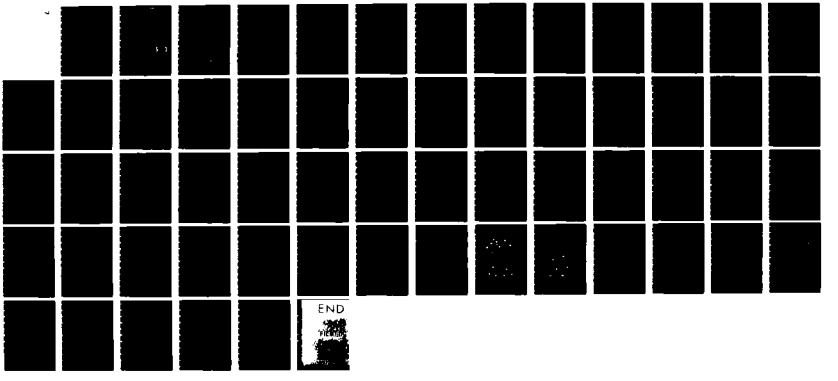
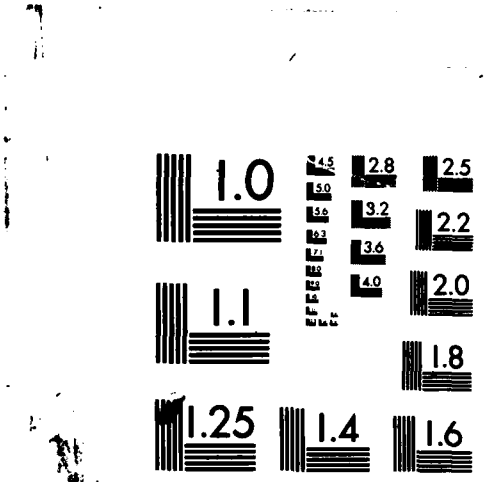
AD-A138 081 A NEURAL NETWORK MODEL OF MULTISTABLE PERCEPTION(U) 1/1
BROWN UNIV PROVIDENCE RI CENTER FOR NEURAL SCIENCE
A H KAWAMOTO ET AL. 15 FEB 84 TR-12 N00014-81-K-0136
UNCLASSIFIED F/G 5/10 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(12)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
#12	AD-A138081	

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
A NEURAL NETWORK MODEL OF MULTISTABLE PERCEPTION	TECHNICAL REPORT
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
Alan H. Kawamoto and James A. Anderson	N00014-81-K-0136

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
CENTER FOR NEURAL SCIENCE BROWN UNIVERSITY PROVIDENCE, RHODE ISLAND 02912	NR-201-484

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Personnel and Training Research Program Office of Naval REsearch (Code 442PT) Arlington, Virginia, 22217	February 15, 1984
	13. NUMBER OF PAGES 57

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. ~~Publication in whole or in part is permitted for any purpose of the United States Government.~~

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTED FEB 21 1984 S B

18. SUPPLEMENTARY NOTES

Submitted for publication: Acta Psychologica.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Concepts
multistable perception
necker cube
neural models
neural nets

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, the major properties and previous models of multistable perception are briefly reviewed. A neural network model based on Hebbian synaptic modification (the brain-state-in-a-boc model of Anderson and colleagues) is shown to satisfactorily account for a number of these properties. We present evidence demonstrating the importance of both the stimulus and the history (both recent and distant) of the system of disambiguate ambiguous stimuli. In addition, some simple extensions are made to allow the dynamic modification of synaptic connectivities during

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

84 02 17 012

DTIC FILE COPY

AD A138081

the course of the stimulus presentation. This enables such properties as the time course of reversals, adaptation, and hysteresis to be simulated.

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

DTIC COPY INSPECTED 3

A Neural Network Model of Multistable Perception

Alan H. Kawamoto

and

James A. Anderson

Department of Psychology

Brown University

Providence, RI., 02912 (U.S.A.)

running title: Multistable Perception

reprint requests: Prof. James A. Anderson
Box 1853
Dept. of Psychology
Brown University
Providence, RI 02912, USA

Abstract

In this paper, the major properties and previous models of multistable perception are briefly reviewed. A neural network model based on Hebbian synaptic modification (the brain-state-in-a-box model of Anderson and colleagues) is shown to satisfactorily account for a number of these properties. We present evidence demonstrating the importance of both the stimulus *and* the history (both recent and distant) of the system to disambiguate ambiguous stimuli. In addition, some simple extensions are made to allow the dynamic modification of synaptic connectivities during the course of the stimulus presentation. This enables such properties as the time course of reversals, adaptation, and hysteresis to be simulated.

A Neural Network Model of Multistable Perception

"The mind cannot fix long on one invariable idea."

John Locke

Introduction

Ever since Necker's report of an observation "which occurs on viewing a figure of a crystal or geometrical solid" a century and a half ago (Necker 1832), figures and drawings which can be perceived in two or more stable configurations have held a curious fascination for numerous people investigating form perception. During this time, a wide variety of stimuli displaying multistability and factors influencing their organization have been reported.

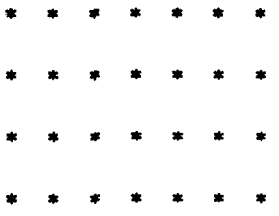
Although there have been numerous reports describing new multistable stimuli, theoretical treatments of this phenomenon have been few (Price 1969). Of the existing approaches, the most widely accepted neural basis is the "satiation hypothesis" proposed by Koehler and Wallach (1944). In the present study, a mathematical model based on Hebbian synaptic modification in a neural network (Anderson, Silverstein, Ritz, and Jones 1977) is used to explore properties of multistable perception.

Before examining the details of the implementation, a brief description of the various types and major properties of visual stimuli which are multistable will be given. This will be followed by a brief review of models which have been previously developed to account for multistable perception. Next, an outline of the neural model will be described and a few simple extensions will be introduced. Then, properties of multistable stimuli will be described in some detail and the manner in which the model explains them will be demonstrated.

Multistable Visual Stimuli

In this review, multistable stimuli will be defined as those visual stimuli which can be perceived in one of two possible configurations and whose perceived organization can, without any physical change of the stimulus, change from one stable configuration to another. We shall restrict our attention to perceptual changes which are not due to "suppression" of parts of a stimulus, but to a global organizational change.

The majority of visual stimuli displaying multistability can be grouped into one of four classes. These are (1) reversible perspective figures, (2) ambiguous figures, (3) figure-ground reversal figures, and (4) gestalt stimuli. Reversible perspective figures include the Necker cube, Mach book, Schroeder staircase, and the Lissajous curve. Examples of ambiguous figures are Boring's wife/mother-in-law, Jastrow's duck/rabbit, and Rock's dog/chef. Rubin's face/vase and Escher's "Devils or Angels" are examples of figure-ground reversal figures. An example of gestalt organization, many of which were introduced by Wertheimer (1938), is illustrated below.



Specifically *excluded* from consideration here will be apparent luminance changes of the stimulus as these do not lead to changes in the perceived *configuration*. Also excluded are those stimuli composed of parts which are all physically present but selectively suppressed such as retinally stabilized images (Pritchard, Heron, and Hebb 1960; Riggs et al. 1953), rivalrous binocular stimuli, and ambiguous random dot stereograms (Julesz

1971). However, we will briefly discuss at the end of this paper, the change in the perceptual organization due to changes in the context of a stimulus. An example would be homographs such as "bat" in nocturnal *bat* and wooden *bat*. In this example, the context disambiguates the stimulus.

With the scope of our study focussed on one class of ambiguous stimuli, let us examine some of the major experimental properties this class of stimuli shows.

(1) There must be two or more possible configurations of the stimulus.

(2) Once a particular configuration is perceived, it remains stable for a period of time before spontaneously changing to the alternative configuration.

(3) The relative amount of time the two different configurations are perceived is a direct function of the probability that a given configuration will be perceived initially (Price 1967b; Sadler and Mefferd 1970).

(4) The rate of fluctuation increases during the first 3 minutes of viewing before reaching an asymptote (Brown 1955; Cohen 1959).

(5) After adaptation to an unambiguous stimulus, an ambiguous stimulus is more likely to be perceived in the alternative configuration (Howard 1961; Virsu 1975).

(6) When viewing a set of stimuli ranging from one unambiguous configuration to another sequentially, hysteresis is observed. That is, the first configuration is maintained beyond where it would have normally ceased to have been evoked (Attneave 1971).

Previous Models

Perhaps the most generally accepted model for these phenomena is the neurally based satiation model first proposed by Koehler and his colleagues (Koehler 1940; Koehler and Wallach, 1944). Spitz (1963), Orbach, Ehrlich, and Heath (1963), and Howard (1961) have also developed models similar to the original "satiation" theory. Based on a mechanism to account for figural after-effects, Koehler and Wallach (1944) proposed,

> According to the first hypothesis a specific figure process occurs whenever a figure appears in the visual field. And this process tends to block its own way if the figure remains for some time in the same location. The second hypothesis states that in reversible figures, the figure-ground relationship will suddenly be reversed when the figure process has altered the medium beyond a critical degree.

An early paper published by McDougall (1906) is close in spirit to our approach. McDougall claimed that "reversible perspective is a special case of a general psychic characteristic, viz. the rapid fatiguability manifested in all the higher levels of consciousness, due to rapid changes in synaptic resistance." Similarly, Howard (1961) regarded satiation as an "auto-inhibitory process occurring at specific synapses in the central nervous system as a result of activity set up there by specific sensory events."

During the past decade, several investigators have presented models based on decision processes in the perceptual system. Vickers (1972) proposed that the initial percept is mediated by an *"optional stopping* decision process" and that the reversals might be accounted for by supposing that the decision process was *cyclic.* This model is based on an observer inspecting a distribution of signal differences (rather than the two original representations) until some criterion is reached. Taylor and Aldridge (1974) proposed a random-walk model based on a "finite majority decision device with memory" related to Selfridge's Pandemonium model (Selfridge 1959). In this version, each "demon" or "cell" selects one of the two possible configurations but can switch decisions instantaneously.

Theoretical Development

Over the past fourteen years one of us (J.A.) has been developing a neural network model based on Hebbian synaptic modification (Anderson 1968; Anderson 1972; Anderson, Silverstein, Ritz, and Jones 1977). Similar approaches have been discussed by Cooper (1974) and Kohonen (1977).

In this section, this approach will be briefly reviewed. First, a brief account of the neurophysiological inspiration and basis will be sketched and the associative memory scheme will be described. Next, the "brain-state-in-a-box" model incorporating non-linear saturating properties of the neural network will be examined. Finally, some simple extensions of the brain-state-in-a-box model will be presented. In the present study, the formalisms and notation in Anderson et al. (1977) will be used.

Associative model.

Two striking properties of cortical organization are its anatomical parallelism and the large degree of interconnection between neurons. Functionally, this results in a highly interconnected, parallel system spatially distributed over a large extent. A neuron is an element in this system and behaves as an analog integrating device.

We assume that

> (a) nervous system activity is usefully represented as the set of simultaneous individual neuron activities in a group of neurons; (b) different memory traces make use of the same synapses; and (c) synapses associate two patterns of neural activity by incrementing synaptic connectivity proportionally to the product of pre- and postsynaptic activity, forming a matrix of connectivities (Anderson et al. 1977, p. 413).

Based on these assumptions, the association of an input vector (stimulus) with an output vector (response) can be modeled as a linear system.

As depicted in fig. 1, a system consisting of two sets, alpha and beta, of N neurons, can be constructed.

---

insert fig. 1 about here

In this figure each of the N neurons in the set alpha synapses on each neuron in beta. Corresponding to each synaptic junction is a synaptic strength, a_{ij}, coupling the *jth* neuron in alpha with the *ith* neuron in beta. Assuming linear integration, the activity of the *ith* neuron in beta is given as

$$g(i) = \sum_j a_{ij} f(j).$$

That is, a given beta cell's firing rate is simply the weighted sum of all the inputs from alpha.

By making a further assumption, that the input vectors, f, are orthonormal, it is easy to demonstrate that an output pattern, g, can be associated with its corresponding input pattern. We assume: (1) the dot product of any input vector with itself is 1 (normality), and (2) the dot product of any input vector with any other input vector is 0 (orthogonality). That is,

$$f_i^T f_i = 1, \quad i = 1, 2, \ldots, K, \text{ and}$$
$$f_i^T f_j = 0, \quad i, j = 1, 2, \ldots, K \text{ and } i \neq j,$$

where ^{T} is the transpose operation. Furthermore, since the dot product of a vector with itself is defined as the length of the vector, each input vector has unit length, with length defined as

$$\| f \| = \left(\sum_i f(i)^2 \right)^{\frac{1}{2}}.$$

If the synaptic weights are assumed to be zero initially, it can be shown that the final values after learning are a function only of the inputs, f_k, their associated outputs, g_k, and their frequency of presentation, p_k. For this synaptic learning (modification) we postulate that to associate pattern f_1 in alpha with g_1 in beta, we need to change the set of synaptic weights according to the product of presynaptic activity at a junction with the activity of the postsynaptic cell.

These connections can be formally described as a matrix A_k, with

$$A_k = g_k \times f_k^T .$$

In this matrix, each element in the m*th* row corresponds to the synaptic weight between the m*th* cell in beta with each of the N cells in alpha. Similarly, each element in the n*th* column of the matrix represents the synaptic weight between the n*th* cell in alpha with each of the N cells in beta.

If we regard the matrix A as a linear operator operating on the input f_1, the response is given as

$$\begin{aligned} A_1 f_1 &= (g_1 f_1^T) f_1 \\ &= g_1 (f_1^T f_1) \\ &= g_1 , \quad \text{from normality assumption.} \end{aligned}$$

Thus, an input f_1 results in the desired output, g_1.

Suppose now, that in addition to associating f_1 and g_1, K such pairs (f_1,g_1), (f_2,g_2), ..., (f_K,g_K), are associated. We now have, after superimposing the effect of each pair of associations on the synaptic junctions, the set of connections given as

$$A = A_1 + A_2 + \ldots + A_K .$$

Even after learning K associations, inputting any of the original input vectors will yield the correct output response if K is less than N because of the orthogonality assumption. For example, we apply our new operator, A, to one of the original input vectors, f_i, to get

$$\begin{aligned} Af_i &= \left(\sum_{k=1}^{K} A_k\right) f_i \\ &= A_i f_i + \left(\sum_{k \neq i} A_k\right) f_i \\ &= g_i + \sum_{k \neq i} g_k \left(f_k^T f_i\right) \\ &= g_i \quad \text{by orthogonality assumption.} \end{aligned}$$

Without any noise, the association is perfect. On the other hand, if the input vectors are not orthogonal, the system will make "mistakes" (Anderson 1977) which are often much more interesting than the perfect association because properties such as abstraction and generalization can be demonstrated (Anderson 1983; Knapp and Anderson 1984).

Brain-state-in-a-box

An extension of this model constructs an associative system which couples a set of neurons, alpha, to itself rather than to a second set. This network is sketched in fig. 2 above. Assuming that every neuron projects to every other neuron, the synaptic weights of these feedback connections can again be defined as the matrix A.

---

insert fig. 2 about here

Under these conditions, an input pattern f is coupled to itself. Since the change in the strength of the synapse a_{ij} is proportional to the product of the activity of the *ith* neuron, f(i), and the jth neuron, f(j),

$$a_{ij} = a_{ji}.$$

If all the synaptic weights are initially zero and their changes are determined by the above scheme, A is a symmetric matrix.

This property of A implies the existence of N mutually orthogonal vectors e_1, ..., e_N such that

$$Ae_i = \lambda_i e_i \quad \text{for } i = 1, \ldots, N,$$

where each lam_i is a real number. The vectors e_1, ..., e_N are called eigenvectors of A, and the scalars lam_1, ..., lam_N are the eigenvalues associated with the corresponding eigenvectors. Since the N eigenvectors are mutually orthogonal, they form a basis set and every N-dimensional vector is a linear combination of the eigenvectors. Consequently, A is completely determined by the set of its eigenvectors and corresponding eigenvalues.

If, as previously assumed, all the synaptic strengths are initially zero, A is determined by its previous inputs. Assuming that K, $K \leq N$, orthonormal inputs are presented p_j times, $j = 1,\ldots,K$, to the system, the synaptic weights are given as

$$A = \sum_{j=1}^{K} p_j f_j f_j^T .$$

In this paper, the set of synaptic connectivities defined by

$$A_j = f_j f_j^T$$

will be defined as the *trace* corresponding to the eigenvector f_j.

In the earlier implementation of this feedback model, it was assumed that the dynamics of the system occur in discrete time. Letting x(t) denote the state vector at a given time t, the succeeding state, x(t+1), is defined as the sum of the activity at time t and the result of the feedback matrix acting on the pattern of activity at time t. That is,

$$x(t+1) = x(t) + A\,x(t).$$

Since all the eigenvalues of A are non-negative, the system is a positive feedback system and the activity is nondecreasing with time. To ensure that the activity does not grow indefinitely large, a nonlinearity (based on the fact that individual cells are limited in their dynamic range) was introduced to contain the activity of the system. This feature can be formalized by assuming that the activity of the system is contained within an N-dimensional hypercube defined by $x_i = \pm C$, $i = 1,\ldots,N$, where x_i is the activity of the *ith* neuron and C is a constant equal to 1. In this system, the vertices or corners are of primary interest. If an eigenvector lies along the diagonal of the hypercube, then both of the corresponding corners will be stable.

Extension

An important limitation of the model is that no mechanism is provided to get the brain state out of the corner it has gotten into. To remedy this we suggest next some assumptions about how brain state vectors might grow and decay. These suggestions will be used in modeling some of the phenomena of multistable perception which cannot be dealt with by earlier versions of the brain-state-in-a-box model.

Iterative scheme. Let us begin by examining the iterative scheme proposed to describe successive patterns of neural activity,

$$x(t+1) = x_{stim} + \sigma x(t) + A(t)\,x(t)$$

where $0 < \sigma < 1$.

Presentation of the stimulus results in a pattern of activity, x_{stim}. The input stimulus is not removed. This pattern of activity can be decomposed and expressed as the linear combination of the eigenvectors of A. In particular, we shall restrict the allowable stimuli to the plane defined by two eigenvectors, f_A and f_B. That is,

$$x_{stim} = \alpha f_A + \beta f_B .$$

Alpha and beta are chosen such that x_{stim} is normal.

The scale factor, sigma, has been introduced to compensate for the magnitude of x(t) relative to the input vector, x_{stim}. Due to the geometry of an N-dimensional hypercube, the magnitude of a vector at one of the corners is proportional to the square root of N. Thus, without any scaling, the contribution of the positive feedback process (i.e. x(t)) can predominate over the input when x(t) is at or near a corner. By scaling the effect of the previous state, x(t), the contributions of the input and the positive feedback become comparable. One important consequence of this scale factor is that the activity of the system does not increase without bound for certain eigenvalues and can even decay to the resting level upon removal of the input.

Initial configuration. A set of initial conditions, x(0) and A(0), is required to begin the iteration. Under most conditions, the initial activity pattern, x(0), will be the zero vector (the resting level) and the initial set of synaptic weights will be the determined by the long-term eigenvalues, lam^{max}, of the eigenvectors of A. These long-term values represent a stable component of the strength of the trace. We assume that the time constant for the modification of this value is slow relative to the dynamic modifications introduced later in this section. Changes in A will not be allowed until the activity of the system reaches a corner. This approximation assumes that since the time spent in corners is generally longer than the time getting there we can neglect transit time. If this assumption is not made, analysis becomes very difficult.

The first few activity patterns are thus

$$x(0) = 0$$

$$x(1) = x_{stim} + \sigma x(0) + A(0)x(0)$$
$$= \alpha f_A + \beta f_B$$

$$x(2) = x_{stim} + \sigma x(1) + A(0)x(1)$$
$$= [1 + (\sigma + \lambda_A)]\alpha f_A + [1 + (\sigma + \lambda_B)]\beta f_B .$$

By induction,

$$x(n) = \sum_{i=0}^{n-1} \left[(\sigma + \lambda_A)^i \alpha f_A + (\sigma + \lambda_B)^i \beta f_B \right].$$

Note that the value of x(n) as n becomes large depends on the sum (sigma+lam). If (sigma+lam) is greater than 1, the component of the corresponding eigenvector grows indefinitely large. If (sigma+lam) is equal to 1, the component is simply (n-1). However, if (sigma+lam) is less than 1, the component approaches 1 / (1 - (sigma+lam)). In other words, it is bounded.

In cases where the components of f_A and f_B are not bounded, as well as some cases where they are, the activity of certain cells in the network might increase beyond their physiological limits. To avoid this situation, we must determine the point where the system reaches a "wall" of the "box," thus limiting the activity of the network. By examining a 2-dimensional system, the linear properties of the network can be studied more closely.

Let the components of the activity vector, x(t), be expressed in terms of the eigenvectors of A, f_A and f_B. Until the state reaches an edge of the box, the components of the activity expressed in terms of the eigenvectors of A are simply

$$\alpha(t) = \sum_{i=0}^{t-1} (\sigma + \lambda_A)^i \alpha$$

$$\beta(t) = \sum_{i=0}^{t-1} (\sigma + \lambda_B)^i \beta .$$

The difference between the components of f_A and f_B of two successive states are

$$\alpha(t+1) - \alpha(t) = \alpha (\sigma + \lambda_A)^t$$

$$\beta(t+1) - \beta(t) = \beta (\sigma + \lambda_B)^t .$$

If the step size is small, the assumption of continuous time can be made and expressions for alpha(t) and beta(t) can be found,

$$\alpha(t) = [\alpha / \ln(\sigma + \lambda_A)](\sigma + \lambda_A)^t$$

$$\beta(t) = [\beta / \ln(\sigma + \lambda_B)](\sigma + \lambda_B)^t .$$

This pair of parametric expressions define a curve in the f_A-f_B plane. In the 2-dimensional system, the edge of the box is reached when the sum of the components of f_A and f_B equal 1,

$$[\alpha / \ln(\sigma + \lambda_A)](\sigma + \lambda_A)^t + [\beta / \ln(\sigma + \lambda_B)](\sigma + \lambda_B)^t = 1 .$$

Once this point is reached, subsequent activity is confined to the edge of the box.

Furthermore, for a given lam_A and lam_B, a family of curves is generated by different values of alpha and beta (components of x_{stim}). One member of this family of curves generates two regions separating those stimuli which end in corner A from those ending in corner B. This parametric curve intersects the edge of the box perpendicularly. Since alpha'(t) and beta'(t) are equal at this point, the activity of the system cannot change and the system will not satiate.

Synaptic modification. Once the activity of the network reaches a corner, the eigenvalues become subject to modification and remain subject to modification as long as the activity remains in a corner. A change in the value of an eigenvalue results in a change in the strength of the *trace* ($f f^T$) associated with the corresponding eigenvector and thus affects each synapse.

If the state vector is in a particular corner, say f_A, the strength of the trace corresponding to that particular state will decrease. That is, the eigenvalue lam_A will

decrease by a fraction of its current value. This property can be formally expressed as

$$\lambda(t+1) = \delta\,\lambda(t), \quad 0 < \delta < 1.$$

where delta is the decay constant. The change in the eigenvalue is given by the following difference equation

$$\lambda(t+1) - \lambda(t) = \delta\,\lambda(t) - \lambda(t).$$

By assuming that the change in lambda is continuous, the following relation is obtained,

$$\lambda'(t) + (1-\delta)\,\lambda(t) = 0.$$

Solving this first order differential equation, lam(t) can be expressed simply as

$$\lambda(t) = \lambda(0)\, e^{-(1-\delta)t},$$

where lam(0) is the value of the eigenvalue at time t=0. The eigenvalue will decay until some minimum value is reached. Arguments for the existence of a minimum are presented in Appendix A.

While the eigenvalue associated with the corner the state vector is in decreases, other eigenvalues increase by a fraction of the difference between their long-term values and their current values. This recovery can be formally expressed as

$$\lambda(t+1) = \lambda(t) + \rho\,(\lambda^{max} - \lambda(t)), \quad 0 < \rho < 1$$

where rho is the recovery constant and lam^{max} is the long-term value of the eigenvalue. Again. the change can be expressed as a difference equation.

$$\lambda(t+1) - \lambda(t) = \rho\,(\lambda^{max} - \lambda(t)).$$

Assuming continuous time, the first order differential equation can be solved to obtain the following expression for the eigenvalue during the recovery phase,

$$\lambda(t) = (\lambda^{min} - \lambda^{max}) e^{-\beta t} + \lambda^{max}$$

With these relationships governing the decay and recovery of the eigenvalues, the values of lam_A and lam_B as a function of time can be plotted. Consider the case when both lam_A and lam_B are initially lam_A^{max} and lam_B^{max}. By the iterative scheme detailed previously, the state vector will reach either corner A or corner B (i.e. the state vector will be f_A or f_B, respectively).

In the example illustrated in fig. 3 below, it is assumed that the state vector reaches corner A first. Thus, lam_A will decrease and lam_B will remain fixed since it is already at its maximum. After a certain period, tau_1, the activity is no longer in a corner as lam_A reaches its minimum. Both eigenvalues remain fixed until the system reaches corner B. (Note that once out of a corner, the state vector cannot return to it without an increase of the eigenvalue associated with that corner.) Once corner B is reached, lam_B decays and lam_A recovers. After decaying for a period tau_2, the state is no longer in corner B as lam_B is now at its minimum. At this point, lam_A no longer increases. This value of lam_A will be one of its "initiation values" as it will be the initial value of lam_A when the state vector again reaches corner A. Note that this first initiation value must be less than lam_A^{max}.

- - - - - - - - - - - - - - - - - - -

insert fig. 3 about here

- - - - - - - - - - - - - - - - - - -

This cycle of decay and recovery of lam_A will continue. Because the initial value of lam_A is smaller than that of the previous cycle, tau_3, the time it takes lam_A to reach its minimum, will be less than tau_1. At this point, the value of lam_B will be less than lam_B^{max}. Thus, tau_4, the length of time before lam_B reaches its minimum for the second time, is less than tau_2. Since the time allowed for lam_A to recover is now shorter than the previous recovery phase, the initiation value will also be smaller than the previous one. Consequently, initiation values of both lam_A and lam_B are strictly decreasing.

To demonstrate that the initiation values approach a limit, it is sufficient to demonstrate that these values are nonincreasing and have a lower bound. It has been shown above that the initiation values are strictly decreasing. Since each eigenvalue must be larger than its minimum, a lower bound exists for lam_A and lam_B. Therefore, asymptotic values for the initiation values of both lam_A and lam_B exist.

Furthermore, it can be shown that the limits are not the minimum values of lam_A and lam_B. (The system would be uninteresting if this were the case.) This will be demonstrated by the following argument. Allow both eigenvalues to decay to their respective minima and then allow both of them to recover. The eigenvalues are plotted as a function of time and the lam_B curve translated so that both minima are vertically aligned as illustrated in fig. 4a. As seen in fig. 4b, the recovery phases are reflected about a vertical line intersecting the eigenvalue minima.

. . . . . .

insert fig. 4 about here

. . . . . .

Consider the case examined previously where the starting values of lam_A and lam_B were their maximum values (points 0 and 0' in fig. 4b) and the first corner reached was A.

While lam_A is decaying, there is no change in lam_B. Since the period of recovery and decay are the same during a given phase, a vertical line through point 0' on the decay curve of lam_B will intersect the recovery curve of lam_A at point 1. By drawing a horizontal line from this point on the recovery curve of lam_A to intersect its decay curve at point 2, the starting point for the modification of lam_A when the system is again in corner A can be found. This cycle is repeated with lam_B. If these vertical and horizontal segments are plotted as illustrated in fig. 5c, it can be seen that a limiting path is approached. This path is bounded by the intersections of the decay and recovery curves for lam_A and lam_B.

A number of other points can be made. First, the same path (horizontal and vertical segments) will be followed if the initial corner was again A, but the initial value of lam_A was less than lam_A^{max}. Second, the state when both eigenvalues are at their respective minima must also be considered. At this point, the system is in unstable equilibrium. It is at equilibrium because the system will remain at this point if it is not perturbed. It is unstable, however, because slight perturbations will drive the system away from this point and toward the same limiting path described above.

It should be noted that not all decay and recovery schemes result in a limiting path. For example, if the decay and recovery occur in constant increments (i.e. linear) with the values bounded below by lam^{min} and above by lam^{max}, the eigenvalues will approach their respective minima if the increment is less than the decrement, and will approach their respective maxima if the increment is greater than or equal to the decrement.

Properties of Multistable Perception

Bias

A number of investigators (Leeper 1935; Fisher 1967) have observed that most bistable stimuli are not equally likely to be perceived in one of the two possible configurations. There is almost always one configuration which is favored over the other. This section examines some possible variables which might lead to this bias and how they might be learned.

To adequately address this issue, however, a measure of bias must initially be agreed on. To date, bias has been empirically determined by assessing (1) the probability that a particular configuration is initially perceived, and (2) the relative duration of two or more configurations over an extended period. These two measures are not independent of each other, and their relationship has been demonstrated by a number of investigators (see Vickers 1972). In this section, the initial configuration will be used as a measure of bias.

Stimulus (alpha and beta). One of the most important factors in determining bias is the stimulus itself. As Fisher (1967) demonstrated using a sequence of figures changing from an unambiguous "man" to an unambiguous "girl," a given ambiguous figure can be regarded as one example from a range of stimuli gradually varying from one unambiguous configuration to another. As Vickers (1972) has indicated, this continuum can be regarded as a decision axis and "the probability of one percept being chosen over another should be a sigmoidal function, resembling the general psychometric type, of the discriminability of the stimulus elements along that dimension which is critical for the emergence of the percept or its alternative" (p. 35). Oyama (1950) and Kuennapas (1957) have shown that a large percentage of the responses given when the stimuli are at one extreme end of the range correspond to the closest unambiguous configuration. While stimuli in the middle of the range are as likely to be perceived in either of the configurations, the closer a stimulus

is toward an end of this range, the more likely it will be perceived as the configuration nearest to it.

In our simulations, we have chosen two Walsh vectors, (1,1.1,1.-1,-1,-1,-1) and (1,-1,1,-1,1,-1,1,-1), as representations of the two unambiguous percepts. Since the dot product of these two vectors is 0, they are mutually orthogonal. The set of stimuli labeled 0-18 ranges from one unambiguous stimulus (0) to the other unambiguous stimulus (18) and lies in the plane defined by the two Walsh functions as illustrated in fig. 5 below. All stimuli are normalized and adjacent stimuli are separated by 5 degrees.

. . . . . .

insert fig. 5 about here

. . . . . .

Kohonen (1977) remarked that "biological memories very probably operate like *adaptive filters*, and their primary purpose seems to be to correct random errors that occur in sensory experiences" (p. 69). One manifestation of this compensatory property is category formation. In forming categories, the prototype of the category can be regarded as the average of the exemplars presented (Knapp and Anderson 1984). Such a process emphasizes the similarities among stimuli, while deemphasizing minor perturbations.

Under the iterative scheme proposed in this paper, a set of simulations run with equal initial eigenvalues showed that stimuli from one half of the range (alpha > beta) always ended in corner A initially, and the other half (beta > alpha) always ended in corner B. When varying amounts of gaussian noise was added to the input, a series of sigmoidal functions were generated. As seen in fig. 6. the greater the amount of noise. the shallower the transition from one category to the other.

. . . . . .

insert fig. 6 about here

- - - - - - - - - - - - - - - - - - - -

Furthermore, the time taken to achieve the initial percept is related to the distance of the stimulus from the "boundary" of the two alternatives along the continuum. Both Shepard and Humphris (unpublished experiment cited in Vickers 1972) and Shepard (1972) have shown that stimuli near the far extremes (unambiguous configuration) have short latencies, and stimuli near the boundary have longer latencies. By counting the number of iterations it takes the activity to reach a stable corner, a measure of latency can be obtained. As illustrated in fig. 7, under the conditions of the above set of simulations (no added noise), the number of steps to a corner increases as one approaches the boundary and is symmetric about this point.

- - - - - - - - - - - - - - - - - - - -

insert fig. 7 about here

- - - - - - - - - - - - - - - - - - - -

System (eigenvalues). Another consideration in determining bias is the role of experience, in both the distant and recent past. Allport (1955) has proposed that perceptual sets are primarily determined by past experience and that the greater the frequency of such experiences, the stronger the set. On the other hand, Epstein and Rock (1960) concluded that in the perception of an ambiguous figure, recency was more important than frequency (or expectation). They argue that a distinction should be made between total past experience and factors in the immediate situation. In this section, the role that total past experience plays will be examined, with the effects of recent experience discussed in later sections on adaptation and hysteresis.

If we regard the neural network as an adaptive filter, we might expect that repeated presentations of certain stimuli during learning trials[1] ultimately affects the response of the filter. After learning for a while, the eigenvectors of A with large eigenvalues will tend to correspond to commonly presented patterns, and the eigenvalue, in a more complex system, will be an estimate of the frequency of presentation.

In the simulations presented here, the long-term values of lam_A and lam_B are assumed to be generated by the learning scheme presented in Anderson et al. (1977). The following results using a normalized stimulus vector demonstrate the significance of these values, or more accurately, the *ratio* of these values in determining the boundary. The sum of lam_A and lam_B was kept constant ($lam_A + lam_B = 0.8$) while their ratio was varied. As seen in fig. 8, as the ratio of lam_A/lam_B increases, the boundary shifts toward f_B. In essence, as the relative familiarity (eigenvalue) of a particular configuration (eigenvector) increases, that stimulus which had been perceived in each of the two configurations with perhaps equal probability is now perceived as the more familiar configuration a greater percentage of the time.

- - - - - - - - - - - - - - - - - - - -

insert fig. 8 about here

- - - - - - - - - - - - - - - - - - - -

Although there may not be, to our knowledge, any studies quantitatively examining the influence of frequency on multistable perception, there does exist a fair amount of qualitative evidence showing that past experience does have an effect (Rubin 1921; Leeper

---

[1] It is important to distinguish the learning phase from test phase. The long-term eigenvalue will be allowed to change only during the learning phase. Although this simplification is incorrect (there is some evidence of learning during the experiments), it is made here for ease in simulation and because it seems not to be a major effect.

1935). Rubin (1921) presented ambiguous figure-ground patterns to subjects and instructed them to see the stimulus in a certain manner. When these patterns were presented later, subjects tended to perceive them in the manner organized previously.

In terms of the neural network model, the trace of an unambiguous pattern f_i can be regarded as the synaptic connectivities defined by $f_i\ f_i^T$, and the strength of this trace as lam_i. If the stimulus presented is somewhere along the continuum between the two unambiguous patterns, the stimulus can be regarded as a weighted sum of these two patterns. Since both traces are represented within the system, *both* will have a positive feedback effect on states immediately following the onset of the stimulus. The initial percept is thus dependent on the relative bias of the stimulus *and* the relative values of the eigenvalues corresponding to the two unambiguous configurations.

Adaptation

By presenting the subject with an unambiguous version of a multistable stimulus, it is possible to selectively adapt that particular configuration. Howard (1961) and Virsu (1977) have examined this effect using reversible perspective figures. Under these conditions, subjects tend to perceive the alternative configuration initially.

This property has also been simulated under the iterative scheme proposed. In the adapting condition (results listed in table 2 of Appendix B), an unambiguous stimulus with components of activity in only one direction (e.g. f_A) is presented. After a period of time (75 iterations), the stimulus is removed. Upon removal of the stimulus, the system decays to its resting level and the eigenvalue lam_A begins to recover. (Adaptation can also be demonstrated if the eigenvalue is not allowed to recover.) Some time after the adapting stimulus has been removed (iteration #83), a stimulus which was biased toward the now adapted configuration is presented. To simulate this procedure, the state of the system at that point of the adaptation simulation (table 2) is used as the initial conditions of the

system.

The results of simulations using the identical stimulus presented to the unadapted and the adapted system are given in tables 1 and 3, respectively, in Appendix B. A number of differences are immediately apparent. First, the initial percept is different. Since the stimulus chosen was biased towards the f_A configuration, this configuration was initially perceived under the unadapted condition. Adapting the system with an unambiguous stimulus corresponding to configuration A resulted in reducing only the eigenvalue corresponding to this eigenvector. Thus, presentation of the same stimulus under the new conditions resulted in the alternative configuration as the initial percept. Another point to note is that the period of time the state is in corner A (the adapted configuration) increases rather than decreases.

Hysteresis

One observation which has been the focus of a number of models (Haken 1979; and Stewart 1978) is the hysteresis manifested upon viewing a sequence of stimuli. Hysteresis in multistable perception refers specifically to the differential response to a stimulus in a sequence of stimuli varying along a continuum as a function of the order of presentation.

Actually, this property has been reported earlier under the term of perceptual set (Wertheimer 1923) or recency (Epstein and Rock 1960). Wertheimer describes a situation in which an observer is first shown a pattern of dots as illustrated below

* * * * * *

a b a b a b

By the gestalt law of proximity, 3 pairs of dots are seen, with "a" the left member of the pair (designated a-b). In successive exposures, the distance between the members of this pair increases, decreasing the distance between members of adjacent pairs of dots. On

viewing this sequence, an observer will perceive the stimulus below

* * * * * *

a b a b a b

as a-b pairs despite the fact that the dots are equidistant from each other. Only if the separation were increased further would the stimulus be seen as a b-a pair. Conversely, if the stimulus configuration biased toward the b-a pair was the initial stimulus with the sequence of stimuli presented in reverse order, the equidistant stimulus above would now be seen as a b-a pair. Thus, the same stimulus gives rise to 2 different percepts depending on the order of presentation.

Attneave (1971) has also explored the consequences of manipulating the order of presentation of a sequence of stimuli. Using a set ranging from the unambiguous "man" configuration to the unambiguous "girl" configuration, Attneave found that a stimulus biased toward one configuration (e.g. the "man") would be perceived as the other configuration (i.e. the "girl") if the sequence began with the unambiguous "girl" configuration.

Simulation of this property of multistable perception is similar to the procedure used in demonstrating adaptation. In both cases, the initial conditions of the state (x(0)) and the system (A(0)) are influenced by the recent history. However, for adaptation, the duration of the stimulus presentation is much longer than for hysteresis.

First of all, recall that each stable state is an eigenvector of the feedback matrix and that the eigenvectors form a basis set. If the input stimulus, x_{stim}, is removed after the system is in a stable state, there will be no component of activity in any direction other than the stable corner. Thus, subsequent states will be confined to changes in the magnitude of the activity vector, with the direction remaining constant. If the long-term eigenvalue associated with the eigenvector is less than 1-sigma, the magnitude of the

activity vector will never be greater than one upon removal of the stimulus,

$$x(t+1) = \sigma f_A + A(t) f_A$$
$$= (\sigma + \lambda_A(t)) f_A$$
$$< f_A .$$

Thus, a stable state can no longer be attained and one would expect that the system would eventually return to its resting level.

If, however, the stimulus is presented briefly, the eigenvalues will not decay very much. Now, the activity of the system has a component of activity in the direction of the previous configuration at the onset of the next stimulus presentation. This component may be sufficiently large to offset both the bias in the stimulus and the bias created by a decrease in the eigenvalue associated with the previous configuration. A sequence simulating hysteresis in presented in Appendix C.

Dynamics

Perhaps more dramatic than the fact that a stimulus can be perceived in two or more configurations is the rapid (seemingly instantaneous) transition from one percept to another. To the extent that the ratio of stable state duration to transition state duration is large, the transition can be regarded as instantaneous. In these simulations, this ratio is sometimes as large as ten to one. An interesting model for operation of a neural system would be a movement of the state vector from stable state to stable state with abrupt transitions between them. This would be a way of discretizing or segmenting a continuous stream of sensory inputs in a way determined by the learning which took place in the past of the system. Presumably, significant states would be maintained for a while, perhaps even beyond their actual duration in the stimulus, so they could drive events and analysis in later parts of the nervous system.

In a dynamic system, the time course of the states the system is in is of particular interest. The iterative scheme described above generates such a plot for the multistable perceptual process. Empirical results (Brown 1955; Cohen 1959) indicate that the time course is marked by a monotonic increase in the rate of reversals during the presentation of the stimulus. The slope of the function (i.e. the rate at which the reversal rate increases) gradually decreases to zero as it reaches an asymptotic level about 3 minutes from stimulus onset.

This property can also be shown by plotting the durations of each phase. As Price (1967) has shown, the mean duration of the initial percept (P1) decreases markedly from its initial duration, with no further decrease in subsequent phases. The mean duration of the reversal phase (P2) is constant through the first 10 cycles of reversals. It was also noted that the asymptotic mean duration of P1 was approximately equal to the constant mean duration of P2.

Our simulations show similar qualitative behavior as illustrated in fig. 9. The duration of the initial percept is longer than the duration of subsequent percepts. Furthermore, the durations corresponding to each of the two possible configurations decrease until an asymptotic level is reached. This is predicted from the existence of eigenvalue minima described in Appendix A.

. . . . . .

insert fig. 9 about here

. . . . . .

As previously noted, the "average of the times during which a particular percept is seen is a direct function of the probability that that percept will occur first" (Vickers 1972). As seen in fig. 10, the duration of the initial percept is a function of the ambiguity

of the stimulus. The more biased a stimulus is toward a particular unambiguous configuration, the longer the initial duration. This plot is qualitatively similar to results of Oyama (1950) and Kuennapas (1957) which are empirical results showing the probability of initial perception.

. . . . . .

insert fig. 10 about here

. . . . . .

Discussion

The brain-state-in-a-box model is a neural network based on modifications at the synaptic level. Although such neurophysiological parameters as synaptic strength are not easily accessible, it is possible that behavior shown by human observers as a result of events at the synaptic level are observable.

In this paper, we have attempted to extend the range of this model by demonstrating that the major properties of multistable perception can be modelled with only a few additional assumptions. We have shown that the configuration initially perceived is a function of both the stimulus and the initial condition of the system. Furthermore, the model can simulate the empirical observation that stimuli very similar to an unambiguous configuration are perceived shortly after stimulus onset whereas stimuli which are more ambiguous require a longer interval. We have also demonstrated that adaptation and hysteresis depend on the recent history of the system by examining the effects of the initial conditions. In addition, we have modelled a number of the dynamic properties of multistable perception.

It is interesting to note that a number of investigators writing about language comprehension (Foss and Hakes 1978; Miller 1981) have proposed that ambiguous sentences are analogous to multistable stimuli. Ambiguities can be due to a word having different meanings as in "Tommy found a *bat* in the attic" (lexical ambiguity), a sentence associated with two different surface structures as in "They are broiling hens" (surface structure ambiguity), or a sentence associated with two different underlying structures as in "The chicken is ready to eat" (underlying structure ambiguity).

As an example of how the model could be extended to this domain, let us consider the problem of lexical ambiguity. Ambiguous words have either the same spelling

(homographs), the same pronunciation (homophones), or both the same spelling and the same pronunciation (homonyms). Given only the spelling of a homograph or the pronunciation of a homophone, only one meaning is eventually accessed, however. Which meaning is accessed is a function of both context and frequency (Simpson 1981).

Using the framework developed here, we can begin to see how the model might account for these findings. A word can be represented as a high-dimensional vector composed of concatenated subvectors representing the graphemic, phonemic, syntactic, and semantic fields of a lexical entry. Therefore, homographs would have identical graphemic fields, homophones identical phonemic fields, and homonyms identical graphemic and phonemic fields. Because a pair of homographs (homophones, homonyms) can have different syntactic and semantic fields, each would be represented as distinct corners. To simulate relative frequency of usage, each word in the lexicon is presented at a given frequency during the learning phase.

After the system has been taught using a number of learning trials, test stimuli are presented. These test stimuli would consist of just the graphemic or the phonemic field of previously learned homographs and homophones, respectively. According to the conventions presented in this study, an unbiased stimulus has components of activity (of the appropriate sign) only along those dimensions which are common to both corners, and no activity along those dimensions in which they differ. Thus, this set of test stimuli can be regarded as unbiased stimuli and we would expect that which meaning gets accessed would be determined by which homograph (homophone) is presented more frequently. However, just as the components of both alternative configurations of an ambiguous stimulus increase until the activity of the system reaches a wall, we would expect *both* meanings of a homograph pair to be activated initially as demonstrated by Swinney (1982).

On the other hand, if the stimuli are equally familiar, the bias provided by those dimensions along which they differ in the two configurations should be sufficient to bias the percept. This bias could be provided by the context of the word and could be implemented as partial activity in the semantic or syntactic fields. Indeed, Swinney (1982) has demonstrated that semantic and syntactic cues do disambiguate homographs.

Moreover, we have shown that a bias created by differential familiarity can be offset by the bias created by different contexts. As Simpson (1981) has shown, which meaning is accessed is a function of both frequency *and* context. Thus, we believe that the approach taken here to study multistable perception can easily be extended to other domains such as comprehension of ambiguous words and sentences.

Financial support for some of this work was provided by a grant from the National Science Foundation to J. A., administered by the Memory and Cognitive Processes section (Grant BNS-82-14728) and by the United States Office of Naval Research (Contract N00014-81-K-0136) to the Center for Neural Science, Brown University. We would like to thank the Center for Cognitive Sciences, Brown University, for computing facilities used in our simulations. We would also like to thank Steve Lehmkuhle and Bill Warren for commenting on an earlier draft of this manuscript.

References

Allport, F.H., 1955. Theories of perception and the concept of structure. New York: Wiley.

Anderson, J.A., 1968. A memory storage model utilizing spatial correlation functions. Kybernetik 5, 113-119.

Anderson, J.A., 1973. A theory for the recognition of items from short memorized lists. Psychological Review 80, 417-438.

Anderson, J.A., J.W. Silverstein, S.A. Ritz, and R.S. Jones, 1977. Distinctive features, categorical perception, and probability learning: Some applications of a neural model. Psychological Review 84, 413-451.

Anderson, J.A., in press. Neural models for cognition. I.E.E.E.: System, Man, and Cybernetics.

Attneave, F., 1971. Multistability in perception. Scientific American 225, 62-71.

Brown, K.T., 1955. Rate of apparent change in a dynamic ambiguous figure as a function of observation time. American Journal of Psychology 68, 358-371.

Cohen, L., 1959. Rate of apparent change of a Necker cube as a function of prior stimulation. American Journal of Psychology 72, 327-344.

Cooper, L., 1974. 'A possible organization of animal memory and learning'. In: B. Lundquist & S. Lundquist (eds.), Proceedings of the Nobel Symposium on Collective Properties of Physical Systems. New York: Academic Press.

Epstein, W., and I. Rock, 1960. Perceptual set as an artifact of recency. American Journal of Psychology 73, 214-228.

Fisher, G.H., 1967. Measuring ambiguity. American Journal of Psychology 80, 541-547.

Foss, D.J. and D.T. Hakes, 1978. Psycholinguistics. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Haken, H., 1979. 'Pattern formation and pattern recognition - An attempt at synthesis'. In: H. Haken (ed.), Pattern formation by dynamic systems and pattern recognition: Proceedings of the International Symposium on Synergetics at Schloss Elmau, Bavaria. New York: Springer Verlag.

Howard, I.P., 1961. An investigation of a satiation process in the reversible perspective of revolving skeletal shapes. Quarterly Journal of Experimental Psychology 13, 19-33.

Julesz, B., 1971. Foundations of cyclopean perception. Chicago: University of Chicago Press.

Knapp, A.G. and J.A. Anderson, submitted for publication. A theory of categorization based on distributed memory storage. Journal of Experimental Psychology: Human Perception and Performance.

Koehler, W., 1940. Dynamics in Psychology. Liveright Publishing Corp..

Koehler, W. and H. Wallach, 1944. Figural after-effects: An investigation of visual processes. Proceedings of the American Philosophical Society 88, 269-357.

Kohonen, T., 1977. Associative Memory: A System Theoretic Approach. Berlin: Springer-Verlag.

Kuennapas, T.M., 1957. Experiments on figural dominance. Journal of Experimental Psychology 53, 31-39.

Leeper, R.W., 1935. A study of a neglected portion of the field of learning - The development of sensory organization. Journal of Genetic Psychology 46, 41-75.

McDougall, W., 1904. Physiological factors of the attention process: IV. Mind 15, 329-359.

Miller, G.A., 1981. Language and Speech. San Francisco: W. H. Freeman.

Necker, L.A., 1832. Observations on some remarkable phenomena seen in Switzerland; and an optical phenomenon which occurs on viewing of a crystal or geometrical solid. Philosophical Magazine, 3 ser. 1, 329-343.

Orbach, J., D. Ehrlich, and H.A. Heath, 1963. Reversibility of the Necker cube: I. An examination of the concept of "satiation of orientation." Perceptual and Motor Skills 17, 439-458.

Oyama, T., 1950. Figure-ground dominance as a function of sector angle, brightness, hue, and orientation. Journal of Experimental Psychology 60, 299-305.

Poston, T. and I. Stewart, 1978. Nonlinear modeling of multistable perception. Behavioral Science 23, 318-334.

Price, J.R., 1967a. Two components of reversal for a rotating skeletal cube: "Conditioned satiation." Australian Journal of Psychology 19, 261-270.

Price, J.R., 1967b. Perspective duration of a plane reversible figure. Psychonomic Science 9, 623-624.

Price, J.R., 1969. Studies of reversible perspective: A methodological review. Behaviour Research Methods and Instrumentation 1, 102-106.

Pritchard, R.M., W. Heron, and D.O. Hebb, 1960. Visual perception approached by the method of stabilized images. Canadian Journal of Psychology 14, 67-77.

Riggs, L.A., F. Ratliff, J.C. Cornsweet, and T.N. Cornsweet, 1953. The disappearance of steadily fixated visual test objects. Journal of the Optical Society of America 43, 495-501.

Rubin, E., 1921. Visuell wahrgenommene Figuren. Copenhagen: Glydendalske.

Sadler, T.G. and R.B. Mefferd, Jr., 1970. Fluctuations of perceptual organization and orientation: Stochastic (random) or steady state (satiation)? Perceptual and Motor Skills 31, 739-749.

Selfridge, O.G., 1959. Pandemonium: A paradigm for learning. In: The mechanisation of thought processes. London: H. M. Stationery Office.

Shepard, M., 1972. 'Decision processed in perceptual organization: effects of proximity and regularity on response frequency and latency'. In: F. F. O'Callaghan (ed.). Pictorial Organization and Shape, Division of Computing Research, C.S.I.R.O. Canberra, A.C.T..

Simpson, G.B., 1981. Meaning dominance and semantic context in the processing of lexical ambiguity. Journal of Verbal Learning and Behavior 20, 120-136.

Spitz, H.H., 1963. Some experiments with a stationary and revolving three-dimensional skeletal cube. Paper presented at the meeting of the Eastern Psychological Association, New York City, April, 1963.

Swinney, D.A., 1982. 'The structure and time-course of information interaction during speech comprehension: lexical segmentation, access, and interpretation'. In: J. Mehler, E.C.T. Walker, and M. Garrett (eds.). Perspectives on Mental Representation: Experimental and Theoretical Studies of Cognitive Processes and Capacities. Hillsdale, NJ: Erlbaum.

Taylor, M.M. and K.D. Aldridge, 1974. Stochastic processes in reversing figure perception. Perception and Psychophysics 16, 9-27.

Vickers, D., 1972. A cyclic decision model of perceptual alternation. Perception 1, 31-48.

Virsu, V., 1975. Determination of perspective reversals. Science 257, 736-737.

Wertheimer, M., 1923. Untersuchengen zur Lehre von der Gestalt: II. Psycholgische Forschung 4, 301-350.

Appendix A

Consider the situation where the state of the system is f_A. That is,

$$x(t) = f_A .$$

According to the iterative scheme, the next state, $x(t+1)$, is

$$\begin{aligned} x(t+1) &= x_{stim} + \tau f_A + \lambda_A(t) f_A \\ &= (\alpha f_A + \beta f_B) + \tau f_A + \lambda_A(t) f_A \\ &= [\alpha + (\tau + \lambda_A(t))] f_A + \beta f_B . \end{aligned}$$

Note that once the state vector is in a corner, the subsequent state is independent of all other eigenvalues. In particular, it is independent of lam_B.

Even after restricting the activity of the state vector to within the boundaries of the hypercube, the subsequent state may again be f_A. Thus, lam_A will be decremented according to the above scheme. This cycle, however, cannot continue indefinitely. At some point, the component of f_A will be less than 1. Since alpha and sigma are constants, the value of lam_A will not decrease below a certain a value,

$$\lambda_A(t) \geq 1 - (\alpha + \tau) .$$

The minimum value of an eigenvalue will probably be larger than the value indicated above. This can easily be seen by examining the 2-dimensional system illustrated in fig. 11. The activity of cell 1 is plotted on the x-axis and the activity of cell 2 is plotted on the y-axis. The two eigenvectors of A are f_A, (-1,1), and f_B, (1,1) and these two orthogonal vectors point to two corners of the box. Due to the constraint on the maximum firing rate

of a cell, a state vector extending into either of the regions, ($x \leq -1, y \geq 1$) and ($x \geq 1, y \geq 1$), will be in a corner. If the state vector is outside the shaded regions, the state will not be in a corner. Thus, the minimum values of lam_A and lam_B, designated lam_A^{min} and lam_B^{min}, respectively, are

$$\lambda_A^{min} = 1 - (\alpha + \gamma) + \beta$$
$$= 1 - \gamma - (\alpha - \beta) \quad \text{and}$$

$$\lambda_B^{min} = 1 - (\beta + \gamma) + \alpha$$
$$= 1 - \gamma + (\alpha - \beta).$$

Although the existence of minima has not yet been formally established for the high dimensional space, our simulations do indicate that such minima do exist.

---

insert fig. 11 about here

Appendix B

The following tables are segments of actual simulations with the columns representing the following values:

1	iteration number
2	number of iterations state vector is in corner A
3	number of iterations state vector is in corner B
4	current value of lam_A
5	current value of lam_B
6-13	x(1),...,x(8), the state of the system

In these simulations, sigma is equal to 0.5, the decay and recovery constants are 0.05 and 0.015, respectively, and lam_A^{max} and lam_B^{max} are 0.8 and 0.4, respectively. In the unadapted condition, the initial values of lam_A and lam_B are their respective maximum values and the initial activity is the zero vector. As seen in table 1, presentation of stimulus 13 leads to configuration A as the initial percept in the unadapted conditions. To simulate adaptation, an unambiguous stimulus (stimulus 0) is presented to the system and the subsequent activity is shown in table 2. After 75 iterations, the stimulus is removed. The eigenvalue, lam_A begins to recover and the system decays to its resting level of activity. Eight iterations from the offset of stimulus 0, a stimulus (stimulus 13) which had been initially perceived as configuration A in the unadapted condition is presented to the system following adaptation. The initial value of lam_A and the activity of the system are obtained from iteration 83 in table 2. Under these conditions, the initial percept is configuration B.

table 1: stimulus 13 in unadapted condition

it#	A	B	lamA	lamB	x(1)	x(2)	x(3)	x(4)	x(5)	x(6)	x(7)	x(8)
0	0	0	0.80	0.40	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0	0	0.80	0.40	0.5	-0.2	0.5	-0.2	0.2	-0.5	0.2	-0.5
2	0	0	0.80	0.40	1.0	-0.3	1.0	-0.3	0.3	-1.0	0.3	-1.0
3	0	0	0.80	0.40	1.0	-0.3	1.0	-0.3	0.3	-1.0	0.3	-1.0
4	0	0	0.80	0.40	1.0	-0.1	1.0	-0.1	0.1	-1.0	0.1	-1.0
5	0	0	0.80	0.40	1.0	0.7	1.0	0.7	-0.7	-1.0	-0.7	-1.0
6	0	0	0.80	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
7	1	0	0.76	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
8	2	0	0.72	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
9	3	0	0.69	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
10	4	0	0.65	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0

table 2: adapting with stimulus 0

it#	A	B	lamA	lamB	x(1)	x(2)	x(3)	x(4)	x(5)	x(6)	x(7)	x(8)
0	0	0	0.80	0.40	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0	0	0.80	0.40	0.4	0.4	0.4	0.4	-0.4	-0.4	-0.4	-0.4
2	0	0	0.80	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
3	1	0	0.76	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
4	2	0	0.72	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
5	3	0	0.69	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
*	*		*									
*	*		*									
*	*		*									
73	71	0	0.02	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
74	72	0	0.02	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
75	73	0	0.02	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
76	74	0	0.02	0.40	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
77	74	0	0.03	0.40	0.6	0.6	0.6	0.6	-0.6	-0.6	-0.6	-0.6
78	74	0	0.04	0.40	0.4	0.4	0.4	0.4	-0.4	-0.4	-0.4	-0.4
79	74	0	0.05	0.40	0.3	0.3	0.3	0.3	-0.3	-0.3	-0.3	-0.3
80	74	0	0.06	0.40	0.2	0.2	0.2	0.2	-0.2	-0.2	-0.2	-0.2
81	74	0	0.07	0.40	0.1	0.1	0.1	0.1	-0.1	-0.1	-0.1	-0.1
82	74	0	0.09	0.40	0.1	0.1	0.1	0.1	-0.1	-0.1	-0.1	-0.1
83	74	0	0.10	0.40	0.1	0.0	0.1	0.0	0.0	-0.1	0.0	-0.1
84	74	0	0.11	0.40	0.1	0.0	0.1	0.0	0.0	-0.1	0.0	-0.1

table 3: stimulus 13 in adapted condition

it#	A	B	lamA	lamB	x(1)	x(2)	x(3)	x(4)	x(5)	x(6)	x(7)	x(8)
0	0	0	0.10	0.40	0.1	0.0	0.1	0.0	0.0	-0.1	0.0	-0.1
1	0	0	0.10	0.40	0.5	-0.1	0.5	-0.1	0.1	-0.5	0.1	-0.5
2	0	0	0.10	0.40	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
3	0	1	0.11	0.38	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
4	0	2	0.12	0.36	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
5	0	3	0.13	0.34	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
6	0	4	0.14	0.33	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
7	0	5	0.15	0.31	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
8	0	6	0.16	0.29	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
9	0	7	0.17	0.28	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
10	0	8	0.18	0.27	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0

Appendix C

The following table is composed of a series of simulations using a sequence of stimuli varying from an unambiguous stimulus to one biased toward the alternative configuration. The stimulus is removed after 3 iterations, and the values of x, lam_A, and lam_B at that point are used as the initial values on the subsequent trial. The first stimulus is stimulus 0, followed by stimulus 1 through stimulus 13. The initial values of lam_A and lam_B are both 0.5. As the simulations show, stimuli biased toward configuration B are perceived as configuration A if the set of stimuli are presented in the above order.

table 4: simulation of hysteresis

stim	it#	A	B	lamA	lamB	x(1)	x(2)	x(3)	x(4)	x(5)	x(6)	x(7)	x(8)
0:	0	0	0	0.50	0.50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	1	0	0	0.50	0.50	0.4	0.4	0.4	0.4	-0.4	-0.4	-0.4	-0.4
	2	0	0	0.50	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	1	0	0.47	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
1:	0	0	0	0.47	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.45	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.43	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.41	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
2:	0	0	0	0.41	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.39	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.37	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.35	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
3:	0	0	0	0.35	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.33	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.32	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.30	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
4:	0	0	0	0.30	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.28	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.27	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.26	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
5:	0	0	0	0.26	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.24	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.23	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.22	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
6:	0	0	0	0.22	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.21	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.20	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.19	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
7:	0	0	0	0.19	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.18	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.17	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.16	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
8:	0	0	0	0.16	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.15	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.15	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.14	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
9:	0	0	0	0.14	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.13	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.13	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.12	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
10:	0	0	0	0.12	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.11	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.11	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.10	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
11:	0	0	0	0.10	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.10	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.09	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	3	0	0.09	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
12:	0	0	0	0.08	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	1	1	0	0.08	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	2	2	0	0.07	0.50	1.0	1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0
	3	2	0	0.07	0.50	1.0	0.9	1.0	0.9	-0.9	-1.0	-0.9	-1.0
13:	0	0	0	0.07	0.50	1.0	0.6	1.0	0.6	-0.6	-1.0	-0.6	-1.0
	1	0	0	0.07	0.50	1.0	-0.1	1.0	-0.1	0.1	-1.0	0.1	-1.0
	2	0	0	0.07	0.50	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0
	3	0	1	0.08	0.47	1.0	-1.0	1.0	-1.0	1.0	-1.0	1.0	-1.0

Figure captions

figure 1. A neural network connecting two sets of neurons, alpha and beta, with the set of synaptic connectivities defined as the matrix A (from Anderson et al. 1977).

figure 2. A neural network feeding back onto itself (from Anderson et al. 1977).

figure 3. Time course of the decay and recovery of eigenvalues lamA and lamB.

figure 4. (a) Plot of lamA and lamB as they are allowed to decay to their respective minima and allowed to recover, (b) The recovery phases are reflected about the vertical lines through the minimum values of lamA and lamB. The reflected plot of lamB is translated horizontally so both minima are vertically aligned. (c) The initiation value of lamA is determined by the vertical line through the recovery curve of lamB with the decay curve of lamA. The value of lamB is found similarly. A graphic demonstration of the asymptotic nature of the initiation values is given by plotting the horizontal and vertical connecting line segments.

figure 5. The set of stimuli used ranges from one unambiguous stimulus (stimulus 0) to the second unambiguous stimulus (stimulus 18). Adjacent stimuli are separated by 5 degrees and all stimuli are normalized.

figure 6. Percentage of initial percept as a function of the stimulus. Each point represents 100 trials with different amounts of gaussian distributed noise added to each dimension off the stimulus vector.

figure 7. Simulation of the time required for perception of the initial configuration. Stimuli similar to unambiguous configurations are perceived shortly after the onset of the stimulus, and stimuli near the boundary require longer

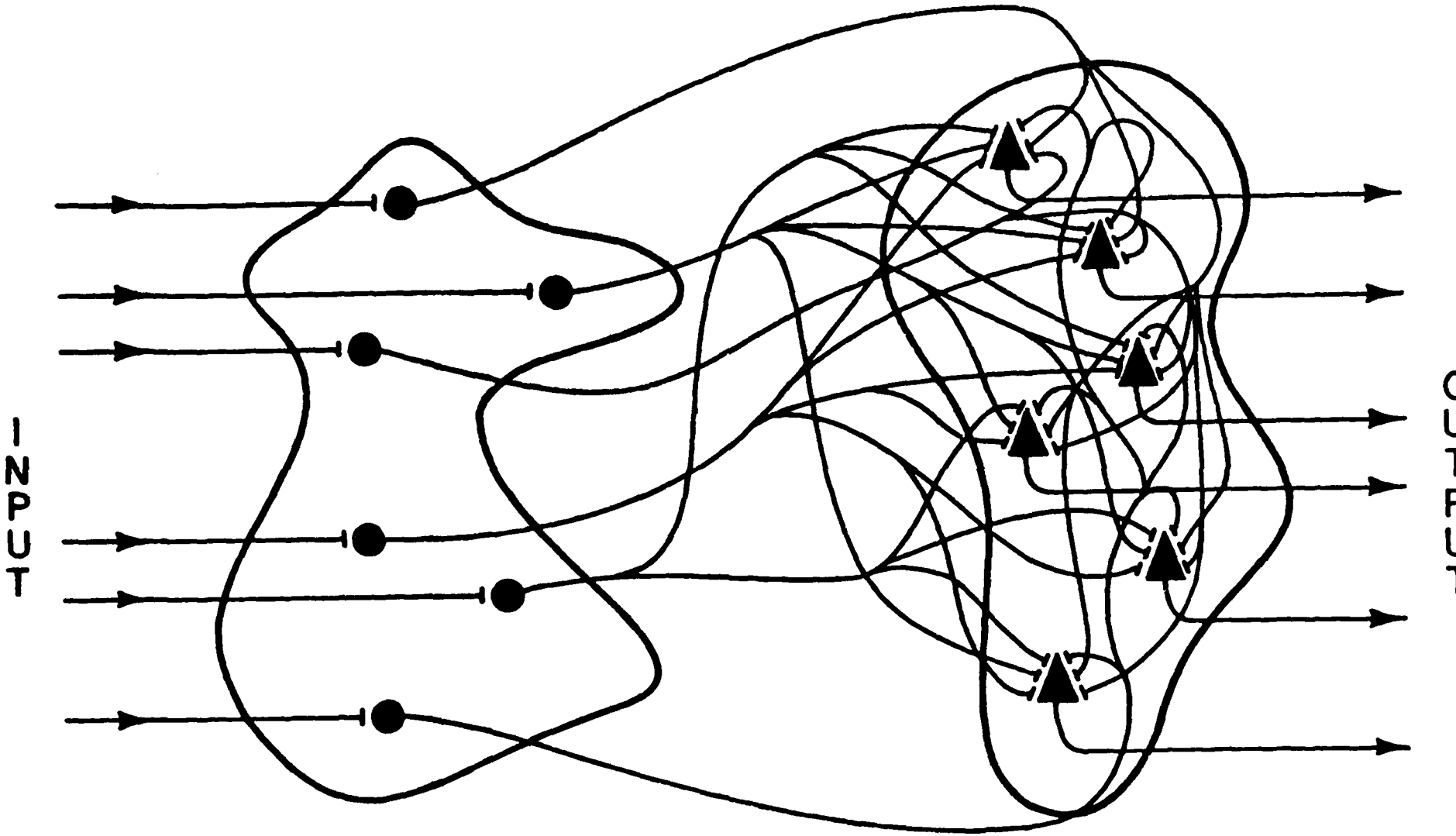
periods of time.

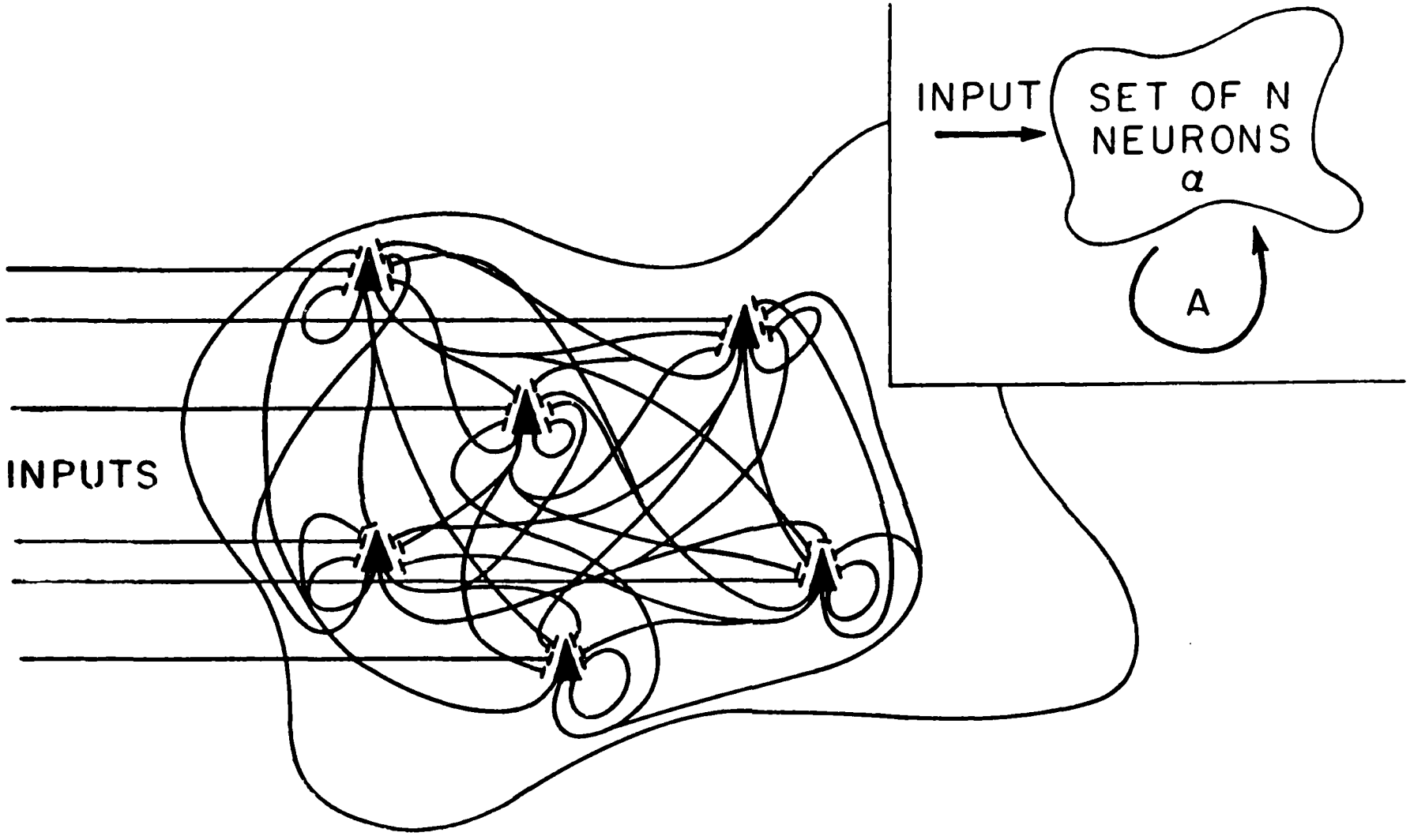
figure 8. Effect of varying the ratio of the initial eigenvalues on the boundary between configurations.

figure 9. The time course of multistable perception. Simulation under dynamic decay and recovery of the eigenvalues.

figure 10. Simulation of the duration of the initial percept as a function of the particular stimulus.

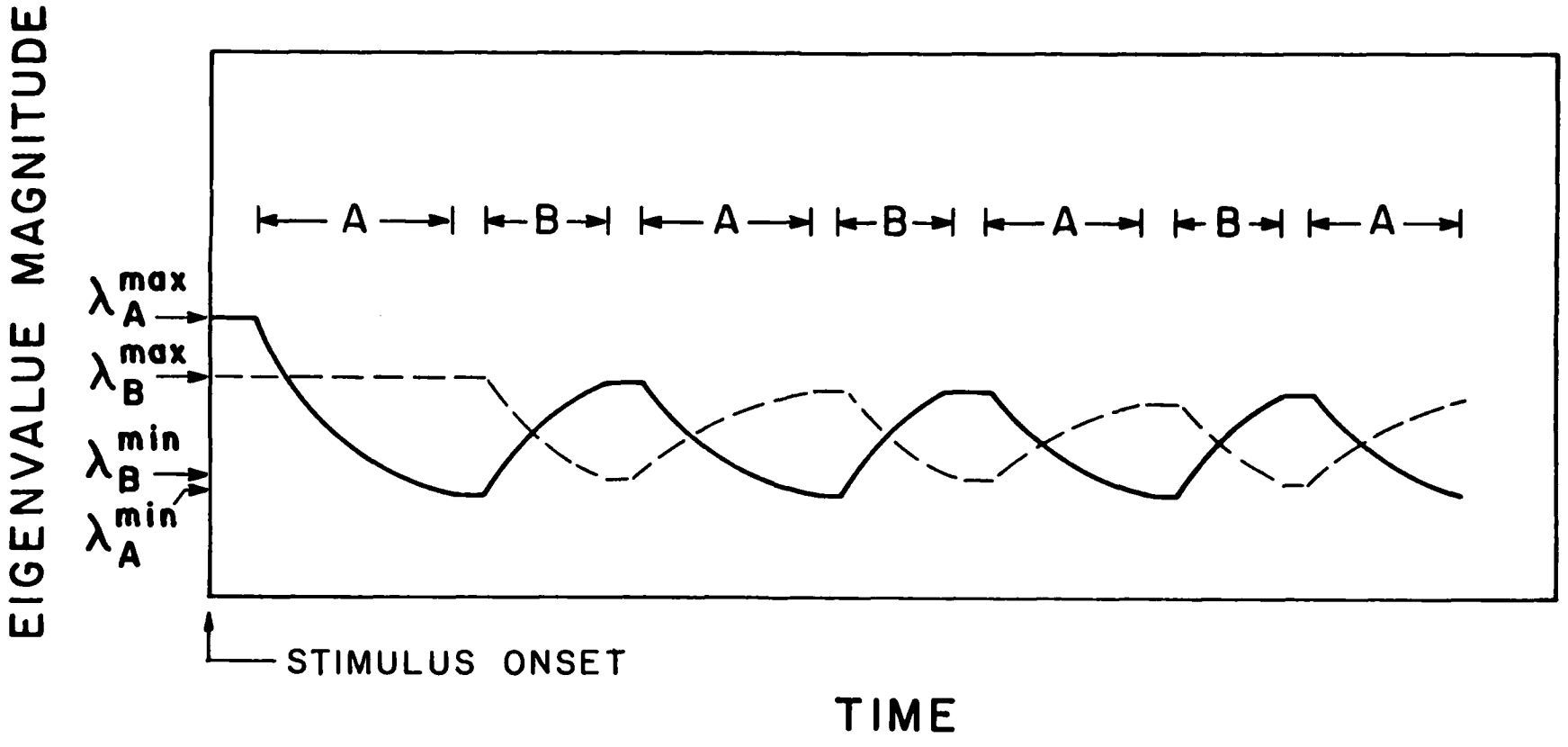
figure 11. A 2-dimensional system with the response of cell 1 plotted on the horizontal axis, and the response of cell 2 plotted on the vertical axis. The firing rate of both cells is bounded below by -1 and above by 1. If a state is in the corner, the subsequent state will also be in a corner if the output, prior to the constraint on the firing rate, is in the shaded region.

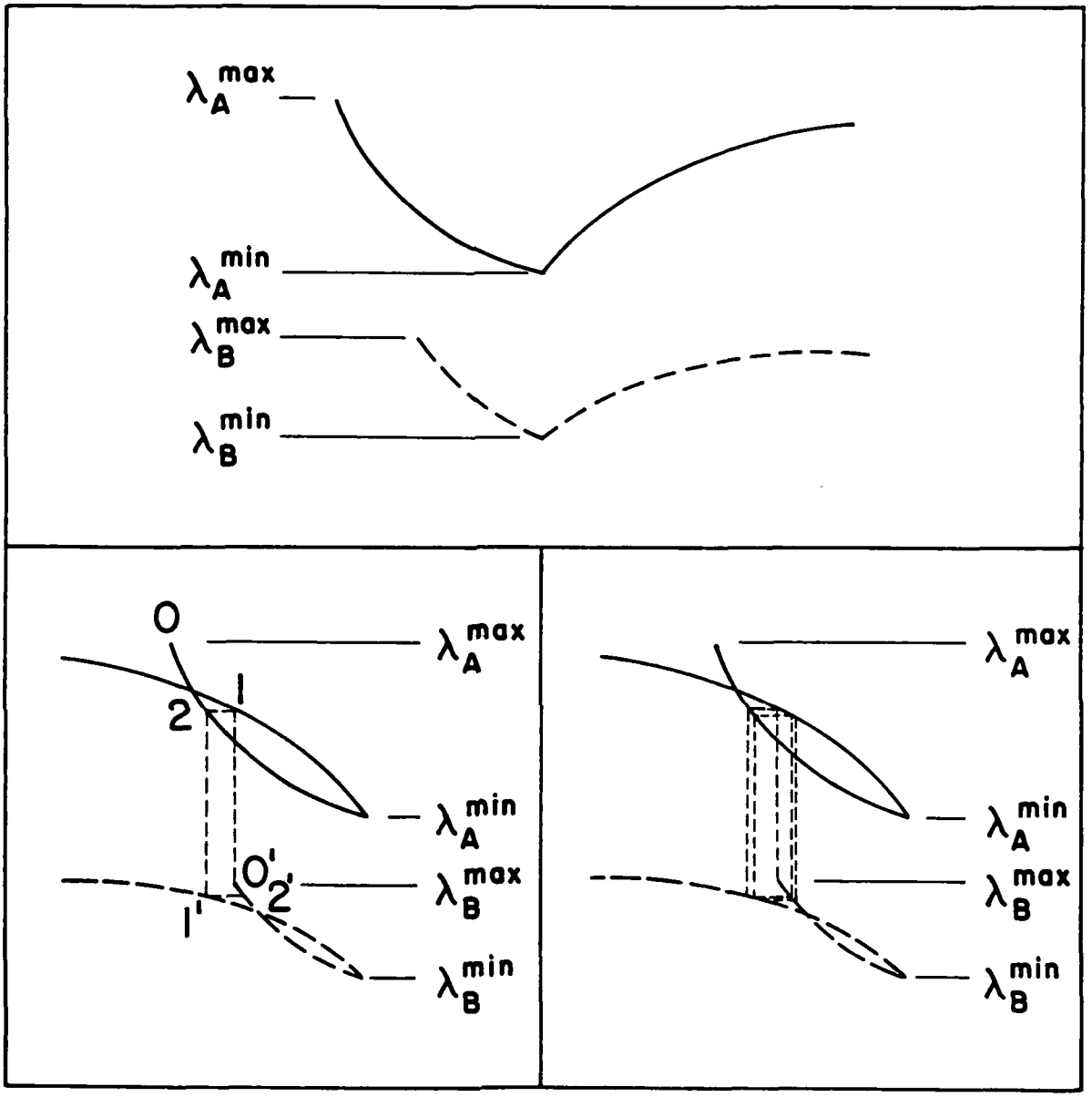

INPUT
OUTPUT
SET OF N NEURONS
α
SHOWS ACTIVITY PATTERN
f̄
SET OF N NEURONS
β
SHOWS ACTIVITY PATTERN
ḡ



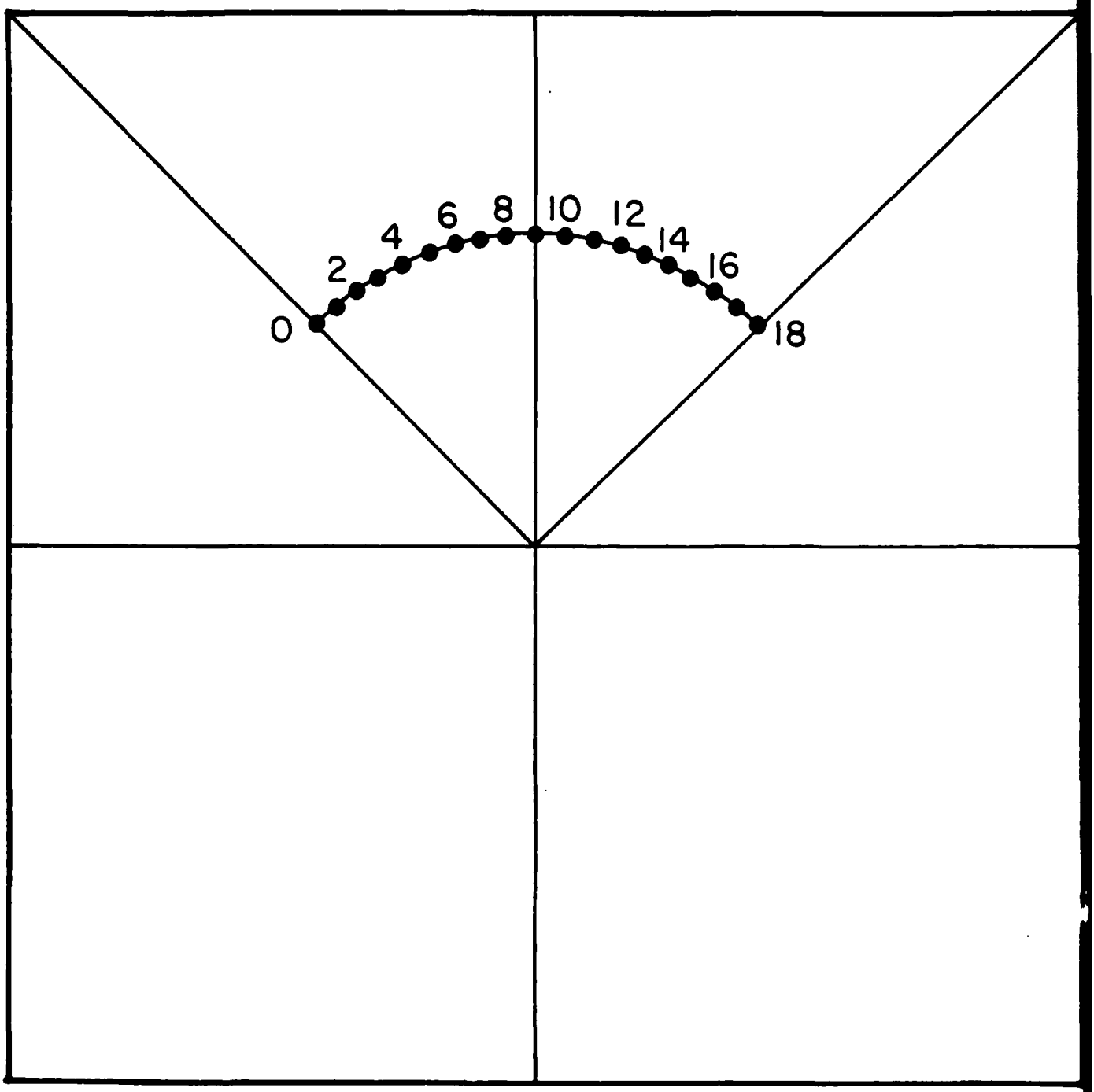
1. SET OF N NEURONS, α
2. EVERY NEURON IN α IS CONNECTED TO EVERY OTHER NEURON IN α THROUGH LEARNING

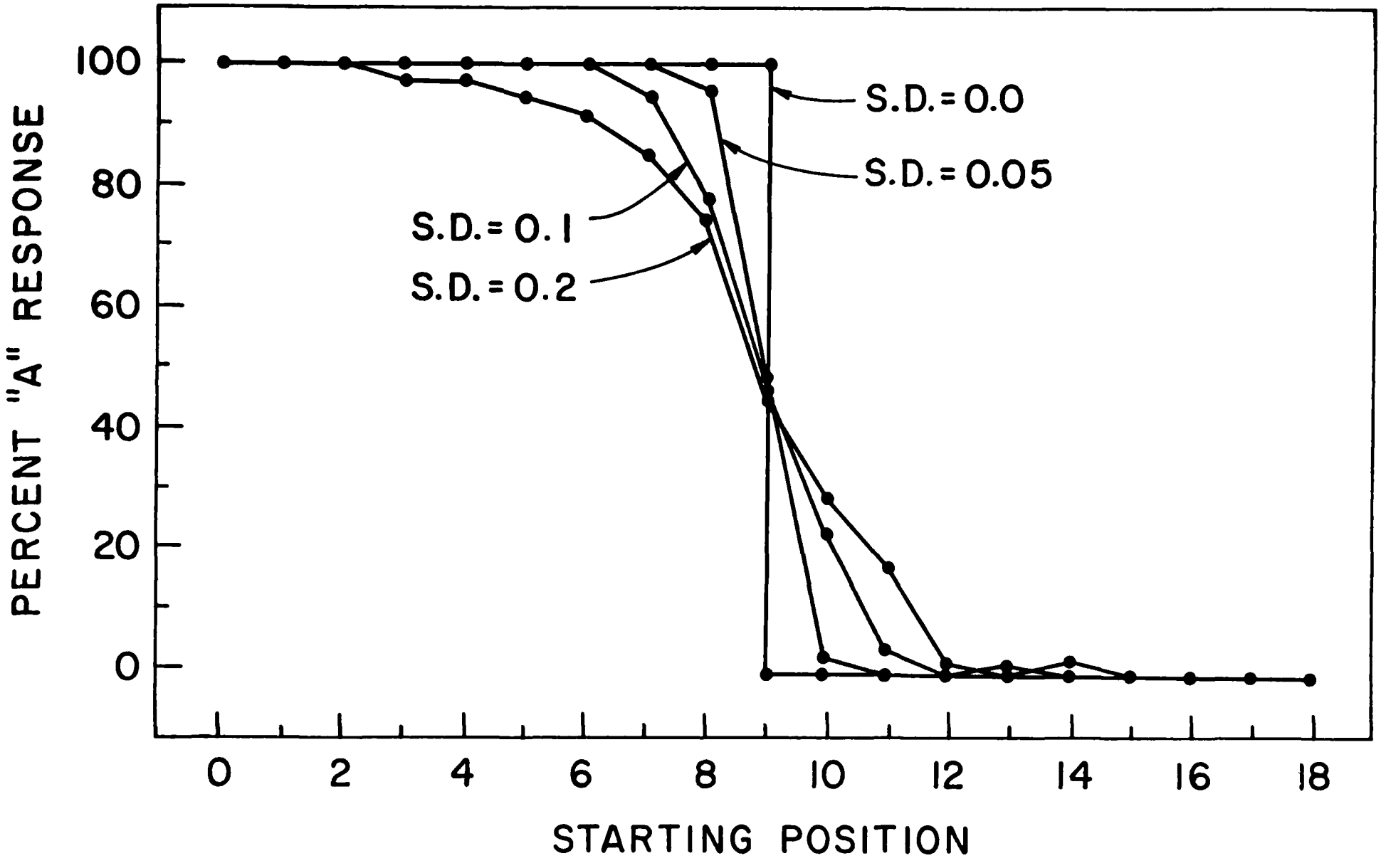
MATRIX OF SYNAPTIC CONNECTIVITIES A

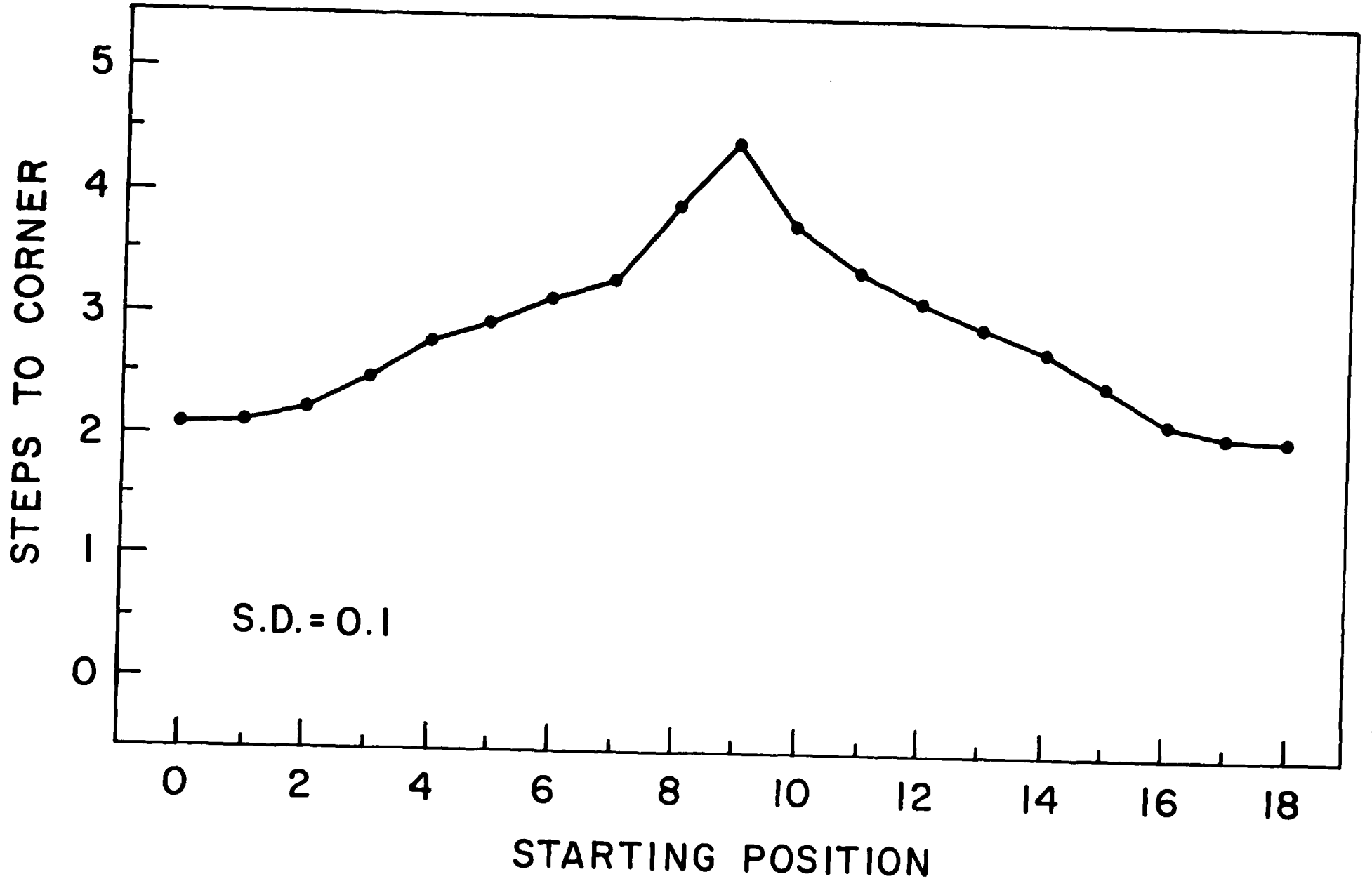

EIGENVALUE MAGNITUDE
A
B
A
B
A
B
A
λ_A^{max}
λ_B^{max}
λ_B^{min}
λ_A^{min}
STIMULUS ONSET
TIME

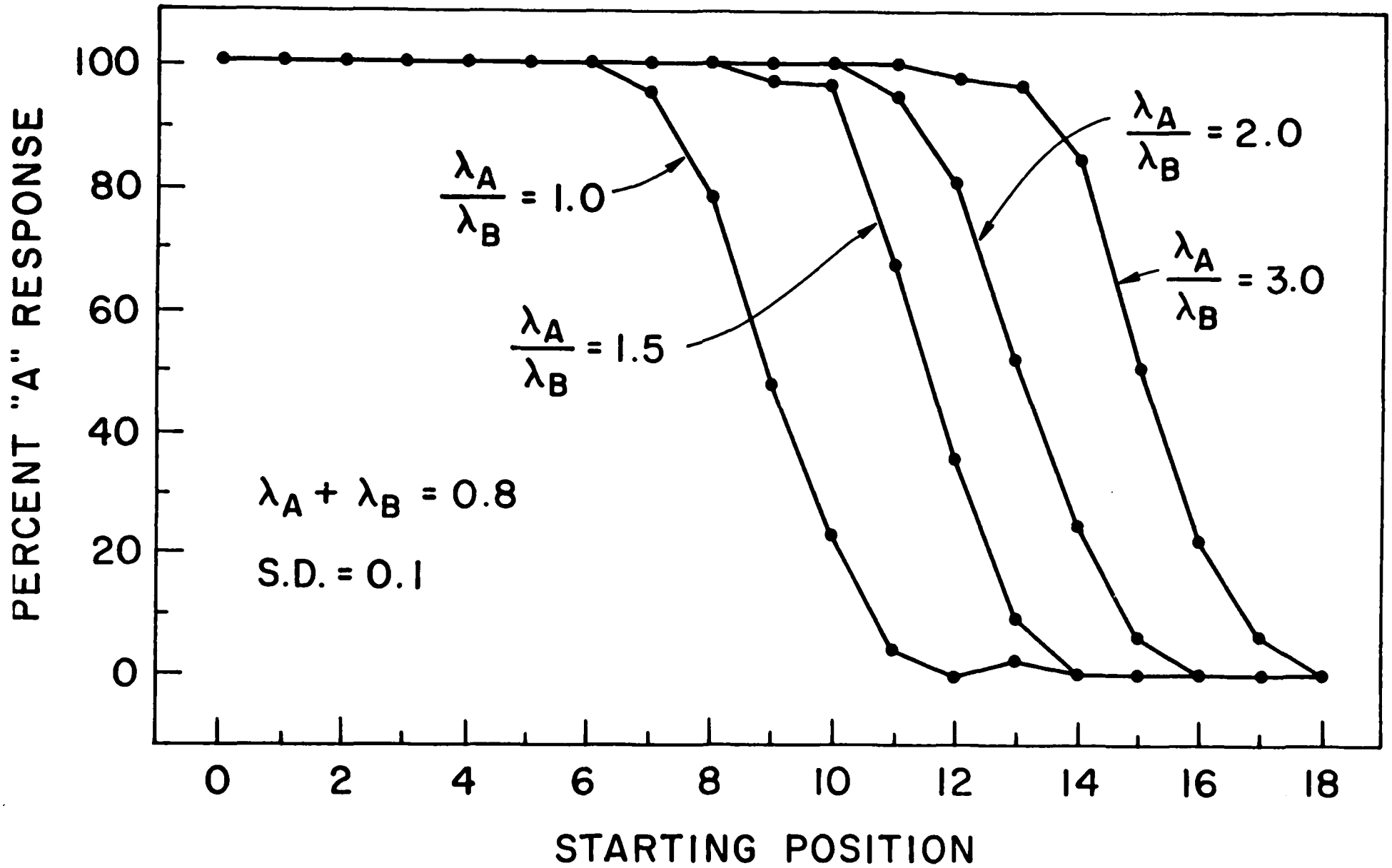

λ_A^{max}
λ_A^{min}
λ_B^{max}
λ_B^{min}
0
1
2
$0'$
$1'$
$2'$
λ_A^{max}
λ_A^{min}
λ_B^{max}
λ_B^{min}
λ_A^{max}
λ_A^{min}
λ_B^{max}
λ_B^{min}

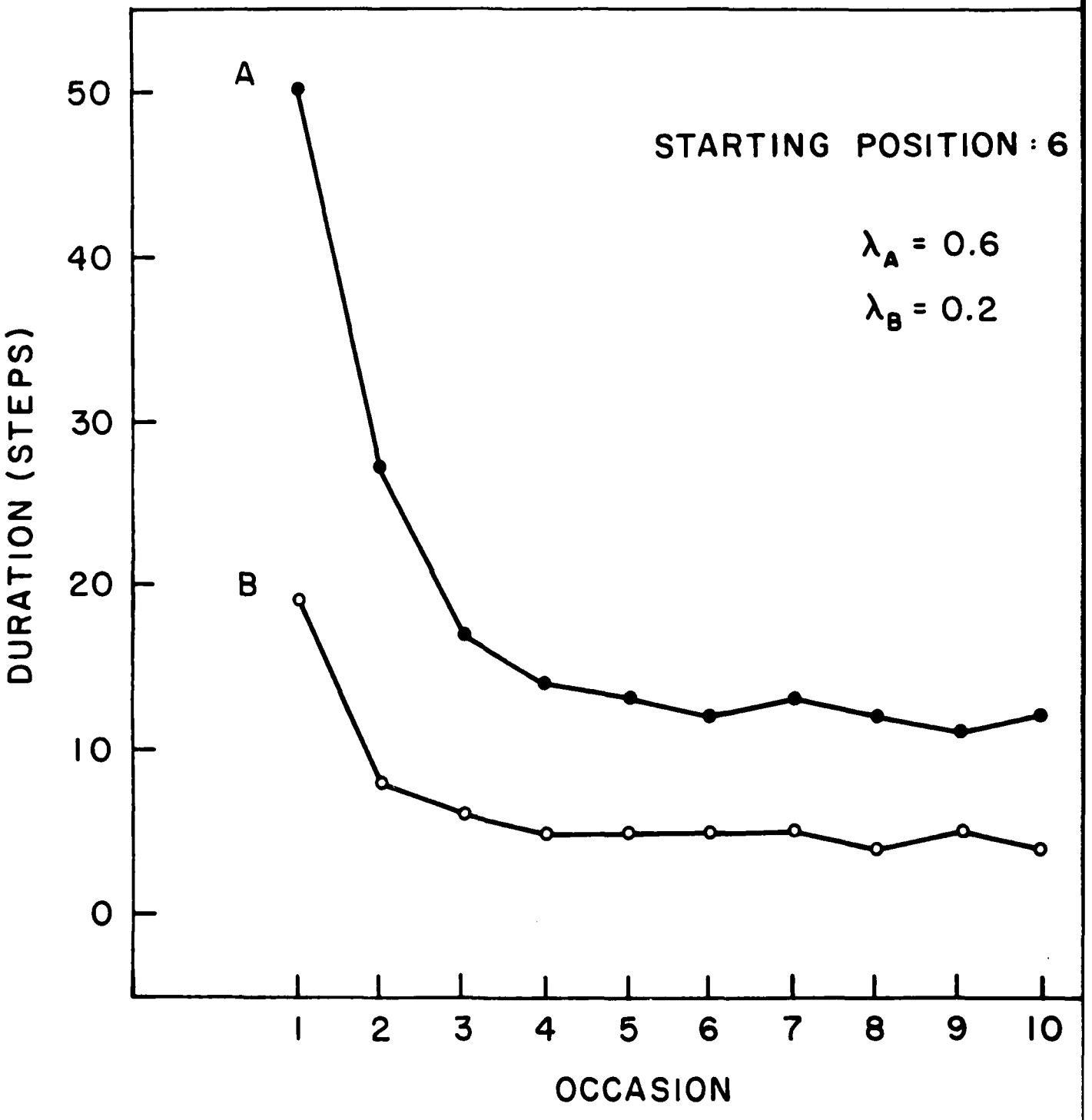
A (1,1,1,1,-1,-1,-1,-1) (1,-1,1,-1,1,-1,1,-1) B

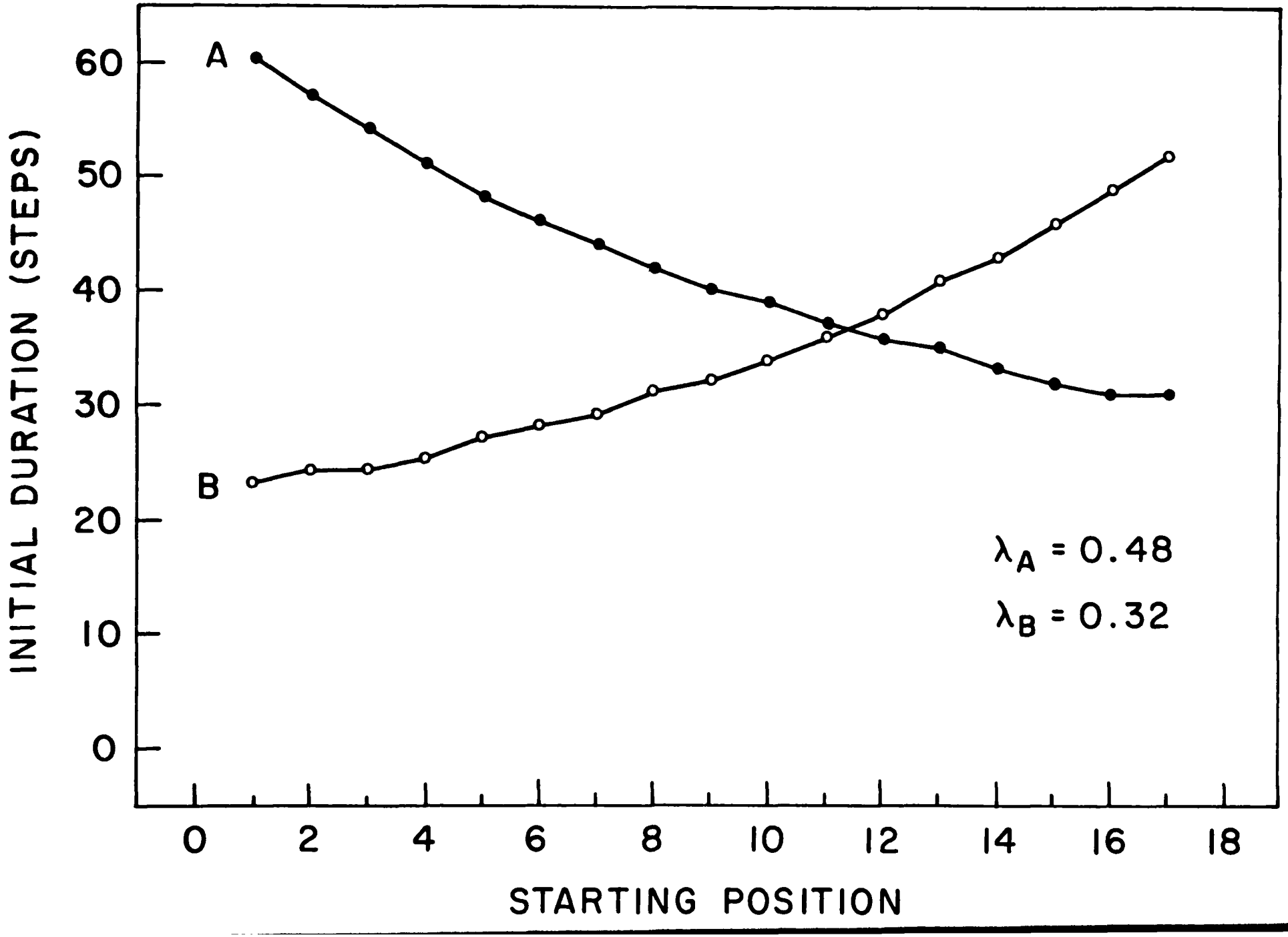


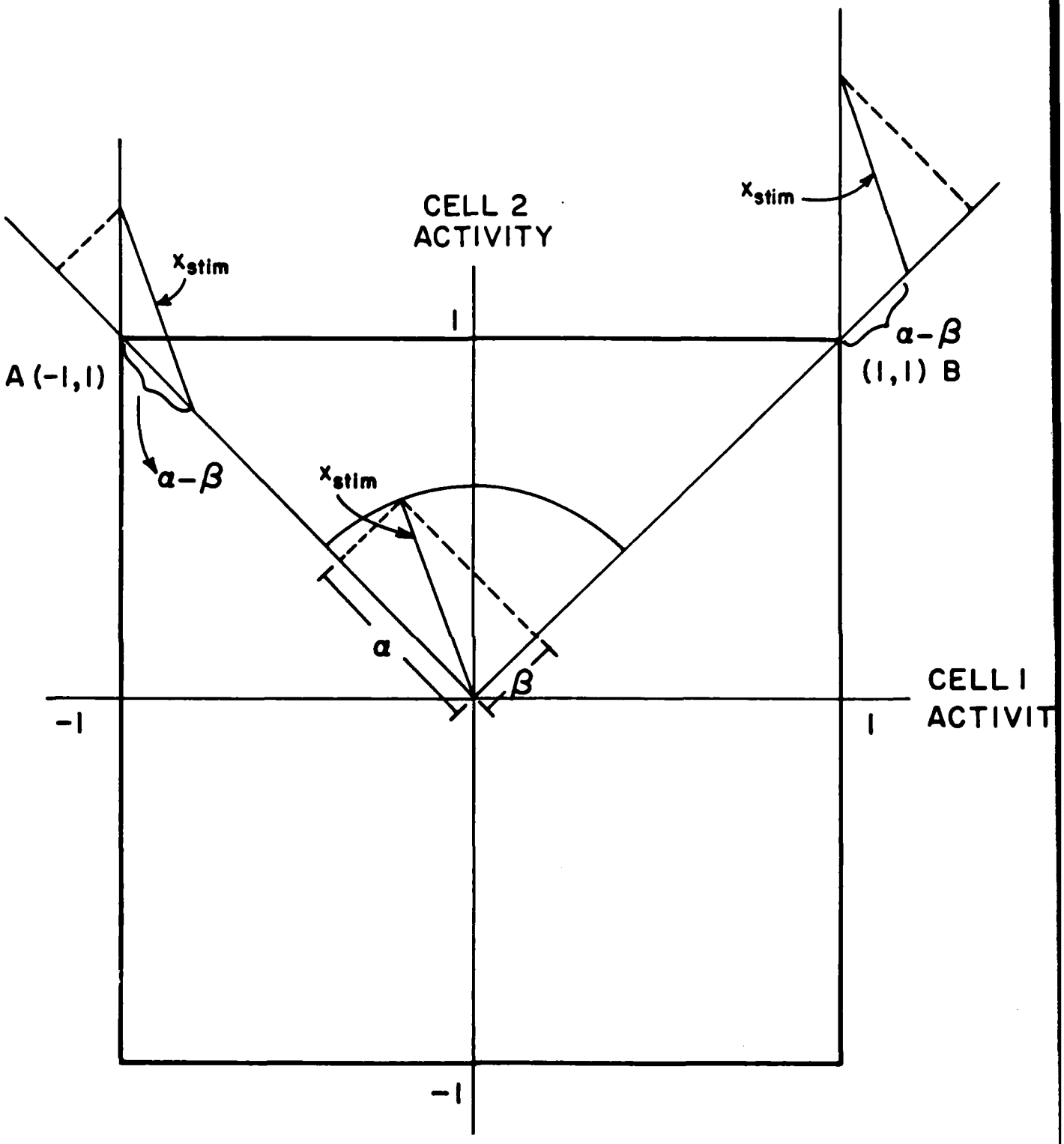

PERCENT "A" RESPONSE
100
80
60
40
20
0
0
2
4
6
8
10
12
14
16
18
STARTING POSITION
S.D. = 0.0
S.D. = 0.05
S.D. = 0.1
S.D. = 0.2

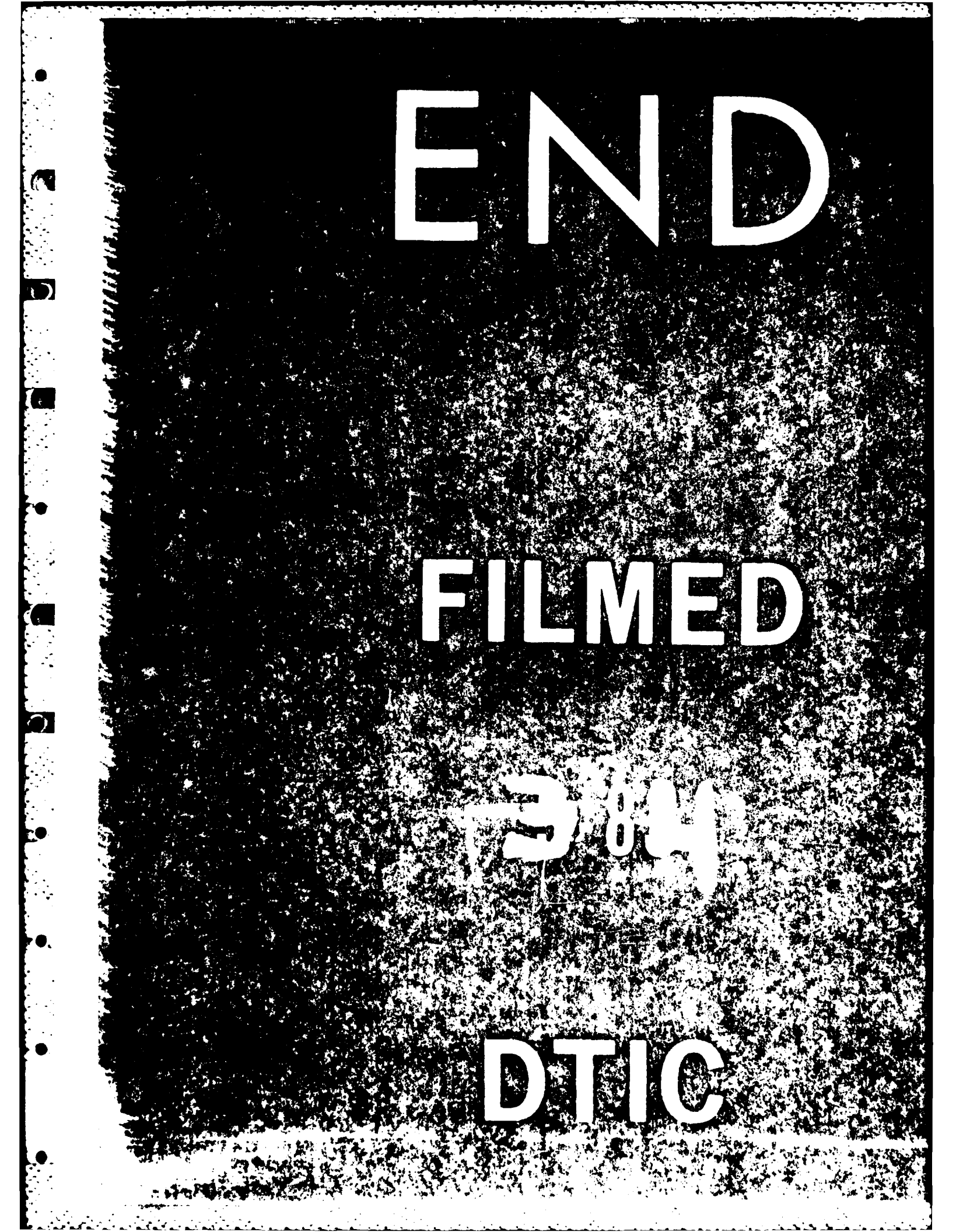

STEPS TO CORNER
5
4
3
2
1
0
S.D. = 0.1
0
2
4
6
8
10
12
14
16
18
STARTING POSITION


PERCENT "A" RESPONSE
100
80
60
40
20
0
0
2
4
6
8
10
12
14
16
18
STARTING POSITION
$\frac{\lambda_A}{\lambda_B} = 1.0$
$\frac{\lambda_A}{\lambda_B} = 1.5$
$\frac{\lambda_A}{\lambda_B} = 2.0$
$\frac{\lambda_A}{\lambda_B} = 3.0$
$\lambda_A + \lambda_B = 0.8$
S.D. = 0.1


A
B
STARTING POSITION : 6
λ_A = 0.6
λ_B = 0.2
DURATION (STEPS)
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
10
OCCASION


INITIAL DURATION (STEPS)
60
50
40
30
20
10
0
A
B
$\lambda_A = 0.48$
$\lambda_B = 0.32$
0
2
4
6
8
10
12
14
16
18
STARTING POSITION


CELL 2
ACTIVITY
x_{stim}
x_{stim}
x_{stim}
1
A (-1,1)
(1,1) B
$\alpha-\beta$
$\alpha-\beta$
α
β
-1
1
CELL 1
ACTIVIT
-1


END
FILMED
DTIC