the properties of sensitization. In a preparation that displays strong sensitization, sensitization will be pronounced at a moderate stimulus intensity and frequency. As the frequency is decreased, sensitization will recover between stimuli, and habituation will appear greater. This effect is related to the problem of recovery from high intensity stimuli discussed below. To test this hypothesis as to the origin of Griffin's data, a lower intensity stimulus should be employed in any preparation that shows Griffin's effect. At some intensity, the effect should reverse, and habituation should appear to be directly related to stimulus frequency. In any given preparation, then, care must be taken to tease out the effects of these processes in order properly to interpret the experimental results.

The characteristics of long-term and below-zero habituation were discussed only sketchily. It was shown that particular functional forms of synaptic modification could be used to realize these effects in the process-level model. Interpretation of these forms, however, had to await the more detailed population-level model. The real uses of the addition of these mechanisms to the process-level model are the ability to test particular functional forms that might generate these effects and the ability to investigate in

detail the ways the effects interact. Because the process-level model can say little about underlying mechanisms, the long-term and below-zero effects were not actively pursued at that level. The population model suggests ways that the effects can be realized in more neural terms, but again, much of the model is couched in terms of inferred synaptic action. More detailed and realistic circuit modeling must be complemented by more detailed models of synaptic action.

The models investigated here suggest one experiment beyond the check discussed above with respect to the effect of stimulus frequency on habituation. This experiment is based on the model's recovery from habituation to high or moderate intensity stimulation. These models have indicated that in order to generate curves of relative habituation having pronounced peaks in early trials in response to moderate or high intensity stimuli, sensitization must have a smaller time constant than central habituation. During recovery, then, sensitization may decay more rapidly than habituation. In that case, the recovery curve following habituation to stimuli intense enough that sensitization is a factor in the response should show an initial dip as sensitization recovers, then a rise as central habituation recovers. Griffin's data, as discussed above, may represent such an effect. I have not seen detailed curves of recovery plotted following habituation to high intensity stimuli, however. These curves should be generated experimentally as a possible test of the two-process theory.

This work, then, has offered a first approach to relevant network models of habituation. Simple circuits and functional forms of synaptic modification have been suggested that duplicate many of the characteristics of habituation. As mentioned in Chapter I, one characteristic of habituation that has not been properly treated by either the TPT or the model theory is the ability to habituate to a regularly repeated stimulus and to dishabituate when the stimulus is given out of step or withheld. This characteristic is dealt with in greater detail in Chapter V with respect to a temporal sequence memory model based on the structure of the dentate gyrus of the hippocampus.

CHAPTER V

A TEMPORAL SEQUENCE MODEL BASED ON THE STRUCTURE OF THE DENTATE GYRUS

One characteristic of habituation mentioned only briefly in earlier chapters is the ability to habituate to a regularly-repeated input and to dishabituate when the stimulus is withheld or its period is altered. This is another of the weak characteristics of habituation that is demonstrated only in certain preparations. As was discussed in Chapter I, however, the hippocampus may be such a system. The observation of time-locking phenomena in the hippocampus (Vinogradova, 1975), together with McLardy's (1959) conjecture that the hippocampus acts as a detector-coder of temporal stimulus properties, prompted an effort to model this characteristic of habituation employing the structure of the dentate gyrus.

The temporal phenomenon described above--called time-locking here--has been studied only seldom. The main sources of data on time-locking at the behavioral level are studies of habituation of human arousal levels. Sokolov (1961) reported the classical time-locking effect in studies on the orienting response. Following habituation to a regularly-repeated, low intensity stimulus, the orienting response returned if the stimulus was withheld. Graham (1973) and Kimmel (1973) discuss further experiments on time-locking using the orienting response. Graham notes that experimental results in this area conflict; some studies show clear differences in habituation when stimuli are presented at fixed rather than variable intervals, while other studies indicate no differences. Groves and Thomp-

son (1970) discuss experiments on temporal conditioning, concluding that if temporal conditioning does occur, it is most likely a "state" phenomenon rather than a property of the stimulus-response system.

In a series of experiments on the characteristics of single units in the hippocampus of rabbits, Vinogradova and her co-workers (Vinogradova, 1970; Vinogradova, 1975 a, b; Vinogradova, Semyonova, Konovalov, 1970) found cells in CA3 and in CA1 that formed extrapolatory responses. In these experiments, unanesthetized rabbits housed in small boxes were exposed to a variety of sensory stimuli. Post-stimulus time histograms (PSTH) were used to check the responses of single units. Cells were found that responded to stimuli of all modalities with either inhibition or excitation. In most cases, responses declined as the stimuli were repeated and were renewed when any stimulus parameter was altered.

Of particular interest is the group of cells that Vinogradova called "extrapolatory cells" (and which I am tempted to call "future detectors"). Such a cell exhibits a bimodal PSTH with an initial peak, followed by a quiet time and then another peak. The second peak may occur three seconds or more after the first. With repeated stimulation, the duration between peaks becomes "locked" to the interstimulus interval, such that the cell's firing rate begins to accelerate before the stimulus is presented. The increase occurs in the absence of the stimulus, then. It should be noted, however, that this result is reported to be weak and is often masked by background activity (Vinogradova, 1975). Another class of cells sensitive to

temporal parameters of the stimulus are those that become locked to the stimulus duration (Vinogradova, 1970). These cells cease responding at the usual time of stimulus offset, even if the duration has been shortened or lengthened.

The existence of a time-locking phenomenon associated with habituation, and the establishment of the existence of time-sensitive cells in the hippocampus lead me to seek a model of the dentate gyrus that might be used as part of an habituation system with time-locking characteristics. A variety of temporal memory models exist that may be drawn on to guide the search. First, it must be noted that the time-locking seen in the hippocampus may not be due to processes intrinsic to the hippocampus at all, but may come from activity in other brain regions. The work reported here focuses on the way that temporal conditioning might arise within the dentate gyrus, in order to investigate the possibilities inherent in the circuitry of the hippocampus. That the time-locking property might arise in another part of the brain is freely acknowledged.

Groves and Thompson (1970) and Sokolov (1975) espouse two different views as to the basis of temporal conditioning. As mentioned above, Groves and Thompson relegate temporal conditioning to the state system, saying that this system becomes entrained to regularly-repeated stimuli. This implies that the state system contains long-period endogenous oscillators that can fix on the period of the stimulus. Such oscillators were not discussed by Groves and Thompson (1970). Long-period oscillators can be constructed using plausible

neural mechanisms (Feldman and Cowan, 1975), and are suggested by the form of the PSTH of Vinogradova's extrapolatory cells. Sokolov (1975), following his stimulus-comparator paradigm, gives circuits using time delays and coincidence detectors to achieve time-locking. This approach is similar to the model presented below, in which a time-delay element is in effect set up that corresponds to the interstimulus interval. Sokolov does not say, however, where long time delays are to come from.

Other temporal sequence memory models exist that seem best suited to time-locking of signals with relatively short interstimulus intervals. In a model due to Grossberg (1969), snapshots of the states of a set of afferents are taken by cells activated at successive instants by a pulse traveling along a command fiber. This system is capable of recording and replaying arbitrary space-time sequences. For long interstimulus intervals, however, a very long or very slow conductor is required to carry the activating impulse. Fukushima (1973) has detailed a model in which successive snapshots of input activity are stored in a neural shift register and combined continuously to form the system's output. This model has the attractive property that each output is generated by the entire input sequence that precedes it. The model is also able to infer from partial information which of a number of stored sequences is to be recalled. Marr (1970) notes that his inferential model for neocortex can handle sequences of inputs by a process of chaining in which an evoked output acts to trigger recall of the output expected next. The model detailed here is similar to Marr's model in some respects,

but Marr's does not have a structure directly applicable to the dentate gyrus. Marr's model would also not be able to handle sequences with long interstimulus intervals. Finally, Spinelli's OCCAM model (1970) is well suited for handling short inputs, again by a snapshot mechanism that examines the afferents at successive moments and stores each moment's signal in a different circuit element. However, Spinelli's model could not directly handle discrete inputs spaced by time intervals that must also be recalled.

This look at available models of temporal sequence memories reveals two common elements. First is the snapshot method that changes a continuous input into a stream of discrete events, however rapid the shutter. The second is the propagation of snapshots to turn time into space. These two techniques can be traced at least to Culbertson (1956) who sketched models similar to those of Fukushima, and to Beurle (1954) who suggested that traveling waves could be used as computational elements. One problem with the propagation methods is that a moving trace must eventually come to the limits of the medium through which it travels. Long input sequences--or long interstimulus intervals--lead to the necessity of unrealistically long chains of cells or impossibly slow propagation. The model detailed below is recurrent, and so needs no long chains of cells for its operation. This model is designed to store and recall discrete inputs separated by relatively long time intervals that are also to be recalled. The model therefore illustrates one way that temporal conditioning of the sort observed by Vinogradova could be realized in the dentate gyrus.

The temporal sequence memory presented here is so designed that once it has been sufficiently exposed to a sequence of inputs separated by particular time intervals, it can reproduce the sequence with proper timing if cued by an initial portion of the sequence. The system thus learns by rote the interstimulus intervals involved in a given repetitive sequence. Predictions generated by the memory may be used as the basis of temporal conditioning, as will be discussed following presentation of the model itself. In the following sections, I first discuss the physiology of the hippocampus to motivate the model's structure. I then detail the operation of the model and present results of its simulation.

Hippocampal Anatomy, Physiology, and Possible Function

As mentioned above, the hippocampus has been studied in detail at the anatomical, physiological, and behaviorial levels. I will be brief in this description of the hippocampus, since a number of excellent reviews and collections of the relevant literature exist to which the reader can refer for more detail. These works include especially the books of Isaacson (1974), and Isaacson and Pribram (1975), the thesis by Segal (1973), and reviews by Kilmer (1975) and by Douglas (1967). These books and papers summarize particularly well the broad spectrum of research and points of view on the hippocampus.

Anatomically, hippocampus is a region of phylogentically old cerebral cortex called archicortex (Chronister and White, 1975).

It is distinguished from the rest of the cortex by differences in cell types and in its layering structure. The hippocampus consists roughly of two U-shaped sheets of cells folded into one another to form separate regions called Ammon's horn and the dentate gyrus. Each of these regions contains major projection cells--called pyramidal cells and granule cells, respectively--as well as a variety of interneurons.

Fibers that enter and that originate within the hippocampus generally remain within parallel slices arranged perpendicular to the long axis of the hippocampus (Andersen, Bliss and Skrede, 1971). Each such slice is called a lamella. Functionally, then, the hippocampus may be considered to be composed of a number of parallel slices. The anatomy of the lamella is shown in Figure 5.1, adapted from the work of Andersen (1964 a). The layers of granule cells and pyramidal cells are indicated with U-shaped bands. The axons of the granule cells form a bundle of fibers, called mossy fibers, that make excitatory contact with the dendrites of pyramidal cells in a region of Ammon's horn called CA3 (Andersen, Blackstad, and Lømo, 1966). These cells in turn excite pyramidal cells in the CA1 region by way of axon collaterals called the Schaeffer collaterals (Andersen, et. al., 1966). Not shown on the figure is a set of fibers arising from the cells of a part of the CA3 region and terminating in the dentate gyrus (Zimmer, 1973; Gottlieb and Cowan, 1973). These fibers spread to make contact with cells in other lamellae. Each region of the hippocampus also contains a powerful recurrent inhibition mediated by interneurons (Andersen, Eccles and Løying, 1964; Andersen, Bruland,

Figure 5.1 Anatomy of hippocampal slice after Andersen, et. al., 1971. A = Ammon's horn, DG = dentate gyrus, En = entorhinal cortex, G = granule cell layer, P = pyramidal cell layer, F = fornix, MF = mossy fiber tract, pp = perforant path, Sch = Schaeffer collateral fiber tract.

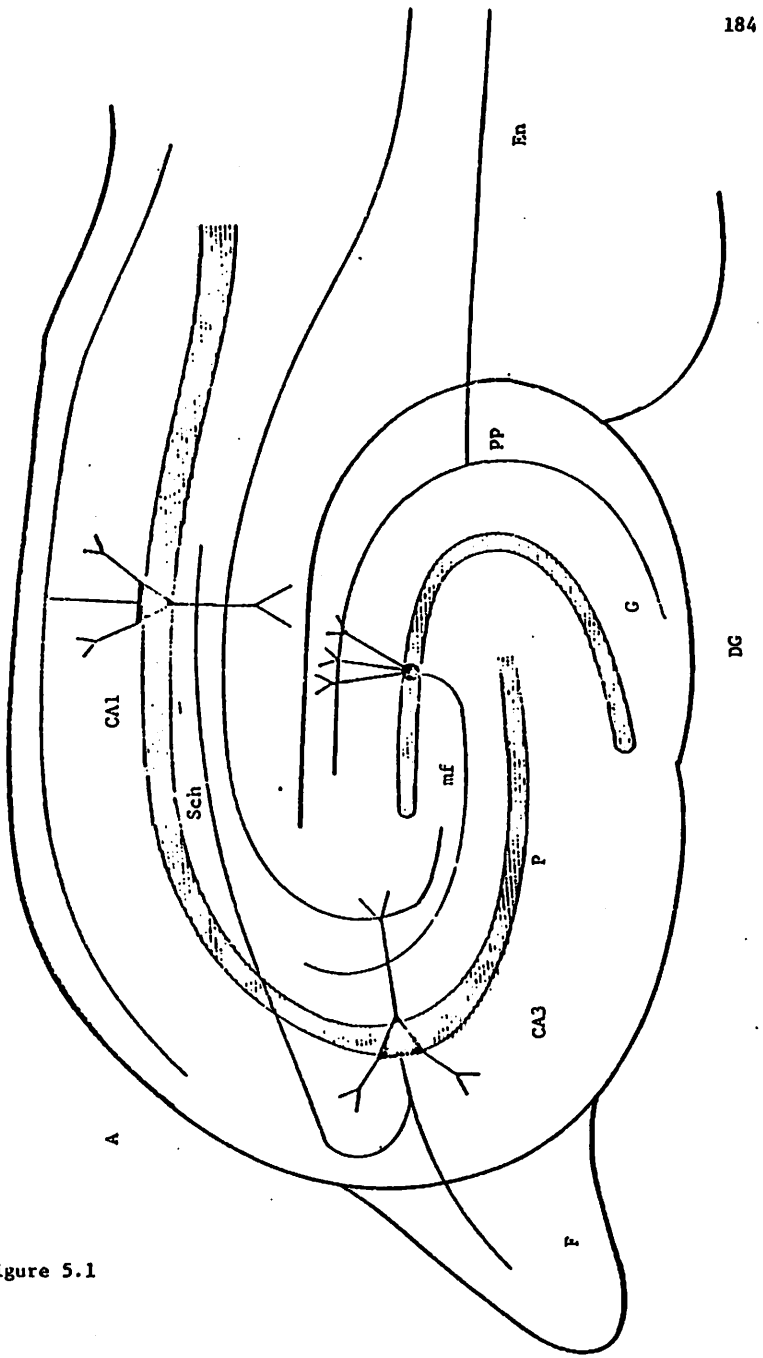


Figure 5.1

and Kaade, 1961). A recurrent excitation system has also been demonstrated in CA3 (Lebovitz, Dichter and Spencer, 1971).

The hippocampus has at least three sources of inputs. The nearby entorhinal cortex gives rise to a pathway called the perforant path (Hjorth-Simonsen and Jeune, 1972; Hjorth-Simonsen, 1972) that excites granule cells and pyramidal cells (Andersen, Holmquist and Voorhoeve, 1966; Segal, 1972; Vinogradova, 1975). This input brings pre-processed sensory information into the hippocampus from cortical association areas (Van Hoesen, Pandya and Butters, 1972). The septum is a second source of input. Axons originating in the septum travel over the fimbria-fornix fiber system to excite cells in all regions of the hippocampus (Andersen, et. al., 1961; Vinogradova, 1975; Mosko, Lynch and Cotman, 1973). This input path may carry information into the hippocampus concerning reinforcement (Olds, 1969), arousal level (Segal, 1973), and bodily state (Covian, 1967). Finally, a number of regions in the brain stem give rise to fibers that make contact with pyramidal cells in Ammon's horn, and granule cells and possibly interneurons in the dentate gyrus (Conrad, Leonard and Pfaff, 1974; Segal, 1975; Moore and Hularis, 1975). This set of inputs may be involved in reinforcement (Claviea and Routtenberg, 1974).

The axons of pyramidal cells in CA1 and CA3 constitute the output of the hippocampus. The CA3 pyramidal cells project to the septum (Raisman, Cowan and Powell, 1966). These cells may also project to the anterior thalamus (Raisman, et. al., 1966) and to other regions of the forebrain and brainstem (Siegel and Tassoni, 1971; Raisman, et.

al., 1966). The pyramidal cells in CA1 may project to the septum (Raisman, et. al., 1966; De France, Kitai and Shimono, 1973) or may send axons exclusively to the subiculum (Andersen, Bland and Dudar, 1973). The ventral CA3 region may also project to the entorhinal cortex (Hjorth-Simonsen, 1971). It should be noted that the results concerning the output of the hippocampus are open to controversy, probably due in part to species differences and to differences between dorsal and ventral regions. These pathways are the subjects of continuing study.

Many behavioral studies have been done aimed at elucidating the function of the hippocampus. One group of studies has involved the recording of hippocampal slow waves and the determination of their relations to behavior (Vanderwolf, Kramis, Gillespie and Bland, 1975; Grastyán, Lissák, Madarász, and Donhoffer, 1959; Elazar and Adey, 1967). Another approach involves the placing of lesions in the hippocampus and subsequently testing to find deficits in behavior (see Isaacson, 1974, for a discussion of lesion experiments in the hippocampus). In a third group of studies, recordings of hippocampal unit activity were taken in a variety of circumstances to determine what single cells might be signalling (Olds, Mink, and Best, 1969; Segal, 1973; Ranck, 1973; Vinogradova, 1975). Interpretations of results of these three types of investigations have lead researchers to posit that the hippocampus is involved in a variety of functions, among which are attention (Grastyán, et. al., 1959), learning and memory (see Izquierdo, 1975, and references therein), hypothesis formation and testing (Isaacson, 1974), and the organization of a cognitive

map of the environment (O'Keefe and Dostrovsky, 1971). Clearly, then, there is little agreement as to what the hippocampus does in overall brain operation.

Two courses of experiments have indicated, however, that the hippocampus is a region that displays plasticity. Studies of the activity of single cells in intact animals have shown that the hippocampal responses to external inputs change during conditioning and extinction (Olds, et. al., 1969; Segal, 1973). Though such changes occur throughout the brain (Hirano, Best, and Olds, 1970), Vinogradova (1975) has indicated that certain regions of the hippocampus itself must contain plastic elements. This conclusion is supported by the results of physiological experiments with intact animals and with isolated hippocampal slices. As mentioned above, direct electrical stimulation of the perforant path results in excitation of granule cells in the dentate gyrus. It has been shown (Bliss and Lømo, 1973; Teyler and Alger, 1976) that following tetanic stimulation of the perforant path its effectiveness in activating granule cells is increased for periods of days or weeks. Lower-intensity and lower-frequency stimulation lead to decreased effectiveness (Alger and Teyler, 1976). The mossy fiber and Schaeffer collateral fiber systems also appéar to show post-tetanic potentiation (Teyler and Alger, 1976). Though these results were obtained under non-physiological conditions, they do indicate that the hippocampus does have a potential for plasticity in physiological situations.

Model System Structure

The possibility that hippocampus is involved in learning and memory and in the direction of behavior at cognitive levels has attracted a number of modelers to this brain region. Models exist that treat the hippocampus (or hippocampus-like circuits) as a temporary memory system (Marr, 1971; Olds, 1969), as a pattern transformer (Mittenthal, 1974; Grossberg, 1973), and as a trainable decisionary system (McLardy and Kilmer, 1970; Kilmer and Olinski, 1974; Wigström, 1974, 1975). Models of hippocampus have also been proposed that deal with its circuit properties, independent of learning (Horowitz, Freeman and Stoll, 1973; Dichter and Spencer, 1969). None of these models deals directly with the possibilities of temporal sequence processing in the hippocampus. As discussed previously, none of the available temporal sequence models can be readily applied to the structure of the hippocampus.

The model presented here is designed around the structure of the dentate gyrus. I assume here that a granule cell in the dentate gyrus that is excited by a system input spreads excitation to neighbors in its lamella, causing a wave of activity to travel down the lamella. Details of the cell coupling and dynamic properties used to produce waves are discussed in the next section following presentation of the model's overall structure in this section.

Like the hippocampus, the model is made up of a number of arrays of cells called lamellae. Each model lamella consists of four lines of neuromimes as indicated in Figure 5.2. G-neuromimes receive the

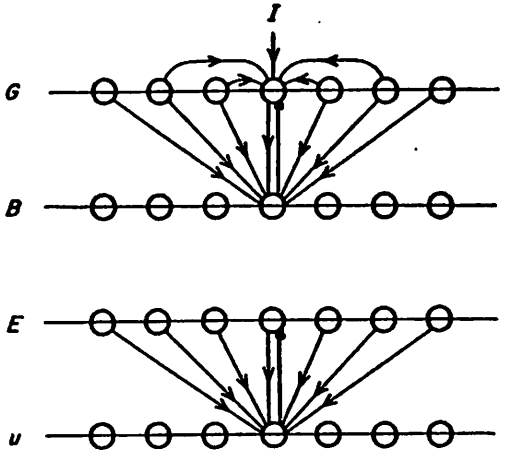


Figure 5.2 Patterns of interconnections between G and B neuromimes and E and U interneuromimes within a lamella. Arrows indicate excitatory influence, dots indicate inhibitory influence.

system inputs and their outputs constitute the outputs of the lamellae. The G-neuromimes are coupled to one another and to B-neuromimes and E-neuromimes in their lamellae. B-neuromimes are excited by the G's and in turn inhibit them. Patterns of connection and connection strengths between G-neuromimes and B-neuromimes are fixed and identical for all G's and B's. Details of these connections are discussed in the following section on wave generation. Coupling between E's and U's is similar to that between G's and B's, with the exception that E's are not coupled to one another. These connections are also fixed and identical for all E- and U-neuromimes.

Connections between G's and E's are assigned with some randomness in the model. Each E is connected to a fixed number of G's in its lamella. These neuromimes are chosen at random from a range of G's centered around the E, as indicated in Figure 5.3a. The strengths of these connections are fixed and identical for all connections. Each E in turn gives rise to two lines, called Z-lines, that run perpendicular to the lamellae, making contact with G's as indicated in Figure 5.3b. These lines represent the fibers of Zimmer (1973) that run from CA3 back to the dentate gyrus. The length of each Z-line is chosen according to a specified probability distribution, F_z. G's in each lamella the Z-line passes through are chosen for contact with fixed probability to a maximum range on either side of the Z-line. These longitudinal lines are the only connections between lamellae, and have the only variable strengths used in the model, as discussed in section four.

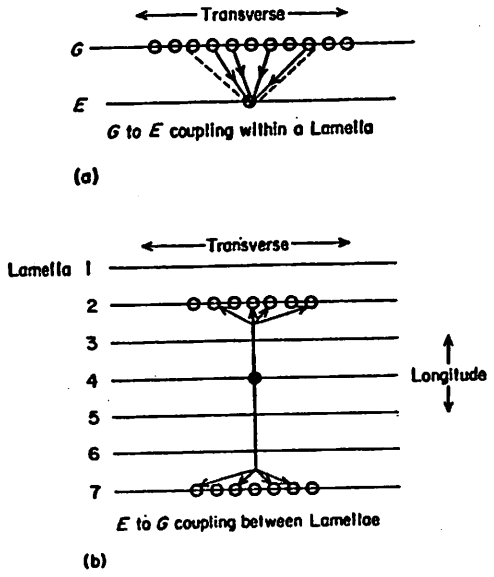


Figure 5.3 Patterns of interconnection between G and E neuromimes. (a) Connections from G to E within a lamella. Dashed lines indicate range of possible contact. (b) Connections from E to G in other lamellae.

Generation and Propagation of Waves in the Model

The wave-supporting substrate of the model has a configuration suggested by hippocampal physiology. As discussed earlier, pyramidal cells in CA3 and granule cells in the dentate gyrus receive excitatory influences from the hippocampal input pathways (Andersen, Bliss, and Skrede, 1971). These cells in turn excite a variety of interneurons whose axons play back onto other pyramidal and granule cells. In both regions, a type of interneuron called a basket cell is assumed to have a profound inhibitory effect on pyramidal and granule cells (Andersen, Eccles, and Loyning, 1964). Basket cells have widely branching axonal arborizations, so excitation of a basket cell by one pyramid or granule cell will depress the activity of others nearby. In CA3, it has been shown that pyramidal cells also excite neighboring pyramids (Lebovitz, Dichter, and Spencer, 1971). We assume here that granule-to-granule excitation also exists in the dentate gyrus, leading us to the circuit configuration for G's and B's shown in Figure 5.2.

The generation of single impulses in nerve cells is a highly nonlinear process involving a buildup of excitatory potentials in dendritic membranes and the subsequent triggering of a propagating action potential (Shepherd, 1974). However, the rate of firing of a single cell and the average firing level of a homogenous population of cells may reasonably be modeled in terms of simple dynamic equations (Amari, 1972; Stein, et. al., 1974; Wilson and Cowan, 1973). Accordingly, each of my model cells obeys a first-order differential equation, the value of which may represent the firing rate of a sin-

gle neuron, or the number of active cells in a population of neurons. I refer to the model cells below in terms of output pulse rates of single cells.

I assume that the output pulse rate of a cell is given by the difference between the total excitatory influence on the cell and some function of the total inhibitory influence on the cell. An initial set of equations for the operation of the G and B cells under these assumptions is:

$$\tau_G \frac{dG_j}{dt} = -G_j + \sum_m \left[G_m^o - \theta_m\right]^+ \cdot w_m + I_j$$

$$\tau_B \frac{dB_j}{dt} = -B_j + \sum_k \left[G_k^o - \Gamma_k\right]^+ \cdot v_k$$

$$G_j^o = {}^M\left[G_j - B_j\right]_m$$

where

G_j represents the total excitatory influence on the G-cell,

B_j represents the B-cell output,

τ_G and τ_B represent membrane time constants,

θ_m and Γ_k are connection thresholds,

w_m and v_k are connection weights,

I_j is the external input to the G-cell,

G_j^o is the output pulse rate of the G-cell,

$$\left[x\right]_m^M = \begin{cases} M; & x > M \\ x; & m < x < M \\ m; & x < m, \end{cases}$$

for $m < 0 < M$, and

$\left[x\right]^+$ is x if $x > 0$ and is zero otherwise.

Note that the output rate of the G-cell is constrained to lie between a maximum value greater than zero and a minimum less than zero. The maximum reflects the fact that nerve cells cannot fire more rapidly than some maximum rate. The minimum is set to a value less than zero under the assumption that the zero level in the model represents a non-zero spontaneous firing rate in the actual nerve cells. Then the minimum value in the model corresponds to a firing rate of zero in the actual cells.

Each G and B cell is driven through weighted threshold connections by the output rates of neighboring cells. That is, neither is influenced by any neighbor to which it is connected until that neighbor begins to fire above a specified rate. The influence of neighbor is weighted by a constant associated with the connection. Note that neighbors firing below threshold rates and at rates below the spontaneous level do not lead to depression of the firing of the cell. Hence cells that are firing below the spontaneous rate are effectively uncoupled from the system. I assume then that information processing and maintenance of a background spontaneous firing rate (the system equilibrium state) are two separable factors in the system dynamics.

A simulation of this system of equation has been carried out using an integration routine written in FORTRAN. The propagating waves illustrated in Figure 5.5 were generated with the connection template shown in Figure 5.2 and the associated weight and threshold profiles of Figure 5.4. All cell values were set initially to zero, representing an undisturbed system displaying spontaneous firing. Following external stimulation at the center of the system, each sufficiently excited G-neuromime stimulated neighboring G's according to the G-G weight and threshold profiles and inhibited a wider range of G's according to the G-B profiles. The indicated weight and threshold settings allow the excitation of each G to build to a high enough level to excite further G's before being countered by a heavy long-lasting inhibition arising from the central weight and threshold of the G-B profile. This central connection corresponds to a strong self-inhibition activated when the cell fires enough.

A wave may travel to the ends of the lamella and die there, or may die before reaching the ends. In the first case, the wave propagates with a constant shape until the end of the lamella is reached. At that point, there are no further G cells to excite and the wave dies. Because of the connection thresholds, the zero level of activity is stable, so no further activity arises in the lamella until a new input is presented. In the second case, inhibition builds that eventually stops the wave. Here, the wave's amplitude decays as it travels until further propagation is impossible. The distance traveled depends on the strength of inhibition and on the relative B time constant. The full memory model to be described below uses waves that die after going a short distance rather than those that

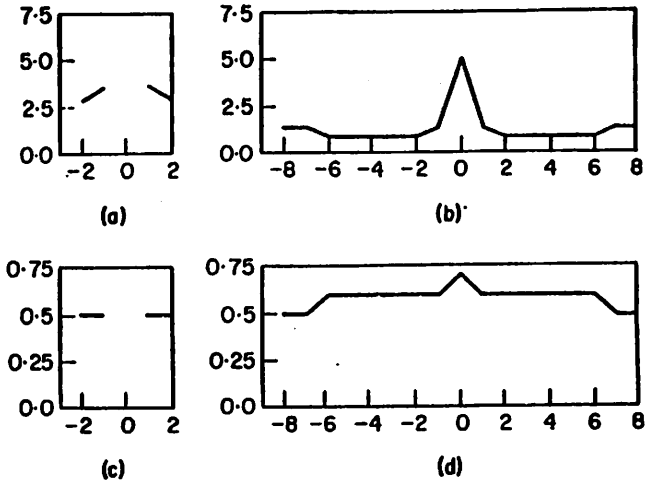


Figure 5.4 Initial weight and threshold profiles. Ordinates indicate number of neighbor relative to cell. (a) G to G weights. (b) G to B weights. (c) G to G threshold. (d) G to B thresholds.

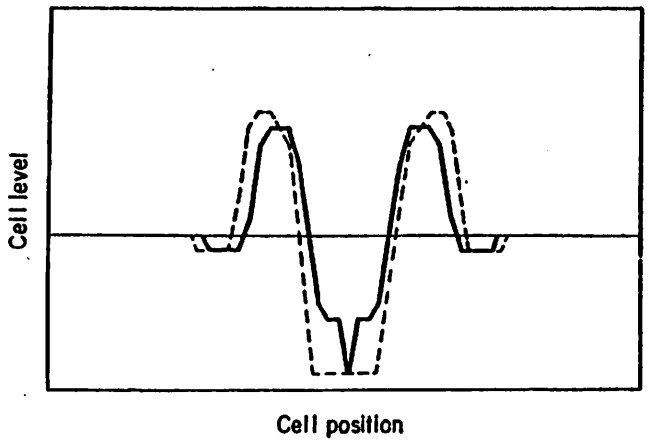


Figure 5.5 Wave produced with profiles of Figure 5.4. Dashed lines indicate progression of wave.

propagate unchanged.

The strong self-inhibition employed above results in a circuit that for long periods after passage of a wave is unable to support another wave. Too long a period of depression decreases the chance that new inputs can enter the lamella, as will be described in section four, so means of shortening the depression time were sought. Two ways to shorten the depression time are to decrease the level of inhibition needed to produce waves and to decrease the inhibition time constant. In order to decrease inhibition levels, an amplitude-dependent nonlinearity is used to allow low levels of inhibition to have a greater effect on the circuit. The amplitude-dependent gain appears in the computation of the input rate as follows:

$$G_j^o = {}^{M}\lceil G_j - (P + W \cdot F(B_j)) \cdot B_j \rfloor_m$$

where

G_j^o, G_j, and B_j are as before,

P is a bias term set close to unity,

W is a weight set greater than one, and

F is the logistic function given by

$$F(x) = \frac{1}{1 + ce^{-d(x-x_0)}}$$

and illustrated in Figure 5.6.

For small values of inhibition, G's output rate is approximately the difference between excitation and inhibition as before. As the

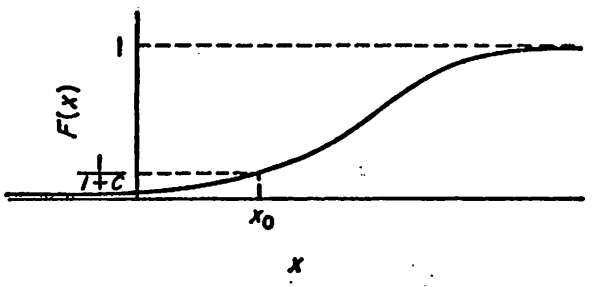


Figure 5.6 Logistic function.

level of inhibition rises, however, its weighting in the output rate calculation increases from unity to 1 + W. In this way low levels of inhibition may have a great effect on the system. The lower values of inhibition decay to zero sooner, leaving the system depressed for a shorter period of time.

A smaller inhibition time constant may be used if different weight and threshold profiles are employed corresponding to a different way of producing waves. With the modified weight profiles, self-inhibition begins to build when the G-cell reaches lower levels of firing, due to the smaller self-inhibitory connection threshold, but builds more slowly than before due to the decreased weight involved. In the previous scheme, waves were produced by allowing cells to build to large firing rates and then depressing them with strong inhibition. In order to produce waves, this inhibition had to last long enough for the excitatory levels to decay below the thresholds necessary for coupling to other cells. Shorter-lasting or weaker inhibition allowed the cells to remain for long periods of time at a level just above the self-inhibition threshold. In order to produce the wave profiles of Figure 5.5, in which each cell fires for a short time and then is silenced, the inhibition time constant had to be great enough to keep the cell depressed until the wave passed out of its neighborhood and its excitatory level decayed considerably. In the new scheme, however, inhibition is applied more slowly, and instead of depressing the cell immediately, pulls the output rate down gradually. Hence inhibitory levels need not remain large for too long, and the inhibition time constant may be

shortened. Figure 5.7 shows waves generated by this scheme. These waves leave the circuit depressed for shorter periods of time, and so are better suited for use in the overall model. Again, these waves may travel to the ends of the lamellae, or may die due to inhibitory buildup before reaching the ends.

The equations employed here are similar to those used by Ratliff (Ratliff, 1965) to investigate the dynamics of lateral inhibition and by Grossberg (Grossberg, 1969) to explore learning and memory in a number of circuit anatomies. In neither case were these equations employed to produce traveling waves of activity, however. The wave generating mechanism in this one-dimensional system is similar to the ones studied by Beurle (Beurle, 1956) and by Wilson and Cowan (Cowan and Wilson, 1973), and Ellias and Grossberg (1975). In each case, the wave peak arises and propagates due to positive feedback between excitatory elements. In Beurle's (1956) studies this feedback is controlled by the refractory properties of the model cells, while in the models of Wilson and Cowan (1973) and Ellias and Grossberg (1975), as in the one presented here, inhibitory buildup stops runaway positive feedback. As stated above, the activity level of a given cell in this model may represent the average firing rate of a nerve cell or the number of cells in a coupled population that are firing at a given moment. The nonlinear inhibition used here may represent a nonlinearity in the way single cells react to inhibition, or the action of a population of inhibitory cells with a distribution of thresholds.

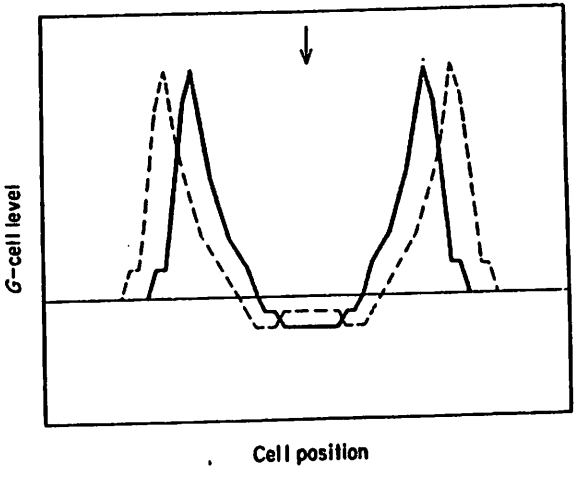


Figure 5.7 Waves produced with modified profiles.

Memory Storage and Recall in the Model

The model is designed to learn and recall both the sets of cells stimulated by system inputs and the time intervals between inputs. It performs this recall by a process of association chaining, in which each input or evoked memory of an input stimulates (after the proper time interval) the set of G's normally stimulated by the next input of the sequence. This set of G's then stimulates the set associated with the next input, continuing the process of recall.

This process is indicated in Figure 5.8 for the case in which the second input immediately follows the first. The model must form associations such that future presentation of the first input will cause immediate stimulation of the second input's G's. This association is handled via the longitudinal lines arising from the E's. The strength of each longitudinal connection is modeled as a first-order differential equation as

$$\tau_z \frac{dZ_{ij}^{km}}{dt} = -Z_{ij}^{km} + \left[(G_{km}^{o}(t) - \theta_G)\right]^{+} \cdot \left[(E_{ij}^{o}(t) - \theta_E)\right]^{+}$$

where Z_{ij}^{km} connects the j^{th} E-cell in the i^{th} lamella with the m^{th} G-cell in the k^{th} lamella,

τ_z is the connection time constant,
G_{km}^{o} is the firing rate of the G-cell,
E_{ij}^{o} is the firing of the E-cell, and
$[X]^{+}$ is equal to x if $x > 0$, and is zero otherwise.

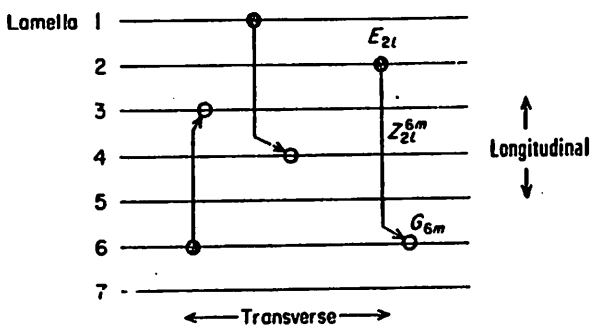


Figure 5.8 Simple association of inputs. Open circles represent G neuromimes excited by present input; hatched circles represent E neuromimes excited by last input's Gs.

Connections are modified according to a form of Hebb modification rule (Hebb, 1949). A connection is strengthened if the E and G it links are simultaneously active above the thresholds θ_G and θ_E, while an unused connection decays to zero at a rate determined by τ_z. Strengths of the longitudinal lines are initially zero, with some exceptions discussed below. A similar connection modification rule was used by Grossberg (1969) in his studies of learning systems.

The thresholds in the modification rule allow the system to discriminate inputs that are to be learned from the waves and from the effects of noise. A system input is assumed to excite a G-cell to a firing level much greater than the maximum reached during passage of a wave. The value of θ_G is set just below this input excitation level to prevent associations from forming between G's excited by waves. Similarly, θ_E is set just below the level of excitation reached by an E-cell when a wave passes by it, so that inputs are associated only with a few E's recently excited by waves.

Thus if an E excited by one of the first input's G's gives rise to a longitudinal line that makes contact with one of the G's excited by the second input, that connection is strengthened according to the learning rule above. Later presentation of the first input will again excite that E, in turn tending to activate the second input's G via the strengthened connection. If the system is properly constructed, as discussed below, each of the second input's G's will with high probability be connected in this way to at least one of the first input's G's. Repeated presentation of the sequence will

result in connections strong enough to excite the G's to the input level of excitation. Presentation of the first input will then cause immediate activation of the second input's G's, effecting recall of the sequence. This chaining process may then continue with the association between the second and third inputs, and so on.

If the next input does not occur immediately, a trace of the first input must be held in the system long enough to form associations as described above, and to code the interval between inputs. Input storage and interval coding are accomplished through the wave action of the G's. Each overall input excites a few G's throughout the system, giving rise to waves moving in the associated lamellae. These waves excite E's as they go, so that an input's G's may be associated with waves generated by the previous input, effecting both storage and interval coding.

This process is illustrated in Figure 5.9. A wave generated by the first input at time T_0 in lamella 4 moves along the lamella, as indicated by the cross-hatched rectangle. Before dieing, the wave will move a distance determined by the inhibitory parameters. In order to handle time intervals longer than those obtainable with a wave moving along a single lamella, some of the longitudinal connections are permanently strengthened initially so that a wave activating such a connection may generate a wave in another lamella. The second wave may continue after the first has died, and may generate more waves. A series of such wave generations is indicated by the dotted lines in Figure 5.9 while arrows show the directions of wave

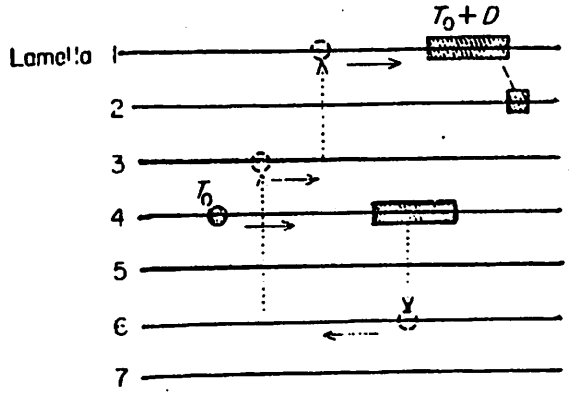


Figure 5.9 Association of inputs and waves.

motion along lamellae. At time T_0 + D a wave generated in this way is at the position shown in lamella 1 when a new input excites a G-neuromime in lamella 2, as indicated by the cross-hatched square. If a longitudinal connection exists between an E excited by the wave at that time and the G excited by the new input, that connection is strengthened according to the system's learning rule. Such a connection is indicated by the dashed line in Figure 5.9.

Now if at some later time the first input is again presented to the system, a wave will be generated at the same place in lamella 4. If this wave causes the same sequence of waves to be generated as before, after an interval of length D, the expected time between the two inputs, a wave will activate the learned longitudinal connection, exciting one of the second input's G's. If each of the second input's G's is associated in this way with at least one wave stemming from the first input, presentation of the first input will cause the activation of the second input's G's after about D seconds. This process may then continue, effecting recall of the sequence.

System Design Constraints

The association system described above can fail in five main ways. The first set of problems concerns the representation of inputs using patterns of moving waves. If the waves all die, or the pattern falls into a cycle, information is lost and proper association becomes impossible. Interference between the wave pattern and the inputs to the network is the second problem. Cells that lie in

the troughs of waves are inhibited, so inputs cannot excite them enough for associations to form. A third difficulty concerns the probability of association. If the system is not properly designed, the probability that a wave can become associated with a given input will be too small for reliable operation. The fourth class of problems concerns interference between wave patterns generated by successive inputs. Finally, the fifth problem arises from the effects of changes in the system's structure brought about by learning. Each of these problems is treated below. In some cases, proper setting of system parameters can alleviate the difficulty. Other problems require that further mechanisms be added to the model.

Pattern persistance and cycling. The first problem arises because all the waves generated by a given input may die out before the next input arrives. This will happen if there are too few opportunities for waves to generate others in other lamellae. Simulations indicate that waves generally persist indefinitely in systems with more than a few prestrengthened connections from each lamella to the others. Second, a given set of waves may fall into a cyclic pattern of movement through the network, thus losing the ability to present the exact interval between inputs. Simulation again shows that in systems with more than a few strong connections between lamellae, wave interactions are complex enough to preclude this possibility. These difficulties, then, will seldom arise in systems with enough prestrengthened connections between lamellae.

Trough interference and association. The second set of problems concerns the interference between a wave pattern and an input, and the probability that all of an input's G's become associated with waves. These two problems are linked through their dependence on the number of waves in the system. First, the input may stimulate a cell that lies in the trough of a wave. Such a cell is inhibited and cannot reach a firing level that allows it to become associated with other waves or to generate a wave itself. That part of the input, then, cannot enter the system and is unavailable for association with other inputs. Second, a stimulated cell may not become associated with any wave in the system. This will occur if there are too few waves or if there are too few longitudinal connections available for association. Formulas dealing with these two sources of failure are derived in Appendix C. There, Table C.2 shows the result of applying the formulas to a system with parameters given in Table C.1, in which waves and stimulated cells are distributed at random. It may be seen that full input entry and full association in such a system are possible only when the numbers of waves and stimulated cells are kept quite low with respect to the total number of cells. The number of waves present in the system depends on the number of cells initially stimulated and on the number of prestrengthened longitudinal connections between lamellae. With too many such connections, too many waves are spawned and input entry becomes a problem. With too few connections, however, the wave pattern may die or fall into a cycle, as discussed previously. Simulation shows that systems with enough prestrengthened connections to avoid cycling and pattern extinction generate far too many waves to

ensure that none of an input's G's is inhibited. Further, the number of waves exhibits oscillations, as shown in Figure 5.10, making the system at times able to accept inputs and unable at other times. Accordingly, means were sought to hold the number of waves at a constant low level, while still avoiding pattern extinction and cycling.

The overall number of waves in the network can be controlled through regulation of wave death. Recall that waves are assumed to travel a certain distance (called the propagation distance), then die due to inhibitory buildup. A feedback mechanism was added to control the number of waves by changing this distance. The propagation distance may be the same at all points in the system, corresponding to an overall setting of inhibitory parameters, or may vary locally, corresponding to local parameter setting. In either case, when the number of waves in the system is small, the propagation distance is made large. This insures that waves will travel far enough to generate other waves before dieing, thus keeping the overall wave pattern from dieing. As the number of waves grows, the propagation distance is decreased to keep waves from generating too many others. In the case of global propagation distance variation, the propagation distance at each point is adjusted according to the total number of waves in the system. With local variation, the distance is set at each point according to the number of waves present in an area about that point.

Figure 5.11 illustrates the result of applying global propagation distance variation to a system with the same parameters as those

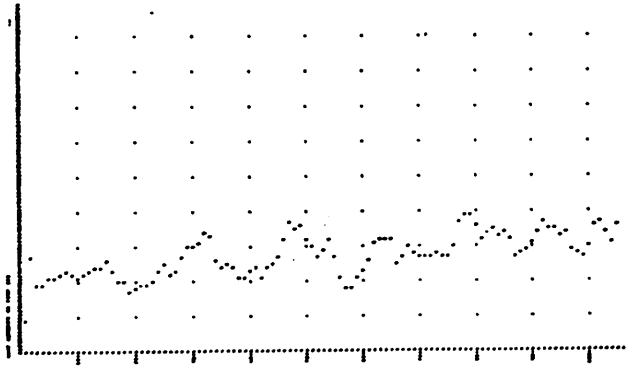


Figure 5.10 Plot number of waves produced in uncontrolled system as a function of time.

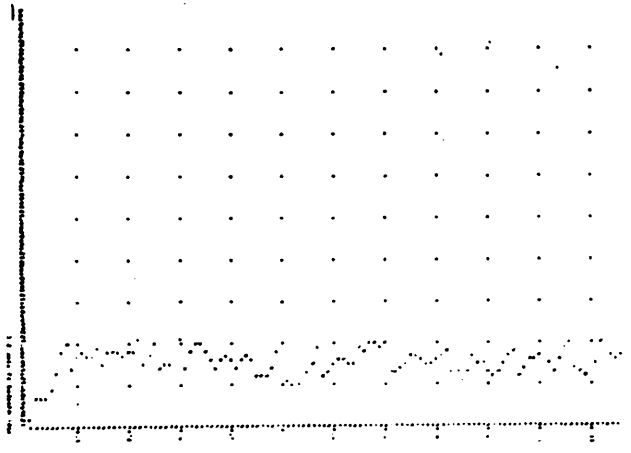


Figure 5.11 Plot of number of waves produced in controlled system as a function of time.

used in Figure 5.10. The number of waves is considerably reduced and is more constant than before. Similar results are obtained using local variation. Local variation has the added advantage that waves are able to spread more evenly through the network, leading to increased local probabilities of association. This method of propagation distance variation thus allows the number of waves to be controlled to satisfy the input entry and association constraints and to keep the wave pattern from dieing entirely or falling into a short cycle.

Inter-sequence interference. Another cause of difficulty is the interference between wave patterns generated by different inputs or by successive presentations of the same input. As an example, consider the sequence formed by regular repetition of a single input, and suppose that there are already waves in the system when this input is first presented. These waves will interact with those generated by the new input, so that the pattern that exists when the input is presented the second time may not be the same as the initial wave pattern. Because of this difference, the third presentation of the input is likely to become associated with still another pattern, and so on. This difficulty stems from the instability of the wave patterns, which is such that two patterns that differ only slightly give rise over time to patterns that are more and more different. Because of this instability, it is possible that a repeated input will not generate the same wave pattern twice over very many repetitions, and hence that no strong connections will form in a short time. The network would thus have stored a large number of memories

of this input and would take a prohibitively long time to train. Even if a short sequence of patterns were formed, if the input was presented with a different initial pattern during recall, the learned sequence of patterns might not arise again, leading to a recall failure.

The first mechanism added to combat this interference problem is one that inhibits all waves when a peak of maximum activity is detected anywhere in the system. This corresponds to a strong blanket inhibition applied when an input is presented or when a strong association is recalled. It's assumed that the strong peaks of activity are able to survive this inhibition and to generate new waves of activity. With the wave-killing mechanism in operation, interference effects are limited to two forms. First are the effects of inhibitory troughs left behind by waves killed when the input was presented. Second is the effect of lingering longitudinal connection inhibition.

Wave troughs give rise to the input entry problem, as discussed earlier. Troughs left after waves are killed also block the propagation of waves generated by new peaks of activity. The effects of these forms of interference may be minimized by use of the proper numbers of waves. Longitudinal connection inhibition left after the wave pattern is killed can affect the development of the new pattern by blocking wave generation. As above, the instability of the wave patterns is such that further patterns generated without the waves whose generation is blocked will differ from patterns generated with those waves. Because of the differences, associations will not readily form and recall may be impossible. A way to obviate this difficultly is to assume that E's have a maximal inhibitory level, and to set all E's to this level when the wave killing mechanism is activated. This corresponds to a saturating inhibitory input applied at that time to all E's. This input effectively eradicates all traces of the E-inhibition due to the previous pattern, allowing the new pattern to develop with interference due only to the wave troughs. The two mechanisms of wave killing and E-inhibition input and saturation can be used to minimize the effects of interference between wave patterns due to successive inputs.

Interference effects of learning. The final difficulty lies in the effects of changes in the system's structure brought about by learning. Because waves travel throughout the system between input presentations, a longitudinal line that has been strengthened through learning is likely to be activated at times when it should not be. Activation of a strong enough connection at the wrong time will cause a wave to arise in the system at a point in the development of the wave pattern where none had been before. Such extra waves will alter the development of the pattern, making proper association and recall impossible.

A way to ensure that longitudinal lines strengthened through learning are activated only at the proper times is to allow E activation only when a particular pattern of waves exists in the network. In this scheme, an E that gives rise to no strengthened connections, that is, one that has never taken part in association or that has

forgotten any associations ever made, may be activated by a single wave that goes by its position, as before. When an E takes part in association, however, that E learns the pattern of waves in an area around its position in the network. Thereafter, the E can be activated only when the wave pattern in the network matches its stored pattern. If the stored pattern is composed of enough waves, it is unlikely to occur at any but the correct time.

This mechanism requires a change in the way the E's are activated. Each E must have strong connections from G's in its lamella, and weak connections from G's in other lamellae. All of these connections must be trainable, such that when a longitudinal line stemming from the E is strengthened, any connection to the E from a G active at or above the wave propagation level is also made strong. At the same time, the E's threshold of activation must increase, so that the entire set of strengthened connections must be excited to activate the E. In this way, a particular wave pattern is stored in the connections to the E from G's in its own and neighboring lamellae.

These mechanisms added to the basic model, then, minimize the effects of the problems arising from fluctuations in the number of waves in the network and from the instability of the wave patterns. The number of waves must be kept small enough to allow inputs to enter the system freely, but large enough to allow proper association and to keep the wave pattern from dieing or cycling. The number of waves is controlled through feedback that decreases the wave propagation distance as the number of waves increases. Pattern instability gives rise to problems of interference between patterns generated by successive inputs and to disturbances due to learning itself. Interference problems are minimized using a wave inhibition mechanism and a strong longitudinal line inhibition, both of which operate when an input is presented to the network. Pattern disruption due to improper activation of longitudinal lines strengthened through learning is prevented by allowing longitudinal lines to be activated only when specific patterns of waves are present in the system. These mechanisms are related to changes in the operation of the basic model, and may be interpreted in terms of the action of particular excitatory and inhibitory mechanisms in the hippocampal system. The means of simulating this system, and results of simulation, are discussed below.

System Simulation

The differential equation formulation of the model, described in section 2, has the disadvantages of being difficult to control and expensive to simulate. It is difficult to know in advance what detailed characteristics the wave generating mechanism must have so that the overall system can function properly. Much work would be needed to find ways to realize a given set of desired characteristics in turns of the dynamic equations. Furthermore, a large system of such equations can be simulated only at considerable expense. In order to simplify the simulation and to make the system easier to control, simulation of the full model was carried out using cells modeled as pairs of finite-state automata, rather than as differential equations.

Each simulated cell consists of a wave-generation portion, representing the action of the G and B cells, and a longitudinal line activation portion that represents the E and U cells. The operation of each cell's wave-generation automaton reflects the generation of wave activity in the differential equation form of the model. These automata are much more easily controlled than the differential equations, however, and may easily be designed to have desired wave propagation characteristics. In this formulation, the entire wave peak is represented by a single cell in a specified state. Each cell therefore represents a number of cells in the differential equation formulation, making this simulation more economical. Similarly, all of the E's activated by a wave peak at a given position in the system are now modeled in the single longitudinal-activation automaton associated with the cell at that position. The structures of these automata are described below.

The wave-generating automaton. The wave-generating automaton of each cell is shown in Figure 5.12. Each state is numbered, and larger numbers represent higher levels of firing in the differential equation formulation. State Q is the resting state, in which a cell remains if undisturbed and to which it returns following perturbation. The highest state, I, represents the level of activation reached when a system input excites a quiescent cell. States W and T are the wave propagation and maximal trough states, respectively, representing the level reached as a wave peak passes the cell and the level of maximal inhibition following the peak. States between the trough state and the quiescent state are trough recovery states, through

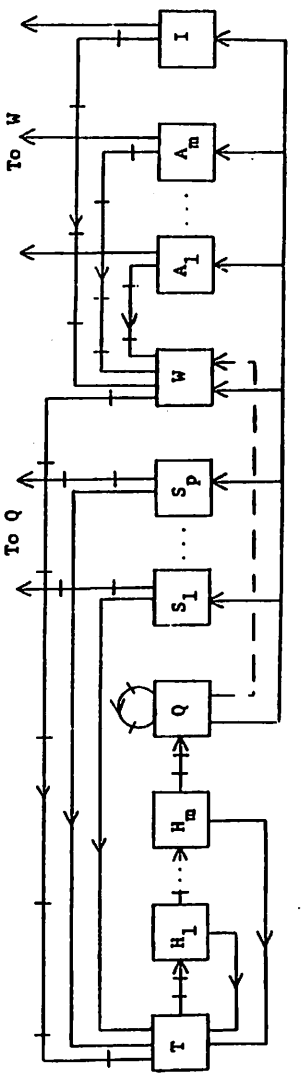


Figure 5.12 Structure of wave-generating automaton. —|— indicates action with no input, — — — action of wave input from neighbors, —— action of input applied to non-quiescent cell. Multiple arrows indicate that state entered depends on size of input.

which a cell goes to quiescence following passage of a wave. The states between the quiescent state and the wave state are states of subactivation to which a cell is sent by small inputs. Finally, the states between the wave state and state of maximal activation are superactivation states representing levels great enough to generate waves but smaller than the state to which a cell is sent by a system input.

Each cell has three sources of input. These inputs are wave-activation inputs from neighboring cells in the same lamella, system inputs, and longitudinal-activation inputs from cells in other lamellae. If a cell is quiescent, that is, if the cell's wave-generating automaton is in state Q, a system input to the cell moves it to state I. From I, the cell goes in one time-step to cell W, the wave state. From state W the cell goes to the maximal trough state, T, at the next time-step, regardless of further inputs. This corresponds to a strong inhibition triggered by the cell's recent activity. If the cell receives no further inputs it moves one state at a time from T to the quiescent state. Any input to the cell while it is in one of the trough states causes it to go back to T again, representing the effect of further inhibition.

In order to propagate waves, each cell receives inputs from the two nearest neighbors in its lamella. A quiescent cell enters the wave state at a given time when one of its neighbors was in the wave state or greater at the time before. The cell next enters the lowest trough

state, as above. In this way, a wave moves through a quiescent region represented by a cell in the wave state followed by cells in the trough states. The single cell in the wave state represents the entire wave peak in the differential equation form of the model. A cell in any trough state is sent to the lowest trough state at the next time-step if one of its neighbors is in the wave state or greater. Thus when two waves collide they annihilate one another, leaving behind only an inhibited region that recovers to quiescence. This action reflects the behavior of colliding waves in the dynamic formulation of the model.

The third source of inputs to a cell is activity communicated over the longitudinal lines from cells in other lamellae. The activity level of a given longitudinal line is an integer calculated in a manner described below. The activity levels of all longitudinal lines to a given cell are summed at each time-step, together with the external input to the cell, to arrive at the total extrinsic input to the cell at that time. If a cell is in any state other than the quiescent state or a subactivation state when it receives a nonzero extrinsic input, it is sent to the lowest trough state at the next time-step. This corresponds to a strong inhibition activated by the cell's recent activity and reinforced by further inputs to the cell. A small input to a quiescent cell will set the cell to a subactivation state, from which it goes to quiescence if it receives no further inputs. Larger inputs set the cell to the wave state or to a superactivation state. The superactivation states decay to the wave state in one step with no further inputs. An input to a cell in a

superactivated state sends it to the lowest trough state at the next time. A strong enough total input can send a quiescent cell to the maximal activation state. In this way, the activation of longitudinal lines can result in the generation of new waves and strong peaks of activity in the system.

The longitudinal-activation automaton. A longitudinal line is activated as a result of the action of both portions of the cell that gives rise to the line. The structure of the longitudinal-activation automaton, denoted LA, is shown in Figure 5.13. The LA automaton has a quiescent state, Q, an activated state, A, and a set of trough states, $T_1 \cdots$ Each LA automaton receives an activating input from the wave-generating automaton in its own cell, and inhibiting inputs from the LA automata associated with neighboring cells in its own lamella. Like the E's in the dynamic form of the model, an LA automaton with no preferred wave pattern can be activated by a wave that passes by its position in the lamella. The LA automaton is inhibited following its own firing or the firing of neighboring LA automata.

To illustrate LA activation, first consider a cell that has no preferred activation pattern, that is, one that has not recently taken part in an association. If the wave-generating portion of that cell is at or above the wave state, and the cell's LA automaton is not in an inhibited state, then the LA automaton is sent to the activated state and the longitudinal lines stemming from the cell are activated. At the same time, the LA automata of neighboring cells receive inhibitory inputs depending on their distances from the cell whose lines

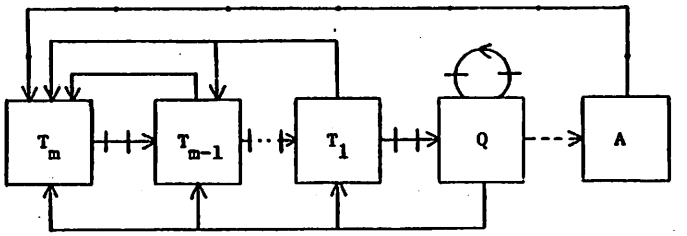


Figure 5.13 Structure of the longitudinal-activation automaton. +++ indicates action with no input, – – – action of inhibitory input from neighbors, ——— action of self-inhibition. Multiple arrows indicate state entered depends on size of input.

became active. Inhibitory inputs are additive, and the greater the total inhibitory input at a given time, the lower the inhibition state to which the LA automaton is sent at the next time. If left unperturbed, the LA automaton recovers one state at a time to quiescence. Note that an inhibited LA automaton can't activate its longitudinal lines or produce inhibition in other LA automata. In this way, a wave of activity alternately activates and inhibits LA automata as it travels along its lamella. This sequence of events is the same for cells with preferred wave patterns, except that activation can occur only if the automaton is quiescent and if the proper wave pattern exists in the network.

The level of activity transmitted from one wave-generating automaton to another through an activated longitudinal line depends on the state of the activating wave-generating automaton and on a weight value associated with the connection between the two automata. In the dynamic form of the model, these weights conform to differential equations. Here, they are modeled as difference equations with values that are increased according to the learning rule presented in section 2, and that fall to zero otherwise. Note, however, that the prestrengthened connections that allow a wave in one lamella to generate waves in others have non-zero values that are fixed for the life of the network. The input to a cell in another lamella due to an activated longitudinal connection is simply the connection weight multiplied by the level above quiescence of the activating wave-generating automaton. Recall that all such inputs are summed at the receiving cell to determine its state at the next time-step. So a

wave in one lamella that activates a strong-enough longitudinal connection can send a cell in another lamella to its highest state, while somewhat weaker connections allow waves to generate other waves in different parts of the system.

This formulation of the model in terms of automata is an abstraction from the dynamic form that is designed to be easier to control and less costly to simulate. Enough of the operation of the differential equations has been embodied in the functioning of the automata, however, that the results obtained with this form of the model can also be achieved using the dynamic system form. Results of simulating the automaton form of the full model are described below.

Simulation Results

A simulation of the system in the automaton form described above was carried out to investigate the effects of interference and to determine the network's memory capacity. Two networks were created that differed in the details of their structure and in their wave control parameters. In order to ensure that associations were made properly and that recall errors were minimized, two further constraints were imposed on the operation of these networks. First, when an input entered the system, the network was searched to find at least two cells active at or above the wave propagation level at the time before that did not already give rise to strengthened connections. These cells were then connected to the cell excited by the input, and the connections were given non-zero values. At the same time, the wave pat-

tern about each of the cells chosen for connection was searched until more than a fixed minimum number of waves were found. The positions of these waves were then recorded to act as the activating configurations for the cells chosen for association.

The first condition ensured that if there were any free cells activated by waves when an input entered the system, then the input would be associated with at least one wave in the network. This corresponds to a network in which each E gives rise to a large number of longitudinal lines ready for association. The second condition ensured that the activating wave configurations were large enough to be unique in the course of development of a wave pattern. Simulations showed that if the number of elements in an activating configuration was more than five or six, the activating pattern was unique to only one overall wave configuration. The first condition may be relaxed to allow associations to form only with some probability. Likewise, the second condition may be changed by allowing a fixed-area search for waves to use in forming the activating configurations. Under the relaxed conditions, however, association failure is possible, and activating configurations too small to be unique may be formed. These sources of error were eliminated here in order to investigate memory capacity and the effects of errors due to interference.

Table 5.1 lists the parameters of the two simulated networks. In each case prestrengthened longitudinal connections were chosen according to an exponential probability distribution. In these simu-

TABLE 5.1

Parameters Used in Simulations of Memory Networks

Network Numbers	N_L	N_C	R	L	D_{max}	D_{min}	a_1	a_2	P_z	P_{zw}
A	9	59	20	4	25	2	12	4	.4	.0045
B	9	59	20	6	25	4	6	2	.4	.0045

ulations, the values of these and all other strengthened longitudinal connections did not decay, so that memory capacity could be evaluated. Connection strengthening parameters were chosen such that four occurrences of strengthening were required to make a connection strong enough to be able to generate a peak of maximal activity. Each system input consisted of two cells chosen from all cells of the network according to a uniform distribution. Input sequences were composed of two inputs repeated at fixed intervals chosen with uniform probability from a range of between thirty and fifty time-steps.

Wave propagation distance was set according to the equation below:

$$D = \max [D_{min}, D_{max} - a_1 \cdot \lfloor N/a_2 \rfloor]$$

where D is wave propagation distance at a given point in the network,

D_{min} and D_{max} are the minimum and maximum allowable propagation distances,

N is the number of waves in a rectangular region centered about the point in question, R cells wide and L lamellae deep, as indicated in Table 5.1,

a_1, a_2 are constants, and

$\lfloor m \rfloor$ denotes the largest integer less than or equal to m.

With this rule, the wave propagation distance decreased in steps of size a_1 for each increase in N of size a_2. In network A, wave control was established on the basis of the number of waves in a region about each cell that is smaller than the region used in network B. In B, the distance was decreased gradually as the number of waves

increased, while in A the decrease was more abrupt due to the larger va;ies pf a_1 and a_2. Wave control was better in A than in B, leading to a more nearly constant number of waves in A. This difference in the wave control effects is reflected in the network's storage and recall success rates, as described below.

Both networks were trained as follows. At the start of each training session, four inputs were chosen to present to the network. Two of the inputs were to form the sequence to be learned and two were to act as "temporal context" during learning and recall. The first context input was presented to the network, and the wave pattern due to that input was allowed to develop. After a time, the first input of the sequence to be learned was presented, followed at the appropriate time by the second input, and then again by the first. This sequence was repeated several times to allow strong associations to form. To test the network's ability to recall the sequence, all waves were then erased from the system and the second context input was presented. As before, the pattern due to this input was allowed to develop for a time and then the newly-learned sequence was presented twice. If the network had learned the new sequence properly, and if interference problems did not occur, the network would continue to regenerate the sequence following cueing. After a complete training session, successful or not, the network was saved to be used in further training. In this way, the effects of storing many sequences in a single network could be evaluated.

Network training was stopped when errors occurred in three successive sequences. These errors could be failures of association or interference from previously stored sequences. After the networks were trained to capacity according to this criterion, each was tested for recall of its stored sequences in order to test for interference between stored sequences. In this testing, the context input was presented as in learning, and the sequence to be recalled was presented twice. As before, successful recall meant that the network would continue to regenerate the sequence after cueing was stopped. As in training, recall could fail due to lack of proper associations or to interference effects.

Results of these simulations are shown in Table 5.2. Approximately the same number of sequences were stored in each of the two networks, and each formed about the same number of strong associations and wave-jump connections. In most cases, recall was established after one or two presentations of the inputs, that is, after about one presentation of the entire sequence. In each network, failures occurred due to context interference (interference from the pattern of waves generated by the previous input), stemming in each case from a context failure during learning. Similarly, association failures during training gave rise to recall failures.

Interference between stored sequences also occurred in each network. In network A, three sequences composed of similar inputs interferred with each other, leading to a recall composed of parts of all three when any was used as a cue. Two other traces in net-

TABLE 5.2

Results of Memory Network Simulations

Network Number	No. Jump Conn.	No. Assoc. Conn.	No. Seq. Stored	No. Interference Failures	No. Assoc. Failures	Success Rate
A	110	262	23	5	2	70%
B	130	235	24	5	4	62%

work A interferred with one another. Two pairs of sequences in network B interferred with one another. In one case, a sequence disrupted the recall of another, but was not itself disturbed. In the second case, a sequence that was not fully stored due to an association failure disrupted recall of another sequence.

In no case did the wave pattern cycle between input presentations. The wave pattern died entirely in only one case. The cause of death was blockage of the new input's waves by the troughs left behind by the previous pattern.

Counting all forms of failure, network A had a success rate of 70%, while B's success rate was 62%. Success rate is measured as the percentage, among all sequences presented to the network before the final three failures, in which there was proper storage and recall. Network A was more successful than B because wave control was better in A. Failures in B often occurred because there were too few or too many waves in the system. The control in A was such that there was more often the proper number of waves to allow association without blocking the inputs.

If failures due to interference from stored sequences are not included, the success rates are 85% and 68%, respectively. Mechanisms may be added that diminish the effects of stored-trace interference by actively erasing the older sequence's connections when interference occurs. Such mechanisms would then increase the success rate of the network, at the expense of the loss of older memories.

Association failures occur when an input enters the system at a time when all waves are at positions that already give rise to strengthened longitudinal connections. This source of error can be minimized by changes in system architecture or by allowing the network to store only up to about half its expected capacity. Passive forgetting due to decay of connection weights and spontaneous loss of stored activation patterns may be employed to erase old memories, keeping the network always at about half capacity. Again, the chance of error is minimized at the expense of the loss of old memories.

Discussion of Results and Applications to Habituation and Conditioning

A temporal sequence memory model based on the structure of the dentate gyrus region of the mammalian hippocampus was proposed in section 2. This memory uses nonlinear propagating waves to generate a representation of its input to effect memorization of a sequence of inputs together with the time intervals between them. Mechanisms were added to this memory model to overcome sources of interference and instability, and the full system was simulated to test the effects of errors and to determine the memory's capacity. Although several types of errors occur in the operation of the network, success rates of 70% can be achieved using proper control of the number of waves in the network. Elimination of interference through active memory erasure yields an 85% success rate.

These error rates are based on only one network, but are supported as approximate error rate figures by similar results from the

other network. Extensive simulation of the wave generating mechanism shows that most networks with a given set of parameters have similar wave generation properties. Since the sources of error are intimately linked with the wave pattern dynamics, most networks having proper control of the numbers of waves will show 60% to 90% success rates. Success rates can be increased through the use of mechanisms that erase old or conflicting memories to decrease the chances of failure due to lack of associations and to interference from stored sequences.

Success rates will also be changed if the forced association rule used in simulation is altered. Under this rule, two or more associations were forced to form between each input cell and the cells activated by waves at the time the input entered the system. However, each activated cell was allowed to associate with only one of the two input cells. As training continued and memory capacity was used up, the chance decreased that there were enough waves in the system to associate with the inputs according to this rule. The forced association rule could be modified to allow each activated cell to form connections to both input cells, and to make connections with some probability. This modified rule corresponds more closely to the learning rule in the original description of the model. Under the modified rule, the chance of association failure is decreased, since any activated cell that does not already have an association connection can contact either or both of the cells activated by the input. At the same time, however, failure may occur under this rule, since associations are formed with some probability. The balance between these two effects could be explored in further simulations.

Note that the modified association rule also extends the memory's effective capacity, since fewer cells are likely to associate with a given input.

As mentioned earlier, this model bears some resemblance to those of Marr (1970) and Grossberg (1969). In Marr's (1970) model, feature-detection cells called codons are used to signal the presence of particular patterns of activity in the overall input. The E-cells here are similar in that they fire in response to particular wave patterns in the network. The outputs of Marr's codons are combined at pattern classification cells through coupling weights. The magnitude of each weight reflects the probability of occurrence of the codon's feature as a part of the class of inputs the classifier cell is to signal. The E's in the current model simply trigger the G's to which they are coupled, rather than activating them to some level depending on the probability that an association should be made. A probabilistic structure such as Marr's would have made training--and simulating--my model more difficult, but might have done away with some of the instability that is this model's major drawback.

That instability would, however, have made Marr's inferential scheme difficult to apply here. Since small differences between wave patterns are rapidly turned into larger ones, the model does not preserve a property of pattern "closeness" for very long. The chance is quite small that an input that is similar to a stored input (for example, a familiar system input that is slightly perturbed by noise or by a cell failure) is able to generate a sequence of

patterns that are at all like those generated by the stored input. When the next input is due, an inferential mechanism would have little to go on to decide whether a stored association is to be recalled or not.

A way around this difficulty is to set particular sequences of wave patterns into the system by making the wave-jump cells themselves inferential. Then, if a pattern of waves is sufficiently like a stored pattern, the next set of waves generated can be more like those generated by the stored pattern itself. This scheme amounts to making error corrections as the wave jumps are produced throughout the interstimulus interval, rather than trying to correct for all the errors that occur between inputs at the time the next input is due. Since pattern closeness is preserved for short times in the model, this sort of running error correction by inference may succeed. Some variety is lost, however, since sequences of wave patterns that come sufficiently close to stored traces will be captured by the stored patterns and will become identical with them. The interesting questions of the meaning of "closeness" in such a system, and of system memory capacity could be explored in further simulations.

The coupled G- and E-cells in this model are also similar to Grossberg's "outstars" (1969). An outstar is a set of cells activated by a single command cell to yield outputs that are records of past inputs. As mentioned earlier, a string of outstars whose command cells are activated sequentially can record and replay arbitrary space-time sequences. In the model presented here, the E-cell acts as command

cell and the G-cell (or cells) to which it connects acts as an output cell. Inputs here do not have arbitrary levels, but take on only the values 'high' and 'low'. The main difference, then, between this model and Grossberg's is the recurrence of the pulses that activate the command cells, as mentioned earlier. This property, together with the pattern-recognition properties of the E-cells, allows a smaller structure to generate properly-timed command cell activations involving long interstimulus intervals.

The memory network discussed here can be used in two ways as part of a system that displays temporal conditioning. According to Groves and Thompson (1970), the state system can become "entrained" to a regularly repeated stimulus. This entrainment results in greater sensitization to a regular stimulus than to one that is not regular; a result that has been reported by Pendergrass and Kimmel (1968). A regular input to the network presented here allows a strong conditioned response to form in the granule cells excited by the stimulus. Their firing rates in response to the conditioned stimulus are bolstered by the action of the stored memory trace. Such enhanced firing levels could lead to increased levels of sensitization, as required by Groves and Thompson (1970). Thus this network can represent one realization of entrainment of the state system to the period of a regular input.

Vinogradova (1975) showed that cells of the dentate gyrus display reactions to stimuli that change form as the stimulation proceeds, but that do not habituate. The reactions of units in CA3, to

which the dentate sends its axons, do habituate, however. Following Sokolov's model theory of habituation, Vinogradova (1975) concluded that incoming stimuli are matched in CAe with stimulus images stored in the dentate gyrus. As the stimulus model is formed in the dentate, the responses in CA3 gradually drop out. Mismatch caused by a change in the stimulus causes CA3 units to dishabituate. Thus the memory network presented here can be used as part of a plausible match/mismatch habituation system in hippocampus that is sensitive to the temporal properties of a stimulus. The units in CA3 can operate according to the match/mismatch schemes of Horn (1967) to produce the desired habituation.

BIBLIOGRAPHY

Alger, B.E., and T.J. Teyler, "Long-term and Short-term Plasticity in the CA1, CA3, and Dentate Regions of the Rat Hippocampal Slice," Brain Res., in press, 1976.

Amari, S.- I., "Characteristics of Random Nets of Analog Neuron-like Elements," IEEE Trans. on Syst., Man, Cyb., SMC-2, 1972, pp. 643-657.

Amari, S.- I., "Learning Patterns and Pattern Sequences by Self-organizing Nets of Threshold Elements," IEEE Trans. Comp., C-21, 1972, pp. 1197-1206.

Andersen, J.A., "A Simple Neural Network Generating an Interactive Memory," Math. Biosci., 14, 1972, pp. 197-220.

Andersen, P., H. Bruland, and B. Kaada, "Activation of Dentate Area by Septal Stimulation," Acta Physiol. Scand., 51, 1961, pp. 17-28.

Andersen, P., J.C. Eccles, and Y. Løyning, "Location of Postsynaptic Inhibitory Synapses on Hippocampal Pyramids," J. Neurophysiol., 27, 1964, pp. 592-607.

Andersen, P., J.C. Eccles, and Y. Løyning, "Pathway of Postsynaptic Inhibition in the Hippocampus," J. Neurophysiol. 27, 1964, pp. 608-619.

Andersen, P., T.W. Blackstad, and T. Lømo, "Location and Identification of Excitatory Synapses on Hippocampal Pyramidal Cells," Exp. Brain Res., 1, 1966, pp. 236-248.

Andersen, P., B. Holmquist, and P.E. Voorhoeve, "Entorhinal Activation of Dentate Granule Cells," Acta Physiol. Scand., 66, 1966, pp. 448-460.

BIBLIOGRAPHY

Andersen, P., T.V.P. Bliss, and K.K. Skrede, "Lamellar Organization of Hippocampal Excitatory Pathways," Exp. Brain Res., 13, 1971, pp. 222-238.

Andersen, P., B.H. Bland, and J.D. Dudar, "Organization of the Hippocampal Output," Exp. Brain Res., 17, 1973, pp. 152-168.

Auerbach, A.A., "Transmitter Release at Chemical Synapses," in Structure and Function of Synapsis, eds. G.D. Pappas and D.P. Purpura, Raven Press, New York, 1972, pp. 137-160.

Beurle, R.L., "Properties of a Mass of Cells Capable of Regenerating Impulses," Phil. Trans. Roy. Soc. Lond. B., 249, 1956, pp. 55-81.

Bliss, T.V.P., and T. Lomo, "Long-lasting Potentiation of Synaptic Transmission in the Dentate Area of the Anaesthetized Rabbit Following Stimulation of the Perforant Path," J. Physiol. (Lond.), 232, 1973, pp. 331-356.

Brookhart, J.M., and E. Fadiga, "Potential Fields Initiated During Monosynaptic Activation of Frog Motoneurons," J. Physiol. (Lond.), 150, 1960, pp. 633-655.

Bruner, J., and D. Kennedy, "Habituation: Occurrence at a Neuromuscular Junction," Science, 169, 1970, pp. 92-94.

Buchwald, J.S., and G.L. Humphrey, "An Analysis of Habituation in the Specific Sensory System," in Progress in Physiological Psychology, 5, ed. E. Stellar and J.M. Sprague, Academic Press, New York, 1973, pp. 1-76.

Carew, T.J., H.M. Pinsker, E.R. Kandel, "Long-term Habituation of a Defensive Withdrawal Reflex in Aplysia," Science, 175, 1972, pp. 451-454.

BIBLIOGRAPHY

Castellucci, V., H. Pinsker, I. Kupfermann, and E. Kandel, "Neuronal Mechanisms of Habituation and Dishabituation of the Gill Withdrawal Reflex in Aplysia," Science, 167, 1970, pp. 1745-1748.

Chronister, R.B., and L.E. White, Jr., "Fiberarchitecture of the Hippocampal Formation: Anatomy, Projections, and Structural Significance," in The Hippocampus, Vol. 1, R.L. Isaacson and K.H. Pribram, eds., New York, Plenum Press, 1975, pp. 9-40.

Claviea, R.M., and A. Routtenberg, "Ascending Monamine-containing Fiber Pathways Related to Intracranial Self-stimulation: Histochemical Flourescence Study, Brain Res., 72, 1974, pp. 25-40.

Conrad, L.C., C.M. Leonard, D.W. Pfaff, "Connections of the Median and Dorsal Raphe Nuclei in the Rat: An Auroradiographic and Degeneration Study, J. Comp. Neurol., 156, 1974, pp. 179-206.

Cowan, J.D., Wilson, H.R., "A Mathematical Theory of the Functional Dynamics of Cortical and Thalamic Nervous Tissue," Kybernetik, 13, 55, 1973.

Covian, M.R., "Studies on the Neurovegetative and Behavioral Functions of the Brain Septal Area," in Structure and Function of the Limbic System, Prog. in Brain Res., Vol. 27, eds. W.R. Adey and T. Tokizane, 1967, pp. 189-217.

Culbertson, J.T., "Some Uneconomical Robots," in Automata Studies, eds. C.E. Shannon and J. McCarthy, Princeton University Press, Princeton, New Jersey, 1956, pp. 99-116.

De France, J., S.T. Kitai, and T. Shimono, "Electrophysiological Analysis of the Hippocampal-septal Projections, I. Response and Topographical Characteristics," Exp. Brain Res., 17, 1973, pp. 447-462.

BIBLIOGRAPHY

Dichter, M., and W.A. Spencer, "Penicillin-induced Interictal Discharges from Cat Hippocampus, II. Mechanisms Underlying Origin and Restriction," J. Neurophys., 32, 1969, pp. 663-687.

Douglas, R.J. "The Hippocampus and Behavior," Psych. Bulletin, 67, 1967, pp. 416-442.

Douglas, R.M., and G.V. Goddard, "Long-term Potentiation of the Perforant Path-Granule Cell Synapse in the Rat Hippocampus," Brain Res., 86, 1975, pp. 205-215.

Eisenstein, E.M., and B. Peretz, "Comparative Aspects of Habituation in Invertebrates," in Habituation, Vol. II, eds. H.V.S. Peeke and M.J. Herz, Academic Press, New York, 1973, pp. 1-34.

Elazar, Z., and W.R. Adey, "Spectral Analysis of Low Frequency Components in the Electrical Activity of the Hippocampus During Learning," E.E.G. and Clin. Neurophys., 23, 1967, 225-240.

Ellias, S.A., and S. Grossberg, "Pattern Formation, Contrast Control, and Oscillations in the Short-term Memory of Shunting On-center Off-surround Networks, Biol. Cyb., 20, 1975, pp. 69-98.

Epstein, R., and L. Tauc, "Heterosynaptic Facilitation and Post-tetanic Potentiation in Aplysia Nervous System, J.Physiol., 209, 1970, pp. 1-23.

Farel, P.B., "Long-lasting Habituation in Spinal Frogs, Brain Res., 33, 1971, pp. 405-417.

Farel, P.B., and R.F. Thompson, "Habituation and Dishabituation to Dorsal Root Stimulation in the Isolated Frog Spinal Cord," Behavioral Biol., 7, 1972, pp. 37-45.

BIBLIOGRAPHY

Farel, P.B., D.L. Glanzman, and R.F. Thompson, "Habituation of a Monosynaptic Response in Vertebrate Central Nervous System: Lateral Column-Motoneuron Pathway in Isolated Frog Spinal Cord," J. Neurophysiol., 36, 1973, pp. 1117-1130.

Feldman, J.L., and J. Cowan, "Large-Scale Activity in Neural Nets II: A Model for the Brainstem Respiratory Oscillator," Biol. Cyb., 17, 1975, pp. 39-51.

Freeman, W.J., Mass Action in the Nervous System, Academic Press, New York, 1975.

Fukushima, K, "A Model of Associative Memory in the Brain," Kybernetik, 12, 1973, pp. 58-63.

Gardner, L.E., "Retention and Over-habituation of a Dual-component Response in Lumbricus Terrestris, J. Comp. Physiol. Psychol., 66, 1968, pp. 315-318.

Goodman, D.A., and N.M. Weinberger, "Habituation in "Lower" Tetrapod Vertebrates: Amphibia as Vertebrate Model Systems," in Habituation, Vol. I, ed. H.V.S. Peeke and M.J. Herz, Academic Press, New York, 1973, pp. 86-140.

Gottlieb, D.I., and W.M. Cowan, "Antoradiographic Studies of the Commissural and Ipsilateral Association Connections of the Hippocampus and Dentate Gyrus of the Rat, I. The Commissural Connectives," J. Comp. Neurol., 149, 1973, pp. 393-422.

Graham, F.K., "Habituation and Dishabituation of Responses Innervated by the Autonomic Nervous System," in Habituation, Vol. I, ed. H.V.S. Peeke and M.J. Herz, Academic Press, New York, 1973, pp. 63-81.

BIBLIOGRAPHY

Grastyán, E., K. Lissák, I. Madarász, and H. Donhoffer, "Hippocampal Electrical Activity During the Development of Conditioned Reflexes," E.E.G. and Clin. Neurophys., 11, 1959, pp. 409-429.

Griffin, J.P., "Neurophysiological Studies into Habituation," in Short-term Changes in Neural Activity and Behavior, eds. G. Horn and R.A. Hinde, Cambridge Univ. Press, Cambridge, 1970, pp. 141-180.

Grossberg, S., "Some Networks That Can Learn, Remember, and Reproduce Any Number of Complicated Space-time Patterns, I," Journal of Math. Mech., 19, 1969, pp. 53-91.

Grossberg, S., "Classical and Instrumental Learning by Neurological Networks," In Progress in Theoretical Biology, 3, Academic Press, New York, 1974

Grossberg, S., "Contour Enhancement, Short-term Memory, and Constancies in Reverberating Neural Networks," Studies in Appl. Math., 52, 1973, pp. 213-257.

Groves, P.M., D. Lee, and R.F. Thompson, "Effects of Stimulus Frequency and Intensity on Habituation and Sensitization Acute Spinal Cat," Physiol. Beh., 4. 1969, pp. 383-388.

Groves, P.M., and R.F. Thompson, "Habituation: A Dual-process Theory," Psych. Rev., 77, 1970, pp. 419-450.

Groves, P.M., and R.F. Thompson, "A Dual-process Theory of Habituation: Neural Mechanisms," in Habituation, Vol. 2, eds. H.V.S. Peeke and M.J. Herz, Academic Press, New York, 1973, pp. 175-205.

Hebb, D.O., The Organization of Behavior, Wiley, New York, 1949.

BIBLIOGRAPHY

Hernández-Péon, R., "Neurophysiological Correlates of Habituation and Other Manifestations of Plastic Inhibition (Internal Inhibition)," in The Moscow Colloquium on Electroencephalography of Higher Nervous Activity, E.E.G. and Clin. Neurophys., 1960, Suppl. 13.

Hinde, R.A., "Factors Governing the Changes in Strength of a Partially Inborn Response, as Shown by the Mobbing Behavior of the Chaffinch (Fringella Coelebs), III. The Interaction of Short-term and Long-term Incremental and Decremental Effects," Proc. Roy. Soc. Ser. B, 153, 1954, pp. 398-420.

Hinde, P.A., "Behavioral Habituation," in Short-term Changes in Neural Activity and Behavior, eds. G. Horn and R.A. Hinde, Cambridge University Press, 1970, pp. 3-40.

Hirano, T., P. Best, and J. Olds, "Units During Habituation, Discrimination Learning, and Extinction," E.E.G. and Clin. Neurophys., 28, 1970, pp. 127-135.

Hjorth-Simonsen, A., "Hippocampal Efferents to the Ipsilateral Entorhinal Area: An Experimental Study in the Rat," J. Comp. Neurol., 142, p. 417.

Hjorth-Simonsen, A., and B. Jeune, "Origin and Termination of the Hippocampal Perforant Path in the Rat Studied by Silver Impregnation," J. Comp. Neurol., 144, 1972, pp. 215-232.

Hjorth-Simonsen, A., "Projection of the Lateral Part of the Entorhinal Area to the Hippocampus and Fascia Dentata," J. Comp. Neurol., 146, 1972, pp. 219-232.

Holmgren, B., and S. Frenk, "Inhibitory Phenomena and Habituation at the Neuronal Level," Nature, 192, 1961, pp. 1294-1295.

BIBLIOGRAPHY

Horowitz, J., W. Freeman, and P. James Stoll, "A Neural Network With a Background Level of Excitation in the Cat Hippocampus," Intern. J. Neurosci., 5, 1973, pp. 113-123.

Horn, G., "Neuronal Mechanisms of Habituation, Nature, 215, 1967, pp. 707-711.

Horn, G., "Changes in Neuronal Activity and Their Relationships to Behavior," in Short-term Changes in Neural Activity and Behavior, eds. G. Horn and R.A. Hinde, Cambridge University Press, Cambridge, 1970, pp. 567-606.

Horn, G., and M.J. Wright, "Characteristics of Transmission Failure in the Squid Stellate Ganglion: A Study of a Simple Habituating System," J. Exp. Biol., 51, 1970, pp. 217-231.

Horn, G., "Novelty, Attention and Habituation," in Attention in Neurophysiology, eds. C.R. Evans and T.B. Mulholland, Butterworths, London, 1974, pp. 230-246.

Isaacson, R.L., The Limbic System, Plenum Press, New York, 1974.

Isaacson, R.L., and K.H. Pribram, The Hippocampus, Vol. 1: Structure and Development, and Vol. 2: Neurophysiology and Behavior, Plenum Press, New York, 1975.

Izquierdo, I., "The Hippocampus and Learning," Progress in Neurobiology 5, Part 1, 1975, pp. 37-75.

Jacklet, J.W., and K. Lukowiak, "Neural Processes in Habituation and Sensitization in Model Systems," in Progress in Neurobiology, 4, ed. G.A. Kerkut and J.W. Phillis, Pergamon Press, 1975.

BIBLIOGRAPHY

Kandel, E.R., and L. Tauc, "Heterosynaptic Facilitation in Neurones of the Abdominal Ganglion of Aplysia Depilans," J. Physiol. (Lond.), 181, 1965, pp. 1-27.

Kandel, E.R., and L. Tauc, "Mechanism of Heterosynaptic Facilitation in the Giant Neuron of Abdominal Ganglion of Aplysia Depilans," J. Physiol. (Lond.), 181, 1965, pp. 28-47.

Kandel, E., V. Castellucci, H. Pinsker, and I. Kupfermann, "The Role of Synaptic Plasticity in the Short-term Modification of Behavior," in Short-term Changes in Neural Activity and Behavior, ed. G. Horn and R.A. Hinde, Cambridge University Press, Cambridge, 1970, pp. 281-322.

Kilmer, W.L., and T. McLardy, "A Diffusely Preprogrammed But Sharply Trainable Hippocampus Model," Int. J. Neuroscience, 2, 1971, pp. 211-248.

Kilmer, W., M. Olinski, "Model of a Plausible Learning Scheme for CA3-Hippocampus," Kybernetik, 16, 1974, pp. 133-144.

Kilmer, W.L., "Biology of Decisionary and Learning Mechanisms in Mammalian CA3-hippocampus: A Review," Int. J. Man-Mach. Studies, 7, 1975, pp. 413-437.

Kimmel, H.D., "Habituation, Habituability, and Conditioning," in Habituation, Vol. 1, eds. H.V.S. Peeke and M.J. Herz, Academic Press, New York, 1973, pp. 219-238.

Kohonen, T., "An Adaptive Associative Memory Principle," IEEE Trans. Comp., C-23, 1974, pp. 444-445.

BIBLIOGRAPHY

Krasne, F.B., and A.M. Roberts, "Habituation of the Crayfish Escape Response During Release From Inhibition Induced by Picrotoxin," Nature, 215, 1969, pp. 769-770.

Krasne, R.B., "Learning in Crustacea," in Invertebrate Learning, Vol. 2: Arthropods and Gastropod Mollusks, eds. W.C. Corning, J.A. Dyal, and A.O.D. Willows, Plenum Press, New York, 1973, pp. 49-130.

Kupfermann, I., V. Castellucci, H. Pinsker, and E. Kandel, "Neural Correlates of Habituation and Dishabituation of the Gill Withdrawal Reflex in Aplysia," Science, 167, 1970, pp. 1743-1745.

Lebovitz, R.M., M. Dichter, W.A. Spencer, "Recurrent Excitation in the CA3 Region of Cat Hippocampus," Intern. J. Neurosci., 2, 1971, pp. 99-108.

Leibovic, K.N., Nervous System Theory, Academic Press, New York, 1972.

MacGregor, R.J., and R.M. Oliver, "A Model for Repetitive Firing in Neurons," Kybernetik, 16, 1974, pp. 53-64.

Machne, X., E. Fadiga, and J.M Brookhart, "Antidromic and Synaptic Activation of Frog Motor Neurons," J. Neurophysiol., 22, 1959, pp. 583-603.

Marr, D., "A Theory for Cerebral Neocortex," Proc. Roy. Soc. Lond., B., 176, 1970, pp. 161-234.

Marr, D., "Simple Memory: A Theory for Archicortex," Proc. Roy. Soc. Lond., B., 262, 1971, pp. 23-81.

Martin, A.R., "A Further Study of the Statistical Composition of the End-plate Potential," J. Physiol., 130, 1955, p. 114.

BIBLIOGRAPHY

McLardy, T., and W.L. Kilmer, "Hippocampal Circuitry," Amer. Psychol., 25, 1970, pp. 563-566.

McLardy, T., "Habituation Deficit and Paucity of Dentate Granule-cells in Some Schizophrenic Brains," IRCS Int. Res. Comm. System 73-3, 16-1-1, 1973.

Mittenthal, Jay E., "Reliability of Pattern Separation by the Cerebellar Mossy Fiber-Granule Cell System," Kybernetik, 16, 1974, pp. 93-101.

Moore, R.Y., and A.E. Halaris, "Hippocampal Inervation by Serotonin Neurons of the Midbrain Raphe in the Rat," J. Comp. Neurol., 164, 1975, pp. 171-184.

Mosko, S., G. Lunch, and C.W. Cotman, "The Distribution of Septal Projections to the Hippocampus of the Rat," J. Comp. Neurol., 152, 1973, pp. 163-174.

Nilsson, N.J., Learning Machines, McGraw-Hill, New York, 1965.

O'Keefe, J., and J. Dostrovsky, "The Hippocampus as a Spatial Map: Preliminary Evidence from Unit Activity in the Freely Moving Rat," Brain Res., 34, 1971, pp. 171-175.

Olds, J., W.D. Mink, and P.J. Best, "Single Unit Patterns During Anticipatory Behavior," E.E.G. and Clin. Neurophys., 26, 1969, pp. 144-158.

Olds, J., "The Central Nervous System and the Reinforcement of Behavior," Am. Psychol., 24, 1969, pp. 114-132.

BIBLIOGRAPHY

Pakula, A., and E.N. Sokolov, "Habituation in Gastropods: Behavioral, Interneuronal, and Endoneuronal Aspects," in Habituation, Vol. II, ed. H.V.S. Peeke and M.J. Herz, Academic Press, New York, 1973, pp. 35-108.

Pearson, J.A., and J.F. MacDonald, "Habituation of the Flexor Reflex: Inhibitory Build-up or Synaptic Depression?", Brain Res., 53, 1973, pp. 451-454.

Petrinovich, L., "A Species-meaningful Analysis of Habituation," in Habituation, Vol. 1, eds. H.S.V. Peeke and M.J. Herz, Academic Press, 1973, pp. 141-162.

Pinsker, H., I. Kupfermann, V. Castellucci, and E. Kandel, "Habituation and Dishabituation of the Gill Withdrawal Reflex in Aplysia," Science, 167, 1970, pp. 1740-1742.

Prosser, C.L., and W.S. Hunter, "The Extinction of Startle Responses and Spinal Reflexes in the White Rat," Am. J. Physiol., 117, 1936, pp. 609-618.

Raisman, G., W.M. Cowan, and T.P.S. Powell, "An Experimental Analysis of the Efferent Projection of the Hippocampus," Brain, 89, 1966, pp. 83-108.

Rall, W., "A Statistical Theory of Monosynaptic Input-output Relations," J. Cell and Comp. Physiol., 46, 1955, pp. 373-411.

Rall, W., "Experimental Monosynaptic Input-output Relations in the Mammalian Spinal Cord," J. Cell and Comp. Physiol., 46, 1955, pp. 413-437.

BIBLIOGRAPHY

Ramon y Cajal, S., The Structure of Ammon's Horn, translated by L.M. Kraft, Charles E. Thomas, Publisher, Springfield, 1968.

Ranck, J.B., Jr., "Studies on Single Neurons in Dorsal Hippocampal Formation and Septum in Unrestrained Rats, Part I. Behavioral Correlates and Firing Repertoires," Exp. Neurol., 41, 1973, pp. 461-531.

Ratliff, F., Mach Bands, Holden-Day, San Francisco, 1965.

Rowell, C.H.F., "Incremental and Decremental Processes in the Insect Central Nervous System," in Short-term Changes in Neural Activity and Behavior, eds. G. Horn, R.A. Hinde, Cambridge University Press, 1970, pp. 237-280.

Segal, M., "Hippocampal Unit Responses to Perforant Path Stimulation," Exp. Neurol., 35, 1972, p. 541.

Segal, M., "The Hippocampus as a Learning Machine," Ph.D. Thesis, California Institute of Technology, Pasadena, California, 1973.

Segal, M., "Physiological and Pharmacological Evidence for a Serotonergic Projection to the Hippocampus," Brain Res., 94, 1975, pp. 115-131.

Segundo, J.P., D.H. Perkel, H. Wyman, H. Hegstad, G.P. Moore, "Input-output Relations in Computer-simulated Nerve Cells," Kybernetik, 12, 1968, pp. 157-171.

Segundo, J.P., and C.C. Bell, "Habituation of Single Nerve Cells in the Vertebrate Nervous System," in Short-term Changes in Neural Activity and Behavior, eds. G. Horn and R.A. Hinde, Cambridge University Press, Cambridge, 1970, pp. 77-94.

BIBLIOGRAPHY

Sharpless, S., and H. Jasper, "Habituation of the Arousal Reaction," Brain, 79, 1956, pp. 655-680.

Shepherd, G., The Synaptic Organization of the Brain, Oxford University Press, New York, 1974.

Sherrington, C.S., The Integrative Action of the Nervous System, Yale University Press, New Haven, 1906.

Siegel, A., and J.P. Tassoni, "Differential Efferent Projections from the Ventral and Dorsal Hippocampus of the Cat," Brain Beh. Evol., 4, 1971, pp. 185-200.

Sokolov, E.N., "Neuronal Models and the Orienting Reflex," in The Central Nervous System and Behavior, ed. M.A. Brazier, 1961, pp. 187-276.

Sokolov, E.N., "The Neuronal Mechanisms of the Orienting Reflex," in Neuronal Mechanisms of the Orienting Reflex, eds. E.N. Sokolov and O.S. Vinogradova, Wiley, New York, 1975, pp. 217-238.

Spencer, W.A., R.F. Thompson, and D.R. Neilson, Jr., "Response Decrement of the Flexion Reflex in the Acute Spinal Cat and Transient Restoration by Strong Stimuli," J. Neurophysiol., 29, 1966, pp. 221-252.

Spencer, W.A., R.F. Thompson, and D.R. Neilson, "Decrement of Ventral Root Electrotonus and Intercellularly Recorded PSPs Produced by Iterated Cutaneous Afferent Volleys," J. Neurophysiol., 29, 1961, pp. 253-274.

Spinelli, D.N., "OCCAM: A Computer Model for a Content Addressable Memory in the Central Nervous System," in The Biology of Memory, eds. K. Pribram and D. Broadbent, Academic Press, New York, 1970, pp. 293-306.

BIBLIOGRAPHY

Stein, R.B., K.V. Leung, D. Mangeron, M.N. Oğuztöreli, "Improved Neuronal Models for Studying Neural Networks," Kybernetik, 15, 1974, pp. 1-9.

Szlep, R., "Change in the Response of Spiders to Repeated Web Vibrations," Behavior, 23, 1964, pp. 203-238.

Tauc, L., "Presynaptic Inhibition in the Abdominal Ganglion of Aplysia," J. Physiol., 181, 1965, pp. 282-307.

Teyler, T.J., and B.E. Alger, "Monosynaptic Habituation in the Vertebrate Forebrain: The Dentate Gyrus Examined in vitro," Brain Res., in press, 1976.

Thompson, R.F., "The Neural Basis of Stimulus Generalization," in Stimulus Generalization, ed. D.I. Mostofsky, Stanford University Press, Stanford, California, 1965, pp. 154-178.

Thompson, R.F., and W.A. Spencer, "Habituation: A Model Phenomenon for the Study of Neuronal Substrates of Behavior," Psych. Review, 73, 1, 1966, pp. 16-43.

Uttley, A.M., "The Informon in Classical Conditioning," J. Theor. Biol., 49, 1975, pp. 355-376.

Thompson, R.F., P.M. Groves, T.J. Teyler, and R.A. Roemer, "A Dual-process Theory of Habituation: Theory and Behavior," in Habituation, Vol. 1, eds. H.V.S. Peeke and M.J. Herz, Academic Press, New York, 1973, pp. 239-272.

Vanderwolf, C.H., R. Kramis, L.A. Gullespie, and B.H. Bland, "Hippocampal Rhythmic Slow Activity and Neocortical Low-voltage Fast Activity: Relations to Behavior," in The Hippocampus, Vol. 2, eds. R.L. Isaacson and K.H. Pribram, Plenum Press, New York, 1975, pp. 101-128.

BIBLIOGRAPHY

Van Hoesen, G.W., D.N. Pandya, and N. Butters, "Cortical Afferents to the Entorhinal Cortex of the Rhesus Monkey," Science, 175, 1972, pp. 1471-1473.

Vinogradova, O., "Registration of Information and the Limbic System," in Short-term Changes in Neural Activity and Behavior, eds. G. Horn and R.A. Hinde, Cambridge University Press, Cambridge, 1970, pp. 95-140.

Vinogradova, O., T.P. Semoyonova, V.Ph. Konovalov, "Trace Phenomenon in Single Neurons of the Hippocampus and Mammillary Bodies," in The Biology of Memory, eds. K. Pribram and D. Broadbent, Academic Press, New York, 1970, pp. 191-222.

Vinogradova, O., "Functional Organization of the Limbic System in the Process of Registration of Information: Facts and Hypothesis," in The Hippocampus, Vol. 2, eds. R.L. Isaacson and K.H. Pribram, Plenum Press, New York, 1975, pp. 3-70.

Vinogradova, O., "The Hippocampus and the Orienting Response," in Neuronal Mechanisms of the Orienting Reflex, eds. E.N. Sokolov and O.S. Vinogradova, Wiley, New York, 1975, pp. 128-154.

Wall, P.D., "Habituation and Post-tetanic Potentiation in the Spinal Cord," in Short-term Changes in Neural Activity and Behavior, eds. G. Horn and R.A. Hinde, Cambridge University Press, Cambridge, 1970, pp. 181-210.

Wickelgren, B.G., "Habituation of Spinal Motoneurons," J. Neurophysiol., 30, 1967, pp. 1404-1423.

Wickelgren, B.G., "Habituation of Spinal Interneurons," J. Neurophysiol., 30, 1967, pp. 1424-1438.

BIBLIOGRAPHY

Wigström, H., "A Model of a Neural Network With Recurrent Inhibition," Kybernetik 16, 1974, pp. 103-112.

Wigström, H., "Associative Recall and Formation of Stable Modes of Activity in Neural Network Models," preprint, 1975.

Wine, J.J., and F.B. Krasne, "Independence of Inhibition and Habituation in the Crayfish Lateral Giant Fiber Escape Reflex," Proc. Seventy-seventh Ann. Convention A.P.A., 1969, pp. 237-238.

Wine, J.J., and F.B. Krasne, "The Organization of Escape Behavior in the Crayfish," J. Exptl. Biol., 56, 1972, pp. 1-18.

Wyers, E.J., H.V.S. Peeke, and M.J. Herz, "Behavioral Habituation in Invertebrates," in Habituation, Vol. 1., ed. H.V.S. Peeke and M.J. Herz, Academic Press, 1973, pp. 1-57.

Zimmer, J., "Ipsilateral Afferents to the Commissural Zone of the Fascia Dentata, Demonstrated in Decommissurated Rats by Silver Impregnation," J. Comp. Neurol., 142, 393, 1973.

Zucker, R.S., "Crayfish Escape Behavior and Central Synapses, I. Neural Circuit Exciting Lateral Giant Fiber, II. Physiological Mechanisms Underlying Habituation," J. Neurophysiol., 35, pp. 599-651.

APPENDIX A

THE HABIT EQUATIONS

In this appendix I derive the equation describing the way the output of the circuit of Figure 2.6 changes with repeated stimulation. This circuit is illustrated again in Figure A.1 for convenience. The assumptions under which the circuit operates are:

1. Cell dynamics do not enter into the circuit's operation. Hence, the circuit equations are:

$$D = I$$
$$S = W_1 D$$
$$O = W_2 D + W_3 S$$

2. Synapses vary according to first-order linear differential equations:

$$\tau_1 \dot{W}_1 = W_{10} - W_1 - a_1 D \qquad W_1(0) = W_{10}$$
$$\tau_2 \dot{W}_2 = W_{20} - W_2 - a_2 D \qquad W_2(0) = W_{20}$$
$$\tau_3 \dot{W}_3 = W_{30} - W_3 - a_3 S \qquad W_3(0) = W_{30}$$

3. The input is a constant.

Under these assumptions the differential equations for the synapses can easily be solved. The equations for W_1 and W_2 may be integrated by inspection to give

$$W_1(t) = W_{10} + (W_1(0) - W_{10})\, e^{-t/\tau_1} - a_1 D(1 - e^{-t/\tau_1})$$
$$W_2(t) = W_{20} + (W_2(0) - W_{20})\, e^{-t/\tau_2} - a_2 D(1 - e^{-t/\tau_2})$$

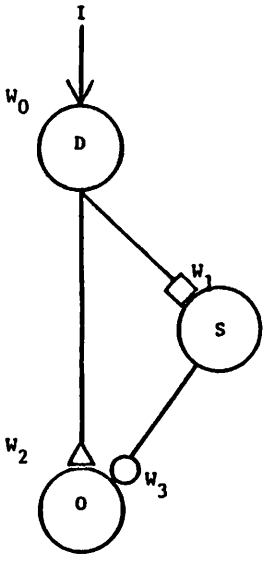


Figure A.1 Lumped model presented in terms of cells and coupling weights. Synapse modification properties are as indicated on the figure.

The form for W_3 is

$$W_3(t) = W_{30} + (W_3(0) - W_{30})\, e^{-t/\tau_3} + a_3\, D\, e^{-t/\tau_3} \int_0^t S(\tau) e^{\tau/\tau_3}\, d\tau$$

From the equations for S and W_1, this integral can be evaluated to give

$$W_3(t) = W_{30} + (W_3(0) - W_{30})\, e^{-t/\tau_3} + a_3\, D(W_{10} - a_1\, D)[1 - e^{-t/\tau_3}] + a_3\, D\, \frac{\tau_1}{\tau_1 - \tau_2}\, (W_1(0) - W_{10} + a_1\, D)[e^{-t/\tau_1} - e^{-t/\tau_3}]$$

Using the facts that D is identical with I, and that all synapses are initially set to their resting levels gives finally

$$W_1(t) = W_{10} - a_1\, I(1 - e^{-t/\tau_1})$$
$$W_2(t) = W_{20} - a_2\, I(1 - e^{-t/\tau_2})$$
$$W_3(t) = W_{30} - a_3\, I(W_{10} - a_1\, I)[1 - e^{-t/\tau_3}] + a_1\, a_3\, I^2\, \frac{\tau_1}{\tau_1 - \tau_3} [e^{-t/\tau_1} - e^{-t/\tau_3}]$$

In the case considered in Chapter III, the decreasing weights varied independent of stimulus intensity. The equations above then reduce to

$$W_1(t) = W_{10} - a_1(1 - e^{-t/\tau_1})$$
$$W_2(t) = W_{20} - a_2(1 - e^{-t/\tau_2})$$
$$W_3(t) = W_{30} + a_3\, I(W_{10} - a_1)[1 - e^{-t/\tau_3}] + a_1\, a_3\, I\, \frac{\tau_1}{\tau_1 - \tau_3} [e^{-t/\tau_1} - e^{-t/\tau_3}]$$

These were the equations employed in the HABIT program.

One interesting feature of these synaptic equations is the way the peak value of the resulting equation for the overall output varies

as a function of stimulus intensity. I consider here the case for which W_1 is fixed. Then a_1 is zero and the equation for W_3 simplifies to

$$W_3(t) = W_{30} + a_3\, W_{10}\, I(1 - e^{-t/\tau_3})$$

The full equation for the output then is

$$\begin{aligned} O(t) &= W_2(t)\, D + W_3(t)\, S \\ &= [W_{20} - a_2(1 - e^{-\tau/\tau_2})]\, I + [W_{30} + a_3\, W_{10}\, I(1 - e^{-\tau/\tau_3})] \cdot W_{10}\, I \\ &= [W_{20} + W_{10}\, W_{30}]\, I + a_3\, W_{10}^2\, I^2(1 - e^{-t/\tau_3}) - a_2(1 - e^{-t/\tau_1})\, I \end{aligned}$$

The initial, or control, value for the output is expressed by the first term in the equation above. The output relative to the control level is

$$O_R(t) = 1 + \frac{a_3\, W_{10}^2\, I}{W_{20} + W_{10}\, W_{30}}\, (1 - e^{-t/\tau_3}) - \frac{a_3}{W_{20} + W_{10}\, W_{30}}\, (1 - e^{-t/\tau_1})$$

This function has a maximum (or a minimum) depending on the last two terms. These terms are of the form

$$F = A(1 - e^{-t/\tau_1}) - B(1 - e^{-t/\tau_2})$$

which has a maximum depending on the terms

$$B\, e^{-t/\tau_2} - A\, e^{-t/\tau_1}$$

This form has a singular point (found by taking the derivative and solving for the time when the derivative is zero) at

$$t^* = \frac{\tau_1 \tau_2}{\tau_2 - \tau_1} \ln\left(\frac{\tau_2}{\tau_1} \frac{A}{B}\right)$$

At that point the second derivative is

$$G = \frac{B}{\tau_2^2}\left(\frac{\tau_2}{\tau_1}\frac{A}{B}\right)^{-\frac{\tau_1}{\tau_2-\tau_1}} - \frac{A}{\tau_1^2}\left(\frac{\tau_2}{\tau_1}\frac{A}{B}\right)^{-\frac{\tau_2}{\tau_2-\tau_1}}$$

In the equation for the circuit output,

$$A = \frac{a_3 W_{10}^2}{W_{20} + W_{10} W_{30}} I \ , \quad B = \frac{a_2}{W_{20} + W_{10} W_{30}} \ ,$$

and τ_1 corresponds to τ_3. Recast in these terms, the equations above are

$$t^* = \frac{\tau_2\tau_3}{\tau_2 - \tau_3} \ln\left(\frac{\tau_2}{\tau_3}\frac{a_3 W_{10}^2}{a_2} I\right),$$

and

$$G = \frac{1}{W_{20} + W_{10} W_{30}}\left[\frac{a_2}{\tau_2^2}\left(\frac{\tau_2}{\tau_3}\frac{a_3 W_{10}^2}{a_2} I\right)^{-\frac{\tau_3}{\tau_2-\tau_3}} - \frac{a_3 W_{10}^2}{\tau_3^2} \cdot \left(\frac{\tau_2}{\tau_3}\frac{a_3 W_{10}^2}{a_2} I\right)^{-\frac{\tau_2}{\tau_2-\tau_3}}\right]$$

For a maximum of relative output, G must be negative. After some algebra, this condition is fulfilled if

$$\tau_3 < \tau_2$$

This was the condition under which the simulations of Chapter III were run. In that case, the value of t^* is positive and represents the point of maximum relative output. Note that this point shifts to later times with the logarithm of stimulus intensity. Thus some

peak shifting occurs in this simple model without the benefit of a second sensitization term with a different time constant.

These results were derived under the condition of constant input. The output due only to the decreasing weight can also be derived for the case of a pulsed input. This corresponds more closely to the way physiological systems are tested. Consideration of only the decremental term corresponds to testing the system at a low stimulus intensity. In order to get an idea of the way a weight governed by a first-order differential equation changes with time in this case, consider the equation

$$\tau\dot{y} = -y + a I \qquad y(0) = 0$$

This is a simplification of the synaptic equations dealt with previously. Information gained from studying this case easily transfers to those cases, however.

The analysis of this equation is based on the behavior illustrated in Figure A.2. There, a pulsed input of height A, period $T_1 + T_2$, and duration T_1 excites the equation started from zero. The value of y at the beginning and the end of each pulse may be expressed as

$$y_i = y_{i-1} e^{-T_2/\tau} \qquad i = 2, 4, 6, \ldots$$
$$y_i = y_{i-1} e^{-T_1/\tau} + a A(1 - e^{-T_1/\tau}) \qquad i = 1, 3, 4, \ldots$$

and $y_0 = y(0) = 0$

uation). The decrease is usually a negative exponential function of the number of stimulus presentations.

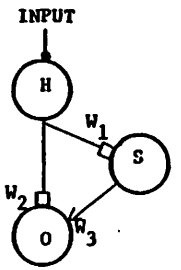
2. If the stimulus is withheld, the response tends to recover over time.

3. The weaker the stimulus, the more rapid and/or pronounced is the habituation, measured as a percent of control (unhabituated) response.

4. Presentation of another (usually strong) stimulus results in recovery of the habituated response (dishabituation).

5. Habituation and dishabituation of the response to a given stimulus exhibit generalization to other stimuli.

On the basis of work done in spinal preparations, Thompson devised a theory of habituation based on the interaction of two processes. In the two-process theory, the waning of response to a particular stimulus is produced by a transmission decrement in the path from stimulus to response. The recovery of response following presentation of another stimulus is the result of a superimposed, transitory sensitization produced by the new stimulus. The simple circuit below illustrates the two-process theory.



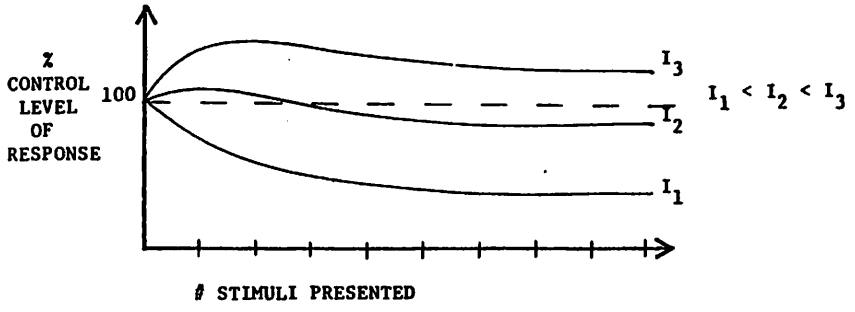
Cells are indicated here by circles. Each cell's output is a function of the sum of its inputs, each weighted by the value of the associated synapse. So:

$$H = F_H(I)$$
$$S = F_S(W_1 \cdot H)$$
$$O = F_O(W_2 \cdot H + W_3 \cdot S)$$

Synapses themselves vary with use to model the processes of decrement and increment. Synapses marked as squares have values that decrease as they are used, so that with successive presentations of an input, the transmission in the direct path from H to O decreases. The synapse marked with an arrow increases with use, simulating the process of sensitization. The values of all synapses return to resting levels if unstimulated, simulating the spontaneous recovery of habituation and sensitization.

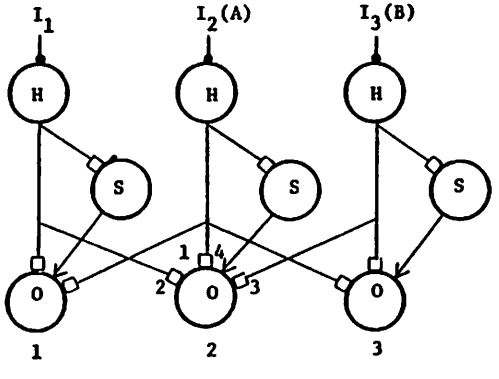
To realize Thompson's theory, synaptic decrement is made independent of stimulus intensity for sufficiently intense stimuli, while increment depends directly on intensity. Hence, small inputs will yield only habituation through the H-O decrement. Larger inputs will

generate both decrement in the H-O path and increment in the H-S-O path, leading to the generation of habituation curves illustrated below, redrawn from the work of Thompson and his colleagues.



With the proper choices of synapse parameters, similar curves may be generated by the UATION program.

In order to simulate the generalization of habituation, the simple unit illustrated above may be used as the basis of a network of units illustrated below.



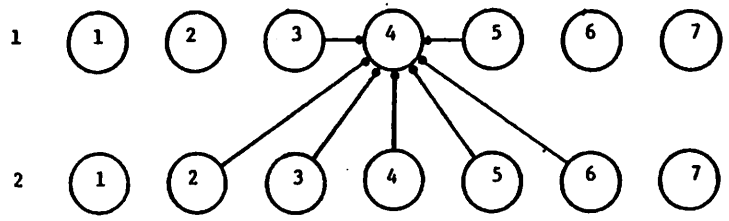
In the basic unit, the synapse values were altered by activity in the channels with which they were associated. Here, the value of a synapse is changed by activity in its own channel and also by activity in nearby channels of the same type. Thus synapse 1 above may be decreased by activity coming through synapses 2 and 3, but will be unaffected by activity at synapse 4. Input A, then, will decrease synapse 1 through activity in synapse 2, causing a decreased response to later presentation of input B. The sensitizing synapses behave in a similar way to simulate generalization of dishabituation, through these connections are not shown here.

Thompson's two-process theory of habituation, then, may be simulated by the networks shown above using the UATION program. With proper choice of network parameters habituation curves such as the ones shown on page 270 may be generated. The program is designed to enable the user to generate these curves and to perform simple habituation experiments.

Part II: Program use and operation. The network simulated by UATION is composed of a number of layers of cells. Each cell may be connected with any cell in any other layer (including its own) but all cells in a given layer have identical connections. The connections from one layer to another are specified in terms of a template. The template gives the value of the connection to a given cell in one layer from its neighbors in another. We are interested in homogenous networks, so the templates are symmetrical left-to-right. The input to a cell in a given layer is simply the weighted sum of the outputs of the cells to which it is connected, according to the templates for that cell's layer. The output of each cell in the network is a simple function of its input. The template is illustrated be-

low. The layer feeding into a connection is called the "output layer" for that connection, while the layer receiving the connection is called the "input layer."

CELL NUMBERING

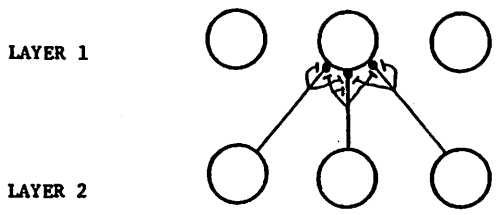


Connections between cells may be fixed or variable. Fixed connections are entirely accounted for by the template mechanism described above, since fixed connections are identical for all cells in a given layer and for all time. Variable connections between cells in two layers are specified identically initially, but must be kept track of individually thereafter. Templates are again used to set up variable connections. As above, such connections between two layers are set up symmetrically. That is, the connections from a cell in a given layer from its leftward neighbors in another layer initially have the same values as the corresponding connections from its rightward neighbors in that layer. These pairs of connections have identical modification parameters, as described below.

Each variable connection obeys a first-order differential equation with a resting value equal to the value in the template that specifies the connection. So each variable connection satisfies:

$$\tau \dot{W} = W_0 - W + \sum_i \alpha_i F_w (C_i)$$

Here τ is the time constant associated with the pair of connections, and W_0 is the resting level. The driving term is a weighted sum of a function of the outputs of the cells whose activities modify the connection value. Each pair of variable connections has an associated template telling which cells in the output layer may modify these connections. A cell in the output layer of a connection may modify that connection's value only if it makes contact with that connection's cell in the input layer, however. This is illustrated below for a system with two elements in the template from layer 2 to layer 1, and with two elements in the template of each pair of variable connections.



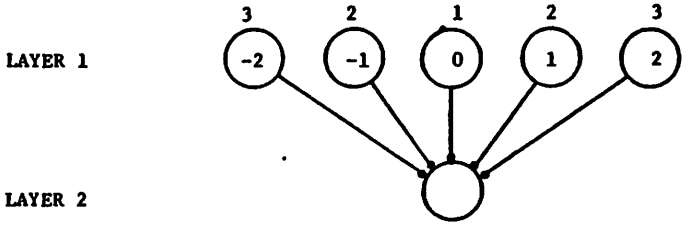
A variety of functions are available for driving the variable weights. Each set of variable connections is given a function that is then used with all of the weights in that set. Further, each left-right pair of a given set of variable weights is given upper and limits. If a weight value moves beyond these limits it is allowed to continue, but the value used in calculation of the input to the associated cell remains equal to the limit crossed by the weight.

This scheme of variable connections was designed as the simplest way to model heterosynaptic facilitation and depression, or a common elements effect. Note that activity in one of a set of connections

can alter the values of connections only in the same set and only those connections on the same cell. This network could be extended to allow activity in any set of connections to alter the values of any other set, on any other cell. Simulation of dendro-dendritic or axo-axonic synapses would then be possible.

Using the program. The program is designed to allow the user to see curves of relative habituation generated by the network he has specified. The network is specified in the parameter input section of the program. This section is set up such that single parameters may easily be altered during a run. The calculation and output sections run the network and display the result. In this section I describe the way networks are set up and run, and the way the output is specified.

A. Template specification. The connection template for each layer is specified in two steps. First a number is given for the total spread of connections in both directions between each pair of layers. If this number is zero, there are no connections between the layers. If it is one, each cell in the output layer connects to the cell with the same number in the input layer, and so on. This is illustrated below.



Numbers on cells in layer 1 give numbering relative to the cell in layer 2. Numbers above the layer 1 cells are the numbers of the associated connections in the template.

At the point in the program where these numbers are to be given, the message

INPUT NUMBER OF CONNECTIONS TO LAYER i FROM EACH LAYER

is printed. The numbers of connections to layer i from all other layers are typed in integer form, for each layer i.

Next the program asks for the weight values associated with each set of connections specified in the templates for each layer, with the message:

INPUT WEIGHT REST VALUES:
FROM LAYER i TO LAYER j

Real numbers are to be typed in for each set of connections. These are the values of the fixed weights and the resting values of the variable weights.

B. Modifiable weights. Each set of weights must be specified as fixed or variable, and the modification parameters of variable weights must be given. Variable weights are referred to using a rather complicated set of conventions described more or less opaquely below.

The program first prints:

INPUT NO. VARIABLE WTS. FOR TYPE SPECIFICATION

If no weights are to vary, a zero is entered. Otherwise, the number of variable weights is given. In that case the message

INPUT WT. LAYERS, TYPE

is printed. Integers of the form I, J, K are to be given, where I is the input layer of the connection, J is its output layer. K gives the type of the synapse, and refers to the number of the function used in varying the connection (see equation on p. 271). The available functions are:

TYPE	FUNCTION FORM
1	1.0
2	1.0/(1.0 + C * EXP(-D * F))
3	$W - A_1$
4	$1.0 - A_1/W$
5	
6	F
7	$F*(W - A_1)$

Here F is MAX(0.0, presynaptic cell value-weight modification threshold)

G is postsynaptic cell firing level

W is current synapse value

A_1, A_u are the weight lower and upper asymptotes

C, D are the weight sigmoid parameters

Fixed connections are given Type 0.

Note that if the firing level of the presynaptic cell falls below the modification threshold of the synapse, no modification is made.

From the types of the synapses a Type Matrix is constructed. The Type Matrix looks like:

$$\begin{matrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{matrix}$$

Entry i, j in this matrix gives the type of connections from layer j to layer i. Each varying connection is also given an Index according to its place in this matrix. The Index of a given set of connections is found by counting the number of nonzero entries in the Type Matrix, beginning at the upper left and proceeding across columns and down rows until the entry corresponding to the set of connections in question is reached. For example, in the matrix

$$\begin{matrix} 1 & 2 & 0 \\ 0 & 4 & 0 \\ 1 & 0 & 3 \end{matrix}$$

the set of connections from layer 2 to layer 1 has type 2 and index 2, and the connections from layer 2 to itself have type 4 and index 3. Variable connections are referred to in terms of their indices in the rest of the program.

Modification parameters for each set of variable weights must be specified. Weight time constants are given following the message

INPUT WEIGHT TIME CONSTANTS:
FROM LAYER i TO LAYER j

Each set of variable weights has an associated set of upper and lower limits, upper and lower asymptotes, and modification thresholds used in the modification functions described earlier. These are put in following the message

INPUT WEIGHT LOWER AND UPPER LIMITS, ASYMS., THRS.
FROM LAYER i TO LAYER j

Sequences of numbers with the form LL, UL, LA, UA, THR for each weight from layer i to layer j, where

LL is the weight lower limit,
UL is the weight upper limit,
UA is the weight upper asymptote,
LA is the weight lower asymptote, and
THR is the weight modification threshold.

All of these parameters for the weights of a given set of connections must be put on a single line.

Type 2 weights have sigmoidal modification functions that are specified for each Type 2 weight in terms of its value at two points selected by the user. The message

INPUT TYPE 2 WEIGHT SIGMOID POINTS, VALUES.

is printed, and the lower point and the sigmoid value there are given, followed by the upper point and its associated value. The sigmoid parameters are then calculated and displayed by the program.

Finally, the modification template and modification gains are put in for each variable set of connections. For each index, the message

INPUT NUMBER OF MODIFICATION GAINS FOR INDEX i (FROM j TO l)

is printed. The letters j and l refer to the layers associated with the set of connections with index i. A number is to be given, following the convention for template specification. Then the message

INPUT MODIFICATION GAINS FOR EACH WEIGHT WITH INDEX i, ONE AT A TIME.

is printed, and the modification gains for each weight are put in by the user. Note each modification gain is numbered according to the index of the set of connections being specified, the number in the connection template of the weight to which it belongs and its number in the modification template for that weight.

At this point in the program, the weights have been fully specified.

C. Cell specification. The cell's output functions are specified here. The functions are sigmoids, and are specified in the same way the Type 2 weight sigmoids are specified. The message

INPUT CELL SIGMOID POINTS, VALUES

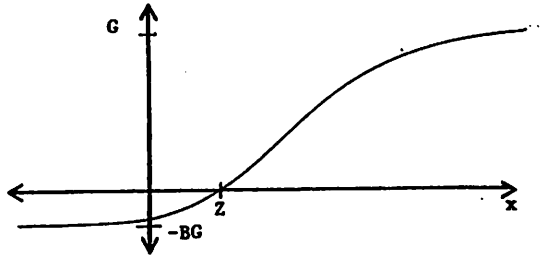
is printed, and upper and lower points and their associated values between zero and one are given, following which the sigmoid parameters are printed out. The program then prints

INPUT CELL SIGMOID GAIN, THR., ZERO (CR)

The cell gain is a multiplier of the sigmoid, which is itself normalized to one. The threshold and zero of the sigmoid are best explained in terms of the sigmoid equation:

$$S(x) = G\left[\frac{1}{1 + c\, e^{-d(x - x_0)}} - \frac{1}{1 + c\, e^{-d(z - x_0)}}\right] = G(S_1(x) - B)$$

This looks like:



The threshold, x_0, is used to shift the sigmoid along the x-axis, and the "zero," z, fixes the point at which the sigmoid crosses the axis. The cell outputs are calculated as

$$O(x) = MAX\ (S(x),\ 0.0)$$

so the zero point given above fixes the point at which the cell output becomes zero. The value of the bias, B, is displayed by the program, for each layer of cells.

At this point the network is completely specified and may be run. It remains to specify inputs and outputs.

D. Stimulus specification. Each stimulus periodically excites a number of cells in the network for a particular duration of time. The cells excited by an input, and the input's duration and period are specified by the user at this point. The program prints

INPUT NO. INPUTS, NO. CELLS PER INPUT, THEN CELLS.

The user first gives the total number of sets of cells to be stimulated, then the number of cells in each set. Following that, the numbers of the cells stimulated in each set of cells are given on separate lines. Note that only the cells in layer 1 may be stimulated from outside the network.

For example, if there are to be two inputs, where input 1 excites cells 1 and 2, and input 2 excites cell 3, the user types

2, 2, 1

1, 2

3

Next, the program asks for the duration and period of each input, which are given in terms of numbers of iterations of the program's calculation section. These numbers are to be given on a single line.

Stimulus intensity is put in following the message

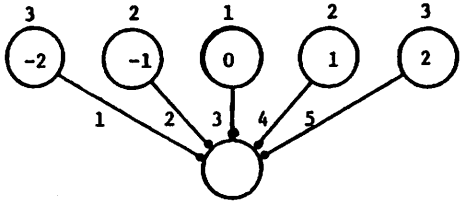
INPUT STIMULUS INTENSITY (CR), 1 TO RESTORE.

"1 TO RESTORE" refers to initialization of the network, and will be explained later (see section H, below). During the initial period of parameter specification this feature is inoperative.

E. Output specification. The program outputs in graphical form the time-courses of selected variable synapses, sampled periodically. The number of weights to be graphed, and the numbers that specify them are input here. The program prints

INPUT NO. WEIGHTS FOR TIME, WEIGHT INDEX, CELL, NUMBER.

The total number of weights to be output is given, then each weight's index, the number of the cell with which it is associated in its input layer, and its absolute number are given. The absolute number of a variable weight refers to its place in the associated weight template, counting from left to right. This is illustrated below, where numbers on and above cells are as in the figure on p. 273, while numbers on the connections are the connection absolute numbers.



Recall that each left-right pair of connections in a template has the same set of modification parameters and resting value, but that each may vary as an individual. The absolute and cell numbers are the way the program keeps track of the individual connections.

Cells are elected for habituation curve output following the message

INPUT NO. CELLS FOR HABITUATION CURVES, CELL LAYER, NUMBER.

These numbers are all entered on a single line. The total output of each of these cells in response to each stimulus presentation is stored for further display.

Now the network form and its inputs and outputs are fully specified. All that remains is to give a set of program control parameters and to initialize the network.

F. Program control modes and initialization. The message

INPUT NTIMES, 1 FOR FREEZE, 1 FOR PROD. MODE., 1 FOR FROZEN CONTROL.

is printed. NTIMES is the period at which weight values are recorded for later graphical output. The program stores the last 40 values of the weights, sampled at this period.

If the next parameter is 1, all variable weights are held at their current values (the network is frozen). A zero here allows the weights to change.

If "Prod. Mode" is set to 1, the program will run until a number of stimuli specified below have been presented to the network, at which time the habituation curves are automatically plotted. If

this parameter is zero, the program runs for a specified number of iterations.

"Frozen Control" refers to two ways of plotting the habituation curves. The first stimulus presentation is used as the control level stimulus. The network response to this stimulus is used in comparison with the responses to the other stimulus presentations to plot the curves of relative habituation. If the "Frozen Control" parameter is set to 1, the network is frozen during the first stimulus presentation, simulating an entirely unhabituated network. If this parameter is zero, the control level includes the effect of the first habituating stimulus.

Note all of these parameters are integers, and are put in on a single line. The network is now initialized, and the message

SYSTEM RESTORED

is printed, followed by

RUN NUMBER 1

The run number is updated each time the system is restored, and is used for purposes of reference.

The network is now ready to roll.

G. The main control point (MCP). The program is now at the main control point, from which all things are possible. From here

single parameters may be altered, the network may be reset or run, and outputs may be called for. The message printed here is

INPUT NUMBER OF ITERATIONS DESIRED, STIMULUS NUMBER,
NUMBER OF STIMULUS PRESENTATIONS DESIRED.

The number of iterations given controls the program's action. If this number is greater than zero, the network is run for the specified amount of time when the "Prod. mode" parameter is set to zero. If "Prod. Mode" is 1, the network is run until the desired number of presentations of the given stimulus have been given. If the number of iterations given is zero, a complete new set of parameters is called for by the program. Numbers less than zero are used to select single parameters or sets of parameters for change, according to the following directory:

NUMBER OF ITERATIONS	PARAMETER
-1	Weight and Habituation Curve Outputs
-2	System Initialization
-3	NTIMES, PROD. MODE, FREEZE, FROZEN CONTROL
-4	Cells for Habituation Curves
-5	Weights for Time Output
-6	Stimulus Intensity
-7	Stimulus Cells, Periods, Durations
-8	Full Cell Sigmoids
-9	Partial Cell Sigmoids
-10	Full Weight Modification Gains

-11	Partial Weight Modification Gains
-12	Type 2 Weight Sigmoid Parameters (Full)
-13	Partial Type 2 Sigmoid Parameters
-14	Full Weight Limits, Asymptotes, Thresholds
-15	Partial Weight Limits, Asymptotes, Thresholds
-16	Full Weight Time Constants
-17	Partial Weight Time Constants
-18	Type Matrix Entries
-19	Full Weight Values
-20	Partial Weight Values
-21	Template Size Specification
-22	Number of Layers, Cells Per Layer, Time Step

Here "partial" refers to the ability to change single values in a set of parameters, rather than having to input a full set of parameters. Messages printed out at these points are generally self-explanatory (hopefully).

H. Secondary control points. In order to make certain sequences of operations easier, secondary control points exist throughout the program. These points generally involve the ability to alter a parameter or set of parameters, then automatically initialize the system. The message

1 TO RESTORE

is printed at such points. A one put in at that time will cause the network to be initialized after the parameter value is altered. A zero returns the control directly to the main control point.

A special control point exists at the end of the habituation curve output section. A control parameter is asked for at that point, taking the following values:

CONTROL PARAMETER	ACTION TAKEN
0	Return to MCP.
1	Go to Specify Input Intensity.
2	Initialize Network, then go to MCP.
3	Go to MCP; next stimulus is treated as new control stimulus for further stimulus presentations.
4	Go to MCP, clear graph of habituation curves.
5	Print "RESET NO. STIM. TO DATE." Input a number less than the total number to date at which to continue plotting habituation curves. This is used to erase points from the habituation graphs, and is useful in generating clean graphs of dishabituation experiments.

A number is printed following the initial message. This number is the size of the habituation graph being displayed, and may not exceed 701. If it nears 701, a 4 should be given here to clear the graph.

I. System filing. To make using the program somewhat easier, the user may store all the parameters necessary to run the program and read them in to the program at any time. In this way the full set of parameters need not be entered by the user every time the program is started.

When first started, or after a zero is put in for number of iterations at the main control point, the program prints

INPUT 1 TO READ STORED PARAMETERS.

A 1 then will read the parameters and initialize the system. A zero allows the user to put in fresh parameters.

In the Output section of the program, the user may write the current parameter values onto the disc over the values stored there. Unfortunately, the program stops there, due to the way the FORTRAN on this machine is set up. However, the program may easily be restarted and run with the parameters just stored.

J. Network calculation. Once a network's inputs and outputs have been specified, calculation may proceed. At each iteration the input to each cell is computed using cell outputs and weight values

from the previous iteration. The outputs of the selected cells and values of selected weights are stored for later graphical output. The values of the variable weights are updated according to the cell firing levels from the previous iteration. Updates are handled according to an Eueler integration formula with a time step specified by the user. The calculation continues for a specified amount of time, or until a specified number of inputs have been presented to the network.

Part III: An example. Below is an example of an actual run of the program showing how inputs are put in, and displaying the form of the output. A set of habituation curves is generated, then a dishabituation experiment is performed. The network involved is the simple three-cell system shown in the figure on p. 269.

```
 PU 1 HABITS

INPUT 1 TO READ STORED PARAMETERS
0

INPUT NUMBER OF LAYERS, NUMBER OF CELLS PER LAYER(S), TIME STEP SIZE.
3,1
.1

INPUT NUMBER OF CONNECTIONS TO LAYER 1 FROM EACH LAYER
0,0,0

INPUT NUMBER OF CONNECTIONS TO LAYER 2 FROM EACH LAYER
1,0,0

INPUT NUMBER OF CONNECTIONS TO LAYER 3 FROM EACH LAYER
1,1,0

INPUT HEIGHT PLOT VALUES
  FROM LAYER 1 TO LAYER 2
1

  FROM LAYER 1 TO LAYER 3
1.

  FROM LAYER 2 TO LAYER 3
.1

INPUT NO. VARIABLE BIAS FOR TYPE SPECIFICATION
3

INPUT NT. LAYERS, TYPE.
2,1,3,3,1 3,3,2,2

INPUT HEIGHT TIME CONSTANTS
  FROM LAYER 1 TO LAYER 2
25.

  FROM LAYER 1 TO LAYER 3
20

  FROM LAYER 2 TO LAYER 3
7.

INPUT HEIGHT LOWER AND UPPER LIMITS, ASYNS., THRS.
  FROM LAYER 1 TO LAYER 2
.5,1.,.5,1.,0

  FROM LAYER 1 TO LAYER 3
.1,1.,.1,1.,0

  FROM LAYER 2 TO LAYER 3
0,.2 ,0.,.3 ,0

INPUT TYPE 2 HEIGHT SIGMOID POINTS, VALUES.
1.,.05,10.,.95

SIGMOID PARAMETERS, HEIGHTS 2
26.5528  0.6542
INPUT NUMBER OF MODIFICATION GAINS FOR INDEX 1 (FROM 1 TO 2)
1

INPUT MODIFICATION GAINS FOR EACH HEIGHT WITH INDEX 1 -ONE HEIGHT AT A TIME.
-10.
```



```
INPUT NUMBER OF MODIFICATION GAINS FOR INDEX 2 (FROM 1 TO 2)
1

INPUT MODIFICATION GAINS FOR EACH HEIGHT WITH INDEX 2 -ONE HEIGHT AT A TIME.
-20

INPUT NUMBER OF MODIFICATION GAINS FOR INDEX 3 (FROM 1 TO 2)
1

INPUT MODIFICATION GAINS FOR EACH HEIGHT WITH INDEX 3 -ONE HEIGHT AT A TIME.
15

INPUT CELL SIGMOID POINTS, VALUES
0 , 05 10 , 95
0 , 05 10 , 95
0 , 05 10 , 95

SIGMOID PARAMETERS
        C           D        MAX SLOPE PT.
    15.00000      0.50000      5.00000
    15.00000      0.50000      5.00000
    15.00000      0.50000      5.00000
INPUT CELL SIGMOID GAIN, THR, SET(S)(S)
1 ,0 ,0
1 ,0 ,0.
1 ,0 ,0.

SIGMOID BIASES
 0.0500
 0.0500
 0.0500
INPUT NO. INPUTS, NO. CELLS PER INPUT, THEN CELLS
3,1,1,1
1
1
1

INPUT DURATION, PERIOD EACH INPUT.
1,21,6,21,15,65

INPUT STIMULUS INTENSITY(S), 1 TO RESTORE
1

INPUT NO. HEIGHTS FOR TIME OUTPUT, HEIGHT INDEX, CELL, NUMBER
3,1,1,1,2,1,1,2,1,1

INPUT NO. CELLS FOR HABITUATION CURVES, CELL LAYER, NUMBER
2,2,1,2,1

INPUT NTIMES, 1 FOR FREEZE, 1 FOR PROG. MODE , 1 FOR FIXED CONTROL
10,0,1,1

SYSTEM RESTORED

    RUN NUMBER   1

INPUT NUMBER OF ITERATIONS DESIRED, STIMULUS  NUMBER,
  NUMBER OF STIMULUS PRESENTATIONS REQUIRED
-1,0,0

INPUT 1 FOR TIME OUTPUTS, 2 FOR HABITUATION GRAPHS, 3 TO FILE PARAMETERS.
3

FILING
 ONE
```

```
IN

RU 1 UNITS

INPUT 1 TO READ STORED PARAMETERS.
1

SYSTEM RESTORED

    RUN NUMBER   1

INPUT NUMBER OF ITERATIONS DESIRED, STIMULUS  NUMBER,
  NUMBER OF STIMULUS PRESENTATIONS DESIRED
1,2,10

WEIGHTS AT TIME  211
  CELL  2  1
    LAYER 1
  0.65688
  CELL  3  1
    LAYER 1
 0.25874
    LAYER 2
  0.25241
HABITUATION CURVE DATA
CELL  2  1    6.6000
  0.952838    0.862072    0.796825    0.749922
 0.716206    0.691970    0.674547    0.662023
 0.653020
CELL  3  1    6.6000
  0.839550    0.592838    0.477045    0.421245
 0.394586    0.381580    0.375100    0.371752
 0.369927

INPUT CONTROL PARAMETER   208
1

INPUT STIMULUS INTENSITY(CR), 1 TO RESTORE.
8.
1

SYSTEM RESTORED

    RUN NUMBER   2

INPUT NUMBER OF ITERATIONS DESIRED, STIMULUS  NUMBER,
  NUMBER OF STIMULUS PRESENTATIONS DESIRED
5,2,10

WEIGHTS AT TIME  211
  CELL  2  1
    LAYER 1
  0.65688
  CELL  3  1
    LAYER 1
  0.25874
    LAYER 2
 1.97000
HABITUATION CURVE DATA
CELL  2  1   48.0000
  0.952838    0.862072    0.796825    0.749922
 0.716206    0.691970    0.674547    0.662023
```



```
CELL  3  1   52.8000
  1.181374    1.417650    1.559463    1.601212
 1.594182    1.575261    1.542098    1.506729
 1.472870

INPUT CONTROL PARAMETER   260
4

INPUT NUMBER OF ITERATIONS DESIRED, STIMULUS  NUMBER,
  NUMBER OF STIMULUS PRESENTATIONS DESIRED.
-6,0,0

INPUT STIMULUS INTENSITY(CR), 1 TO RESTORE.
1.
1

SYSTEM RESTORED

    RUN NUMBER   3

INPUT NUMBER OF ITERATIONS DESIRED, STIMULUS  NUMBER,
  NUMBER OF STIMULUS PRESENTATIONS DESIRED.
1,2,10

WEIGHTS AT TIME  211
  CELL  2  1
    LAYER 1
  0.65688
  CELL  3  1
    LAYER 1
  0.25874
    LAYER 2
  0.25241
HABITUATION CURVE DATA
CELL  2  1    6.6000
  0.952838    0.862072    0.796825    0.749922
 0.716206    0.691970    0.674547    0.662023
 0.653020
CELL  3  1    6.6000
  0.839550    0.592838    0.477045    0.421245
 0.394586    0.381580    0.375100    0.371752
 0.369927

INPUT CONTROL PARAMETER   208
0\0\1

INPUT STIMULUS INTENSITY(CR), 1 TO RESTORE
10
0

INPUT NUMBER OF ITERATIONS DESIRED, STIMULUS  NUMBER,
  NUMBER OF STIMULUS PRESENTATIONS DESIRED.
1,2,11

WEIGHTS AT TIME  233
  CELL  2  1
    LAYER 1
  0.65167
  CELL  3  1
    LAYER 1
  0.25824
    LAYER 2
```

APPENDIX C

PROBABILITIES OF INPUT ENTRY AND ASSOCIATION

I calculate here the probability that an input stimulates no cell that lies in the trough of a wave, and the probability that a given cell stimulated by an input becomes associated with at least one wave in the system. The following parameters are used:

N_L = number of lamellae,

N_C = number of cells/lamela,

N_H = number of cells stimulated by an input,

N_W = number of waves in the system,

M_W = total number of cells/wave,

M_T = number of cells in a wave trough,

M_R = number of cells in another lamella that a cell in a given lamella is able to form connection to,

$P_Z(i,j)$ = probability that a longitudinal line from lamella i reaches lamella j,

P_Z = probability that a longitudinal line makes contact with a given cell (out of a total of M_R) in a lamella it reaches,

P_E = probability that a wave can activate a given longitudinal line.

To find the probability that an input will stimulate a cell in the trough of a wave, note that each wave trough uses M_T cells, so that the total number of depressed cells is $N_W \cdot M_T$. So

p(a given stimulated cell lies in a wave trough) =

$$P_{ST} = \frac{N_W \cdot M_T}{N_L N_C}$$

If waves and stimulated cells are distributed at random through the system,

p(k stimulated cells fall in wave troughs) =

$$\binom{N_H}{k} P_{ST}^{k} (1 - P_{ST})^{N_H - k}$$

So the probability that all stimulated cells fall outside wave troughs is

$$P_{FE} = (1 - P_{ST})^{N_H}$$

To find the probability that a given stimulated cell is connected with at least one wave in the system, we assume first that an input is equally likely to stimulate a cell in any of the N_L lamellae. Then

p(a stimulated cell is connected to at least one wave when there are N_W waves in the system) =

$$P_H(N_W) = \frac{1}{N_L} \sum_{i=1}^{N_L} P_C(i, N_W),$$

where

$P_C(i, N_W)$ = p(a stimulated cell in lamella i is connected to at least one wave, given N_W in the system).

The waves may be distributed through the network in many ways. Each such distribution will be called a configuration, and configurations will be denoted C_j, $j = 1, \ldots, M_C$, where M_C is the total number of configurations of N_W waves distributed in N_L lamellae. Associated

with each C_j is a set of numbers $K_j = (k_{j1}, \ldots, k_{jN_L})$ giving the number of waves in each lamella for that configuration.

Hence

$$P_C(i, N_W) = \sum_{j=1}^{M_C} P_C(i, N_W|C_j)\, P(C_j),$$

where

$P(C_j)$ = probability that configuration C_j occurs,

Now

$P_C(i, N_W|C_j)$ = p(a stimulated cell in lamella i is connected with at least one wave in the system given N_W waves in configuration C_j)

= 1 - p(a stimulated cell in lamella is connected with no wave, given N_W waves in configuration C_j)

= $1 - P_W(i, N_W|C_j)$,

where

$$P_W(i, N_W|C_j) = \prod_{m=1}^{N_L} [1 - P_A(m, i)]^{k_{jm}}$$

and

$P_A(m, i)$ = p(a given wave in lamella m makes contact with a given stimulated cell in lamella i),

$P_A(m, i)$ = p(the wave is at a cell in lamella m that sends a z-line to the stimulated cell in lamella i) · p(the E-cell associated with the z-line can be activated),

$$= \frac{M_R}{N_C} \cdot F_z(m, j) \cdot P_E.$$

The first two terms in the final expression give the probability of horizontal and vertical connection from a cell in lamella m to one in lamella j, assuming that each G-cell is connected to exactly one E-cell in its lamella.

These expressions are evaluated for a system with parameters given in Table C.1, with the results shown in Table C.2.

TABLE C.1

Parameters Used in Evaluation of Association

and Input Entry Probabilities

$N_L = 6$

$N_C = 60$

$M_R = 30$

$M_T = 3$

$P_z = .9$

$P_E = .66$

$F_z(i,j) = 1.0;\ 0 < |i-j| < 3$

$0.0;$ otherwise

TABLE C.2

N_S \ N_W	6	7	8	9	10	11	12
1	0.71 / 0.95	0.76 / 0.94	0.79 / 0.93	0.82 / 0.92	0.85 / 0.91	0.87 / 0.91	0.89 / 0.90
2	0.50 / 0.90	0.57 / 0.88	0.63 / 0.87	0.68 / 0.85	0.73 / 0.84	0.76 / 0.82	0.79 / 0.81
3	0.36 / 0.86	0.43 / 0.83	0.50 / 0.81	0.57 / 0.79	0.62 / 0.77	0.67 / 0.75	0.71 / 0.73
4	0.25 / 0.81	0.33 / 0.78	0.40 / 0.76	0.47 / 0.73	0.53 / 0.71	0.59 / 0.68	0.63 / 0.65

Association and Input Entry Probabilities

Entries are of the form: probability of association/probability of entry