Psychological Review
1978, Vol. 85, No. 6, 592-596

Do All Neural Models Really Look Alike? A Comment on Anderson, Silverstein, Ritz, and Jones

Stephen Grossberg
Department of Mathematics
Boston University

Several of the formal approaches that are used to explain psychophysiological phenomena lead to different properties and principles of organization. These approaches include computer, linear, and nonlinear models. The present note illustrates this by citing differences between the linear theory of Anderson et al. and the nonlinear theory of Grossberg. For example, when linear signals are joined to neural mechanisms, the resultant model is unstable and is unable to adapt or retune its sensitivity in response to changing input patterns. Nonlinear signals overcome these difficulties and also compute reflectances, a Weber–Fechner law, and complex decision schemes that transform input patterns before they are stored in short-term memory. The nonlinear theory also expresses organizational principles that have been used to analyze a class of problems in perception, reinforcement, sensory-motor development, and goal-oriented cognitive behavior.

Noise-Saturation Dilemma: A Universal Constraint on Cellular Information Processing

The article by Anderson, Silverstein, Ritz, and Jones (1977) suggests that concepts from linear system theory, notably the eigenvalues and eigenvectors of a symmetric matrix, can be used to describe neural mechamisms related to distinctive features, categorical perception, and probability learning. The authors note that the assumption of linearity is among some "grievous oversimplifications of the physiology" (Anderson et al., 1977, p. 444). I will indicate below that when linear signals are joined to well-known neural mechanisms, they are inherently unstable. Stability can only be salvaged using nonlinear mechanisms. Moreover, these nonlinear mechanisms have previously been mathematically analyzed and exhibit phenomena such as those reported by the authors, but for reasons that are not best understood by means of eigenvalues and eigenvectors.

The linear models also do not exhibit certain other properties that occur in the nonlinear models as well as in a diverse body of psychological data (Grossberg, 1978a). These properties solve a universal problem about information processing to which all cellular systems are subjected, namely, the noise-saturation dilemma. This dilemma is easy to state: Let a pattern of inputs $(I_1, I_2, \ldots, I_n)$ excite a collection of cells $v_1, v_2, \ldots, v_n$. All cells experience in vivo a certain amount of noise, and all cells possess only finitely many excitable sites. If the inputs in the pattern are too small, they get lost in the noise. If they are too large, they can turn on all the excitable sites in all the cells; that is, saturation, or sensitivity loss, occurs. How can cellular systems balance between the two deadly extremes of noise and saturation? The answer is that competitive interactions between the cells or their inputs cause the cells to automatically adapt, or retune, their sensitivity in response to fluctuations in the size of feedback signals or inputs (Grossberg, 1970, 1972a, 1973; Sperling, 1970). Feedback signals exist, for example, when the cells reverberate activity in short-term memory. We will indicate below why a self-tuning system in which feedback signals occur is nonlinear. Thus, linear models omit a fundamental principle of neural design: They cannot retune, or adapt, themselves to fluctuating input demands. Another way of saying this

This research was supported in part by the National Science Foundation (Grant NSF MCS 77-02958).

Requests for reprints should be sent to Stephen Grossberg, Department of Mathematics, 264 Bay State Road, Fourth Floor, Boston University, Boston, Massachusetts 02215.

 0033-295X/78/8506-0592$00.75

is that linear models omit the nerve cells from neural models, but cells are there for important functional reasons.

Anderson et al. (1977) seem to have a misconception about recent neural data that prevents them from addressing the noise-saturation dilemma. They cite recent data of Creutzfeldt, Innocenti, and Brooks (1974), which they say suggests that "most cells in primary visual cortex . . . are not strongly coupled together, again implying a good deal of individuality of cell response. . . . This assumption allows us to represent these large-scale activity patterns as vectors of high dimensionality with *independent components*" (Anderson et al., 1977, p. 415; italics mine). However, Creutzfeldt (1976) writes "Let us assume that the cortical network consists of a system of short-ranging excitatory and of wider-ranging inhibitory interactions among cortical neurons" (p. 457). It is shown below that if the cells in the cortex did have independent components, then they could not easily survive the noise-saturation dilemma; but if they have short-range excitatory interactions and longer range inhibitory interactions, then they can completely solve the noise-saturation dilemma.

Positive Eigenvalues Cause Catastrophes in Linear Systems

Anderson et al. (1977) consider linear systems in which positive eigenvalues exist. Their goal is to selectively amplify the neural activities of certain internal states, or representations, at the expense of other states. A simple example illustrates the main idea. Let there be n cells v_i, where $i = 1, 2, \ldots, n$, and let $x_i(t)$ be the activity or potential of v_i at time t. Suppose that the rate of change of x_i, namely, $\dot{x}_i$, satisfies the equation $\dot{x}_i = a_i x_i$, where $i = 1, 2, \ldots, n$ and where $a_1 \geq a_2 \geq \cdots \geq a_n > 0$. This is a linear system, albeit a very simple one, with positive eigenvalues $a_1, a_2, \ldots, a_n$. If, for example, $a_1 > a_2$, then $x_1(t)$ becomes much larger than all other $x_i(t)$ as t becomes large, since each $x_i(t) = x_i(0)e^{a_i t}$. This is the main use of positive eigenvalues in the Anderson et al. article. However, because the eigenvalues are positive, all $x_i(t)$ approach infinity as t becomes large, which is physically meaningless. The authors realize this and note that "unfortunately the desirable features of positive feedback are exactly the ones that cause catastrophes" (Anderson et al., 1977, p. 427). To overcome these catastrophes, the authors prevent the activities from changing by fiat after they reach a certain maximum or minimal value. This rather nonphysical property they call "hard saturation" (Anderson et al., 1977, p. 444). In vivo, saturation also occurs, but it is of a different type. If neural saturation is added to this linear model, then differential amplification is destroyed, and sensitivity loss as well as noise amplification occurs.

To see this, we note that in vivo the equation for a cell's activity, or voltage $V(t)$, often has the form

$$C\frac{\partial V}{\partial t} = (V^+ - V)g^+ + (V^- - V)g^- + (V^p - V)g^p, \quad (1)$$

where C is a capacitance; the constants V^+, V^-, and V^p are excitatory, inhibitory, and passive saturation points, respectively, and g^+, g^-, and g^p are conductances that can be changed by inputs (Hodgkin, 1964; Katz, 1966). We adopt the convention that $V^+ > V^p \geq V^-$. The voltage $V(t)$ remains between V^+ and V^-. Saturation can occur in Equation 1, for example, if g^+ becomes so large that $V(t)$ approaches V^+. Then, $(V^+ - V)g^+ \cong 0$, so g^+ thereafter has little influence on $\frac{\partial V}{\partial t}$. How does neural saturation influence the linear system $\dot{x}_i = a_i x_i$, where $i = 1, 2, \ldots, n$? For example, let $\dot{x}_i = a_i x_i(1 - x_i)$, where $i = 1, 2, \ldots, n$. The term $1 - x_i$ is analogous to the term $V^+ - V$ in Equation 1. Suppose that all x_i start out positive but can have any values whatever between 0 and 1. Then, all x_i approach 1 as t increases. No matter how small or different the x_i are initially, they all eventually saturate at 1. As t increases, there is a complete loss of sensitivity to the initial differences in the x_i. Noise amplification also occurs, since x_is that are initially very small, say due to a little prior noise, become maximally large as t increases.

A Solution: Competition in Mass-Action Systems

The problem of how to avoid these pathologies was solved in Grossberg (1973). Below I sketch relevant aspects of the solution and show how differential amplification occurs in nonlinear neural networks whose linearizations, near their equilibrium point, have negative eigenvalues. Moreover, these nonlinear networks automatically adapt, or retune, their sensitivity in response to fluctuating input demands.

First, we provide a heuristic derivation of Equation 1 to clarify its statistical interpretation. Suppose that n cells v_i are given and that each cell v_i is subjected to an input $I_i(t)$. Let B be the total number of excitable sites in v_i; let $x_i(t)$ be the number of excited sites or the potential at time t; and let $B - x_i(t)$ be the number of unexcited sites at time t. Suppose that excited sites spontaneously become unexcited at rate A. Also let unexcited sites be excited by mass action at a rate jointly proportional to their number $B - x_i(t)$ and the input intensity $I_i(t)$. Then,

$$\dot{x}_i = -Ax_i + (B - x_i)I_i, \tag{2}$$

with $0 \leq x_i \leq B$, where $i = 1, 2, \ldots, n$. Each x_i in Equation 2 saturates at B as each I_i increases. How can this be prevented? This sensitivity loss can be explicated as follows. Often the information in an input pattern $(I_1, I_2, \ldots, I_n)$ resides in the relative input sizes $(\theta_1, \theta_2, \ldots, \theta_n)$, where $\theta_i = I_i I^{-1}$, and $I = \sum_{k=1}^{n} I_k$ is the total, or background, input activity. For example, in vision, these relative sizes are the "reflectances" (Cornsweet, 1970) of the pattern. How can a system be designed that remains sensitive to the ratios θ_i as the background input intensity I is parametrically increased?

In order for each v_i to compute a ratio θ_i, it must know what all the inputs are. Writing $\theta_i = I_i(I_i + \sum_{k \neq i} I_k)^{-1}$, it is clear that increasing I_i increases θ_i and that increasing any I_k, where $k \neq i$, decreases θ_i. In other words, I_i "excites" θ_i, as in Equation 2; whereas all I_k, where $k \neq i$, "inhibit" θ_i. Thus, the inputs must compete at each v_i in order to prevent saturation and thereby compute the reflectances. When this intuition is translated into mass-action dynamics, we find a feed-forward on-center (excite v_i) off-surround (inhibit all v_k, $k \neq i$) interaction pattern that obeys a law like Equation 1 in which $V = x_i$, $C = 1$, $V^+ = B$, $V^- = V^p = 0$, $g^p = A$, $g^+ = I_i$, and $g^- = \sum_{k \neq i} I_k$; namely,

$$\dot{x}_i = -Ax_i + (B - x_i)I_i - x_i \sum_{k \neq i} I_k. \tag{3}$$

The new term $-x_i \sum_{k \neq i} I_k$ says that excited sites at v_i (which number x_i) are inhibited (note the minus sign) at a rate proportional to the total inhibitory input (which is a sum of inputs from the off-surround of v_i). Thus, the off-surround automatically changes the gain, or decay rate, of the system by multiplying x_i. This multiplication of x_i by inputs is called a *shunt*. Additive models, such as the Hartline–Ratliff model, do not automatically retune themselves (Ratliff, 1965).

In response to fixed reflectances $(\theta_1, \theta_2, \ldots, \theta_n)$ and a fixed background input I, the x_is in Equation 3 approach the equilibrium activities:

$$x_i = \theta_i \frac{BI}{A + I}. \tag{4}$$

(Set $\dot{x}_i = 0$ and solve for x_i.) By Equation 4, no matter how large I becomes, each x_i is proportional to θ_i. There is no saturation because the off-surround automatically retunes x_i's sensitivity as I increases. Note also that term $BI(A + I)^{-1}$ has the form of the Weber–Fechner law. This sensitivity shift can also be described in the following terms: Write I_i in logarithmic coordinates as $K = \log I_i$, and let $L = \sum_{k \neq i} I_k$ be the off-surround input. Then, Equation 4 can be rewritten as

$$x_i(K, L) = \frac{Be^K}{A + e^K + L}. \tag{5}$$

Using Equation 5, we can prove that a shift in L shifts x_i's response curve without causing any sensitivity loss. In fact, if L is changed from $L = L_1$ to $L = L_2$, then the shift is predicted to be

$$S = \ln[(A + L_1)(A + L_2)^{-1}], \tag{6}$$

since

$$x_i(K + S, L_1) = x_i(K, L_2) \quad \text{for all } K \geq 0. \tag{7}$$

A shift of this kind occurs in intracellular data taken from bipolar cells in the mudpuppy retina (Werblin, 1971). In summary, the saturation problem is overcome by mass-action systems undergoing competitive interactions. In the language of neural networks, these systems are said to obey the shunting law (Equation 1) in an on-center, off-surround anatomy. Cells in such networks automatically compute reflectances, a Weber–Fechner law and shift their response curves when the inputs to their off-surrounds are parametrically changed.

System 3 is linear in the variable x_i, despite the nonlinear occurrence of I in Equation 4 because it experiences no feedback signals. Such signals are, for example, needed to keep short-term memory traces active after external inputs terminate. To prevent saturation, the feedback signals should also be distributed

in an on-center off-surround anatomy. Then, Equation 3 is replaced by the nonlinear system

$$\dot{x}_i = -Ax_i + (B - x_i)[f(x_i) + I_i] - x_i[\sum_{k \neq i} f(x_k) + J_i], \quad (8)$$

where $i = 1, 2, \ldots, n$. Term $(B - x_i)f(x_i)$ describes how a positive feedback signal $f(x_i)$ from v_i to itself excites the unexcited sites $B - x_i$ by mass action. The inhibitory term $-x_i \sum_{k \neq i} f(x_k)$ describes the switching off of excitation at v_i by inhibitory signals $f(x_k)$ from all v_k, where $k \neq i$. Term I_i is the total excitatory input, and term J_i is the total inhibitory input at v_i.

How should the feedback signal $f(w)$ be chosen as a function of cell activity w? Grossberg (1973) proves that linear signals [e.g., $f(w) = Cw$] and slower-than-linear signals [e.g., $f(w) = Cw(D + w)^{-1}$] both amplify noise and are therefore unacceptable. Faster-than-linear signals [e.g., $f(w) = w^2$] suppress noise but contrast enhance the system so drastically that it acts like a choice, or finite-state, machine exhibiting such phenomena as categorical perception. Sigmoid, or S-shaped, signals have very desirable properties, since they cause a *quenching threshold* (QT) to exist. Activities x_i that are smaller than the QT are treated like noise and quenched, or suppressed. Activities larger than the QT are contrast enhanced and stored in short-term memory. Because the QT exists, varying the network's arousal level can retune which activities will be stored in short-term memory, for example, a low arousal level can generate a rapid choice. This is again an adaptational effect. Sigmoid signals occur, for example, in the data of Freeman (1975) that Anderson et al. cite.

Now we make a crucial point. Every signal $f(w)$ that suppresses noise in Equation 8 is nonlinear at small values of w. Thus, if we shut off the inputs I_i and J_i and linearize Equation 8 near the equilibrium point $x_i = 0$, where $i = 1, 2, \ldots, n$, then we find $\dot{x}_i = -Ax_i$, where $i = 1, 2, \ldots, n$, which has the negative eigenvalue $-A$ despite the system's ability to differentially amplify certain activities at the expense of others. Positive eigenvalues can occur when the system is linearized near other values of the x_i. The nonlinear systems can hereby blend together several reaction tendencies in a way that transcends the capabilities of a linear system.

System Equation 8 is a very special example, to be sure, but our conclusions can be generalized to large classes of nonlinear neural networks (Ellias & Grossberg, 1975; Levine & Grossberg, 1976). Indeed, I have recently proved that every competitive system induces a decision scheme that can be used to globally characterize its dynamics, in particular, its noise amplification properties (Grossberg, 1978b).

By contrast with the Anderson et al. linear theory of distinctive features, nonlinear interactions have been broadly used in other neural models of feature extraction (Grossberg, 1976a, 1976b; Pérez, Glass, & Shlaer, 1975; Von der Malsburg, 1973). Von der Malsburg (1973) and Pérez et al. (1975) used computer studies to illustrate how a short list of stimulus patterns can retune the responsiveness of a small number of feature detectors. In Grossberg (1976a, 1976b), it is shown how a stable, globally self-consistent hierarchy of codes can develop in response to an arbitrary input environment if suitably defined feedback expectations, antagonistic rebounds, coding, pattern learning, and matching mechanisms exist—all of which are defined by nonlinear laws. Also an interpretation and extension of stimulus sampling theory using nonlinear mechanisms has been given (Grossberg, 1972b, 1972c, 1974). A general analysis of the nonlinear feedback dynamics of systems undergoing plastic (long-term memory) changes has been established (e.g., Grossberg, 1971, 1972d, 1974) and does not require the simplifying symmetry assumption $\Delta a_{ij} = \Delta a_{ji}$ of the Anderson et al. analysis. Finally, these results have been synthesized to suggest a theory of human memory in which a few nonlinear principles that express solutions to environmental pressures to which a surviving species must adapt are used to attack a class of problems in perception, sensory–motor development, and goal-oriented cognitive behavior (Grossberg, 1978c).

The mechanisms and underlying principles in the linear versus nonlinear models are as different, say, as the mechanisms in computer versus linear models. Deciding between these alternative descriptions is an important problem for psychophysiology.

References

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

Cornsweet, T. N. *Visual perception*. New York: Academic Press, 1970.

Creutzfeldt, O. The brain as a functional entity. In M. A. Corner & D. F. Swaab (Eds.), *Perspective in brain research: Progress in brain research* (Vol. 45). Amsterdam, The Netherlands: Elsevier/North-Holland Biomedical Press, 1976.

Creutzfeldt, O., Innocenti, G. M., & Brooks, D. Vertical organization in the visual area (Area 17) in the cat. *Experimental Brain Research*, 1974, *21*, 315–336.

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

Freeman, W. J. *Mass action in the nervous system*. New York: Academic Press, 1975.

Grossberg, S. Neural pattern discrimination. *Journal of Theoretical Biology*, 1970, *27*, 291–337.

Grossberg, S. Pavlovian pattern learning by nonlinear neural networks. *Proceedings of the National Academy of Sciences*, 1971, *68*, 828–831.

Grossberg, S. Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, 1972, *10*, 49–57. (a)

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

Grossberg, S. A neural theory of punishment and avoidance, II. *Mathematical Biosciences*, 1972, *15*, 253–285. (c)

Grossberg, S. Pattern learning by functional-differential neural networks with arbitrary path weights. In K. Schmitt (Ed.), *Delay and functional-differential equations and their applications*. New York: Academic Press, 1972. (d)

Grossberg, S. Contour enhancement, short term memory, and constancies in reverberating neural networks. *Studies in Applied Mathematics*, 1973, *52*, 217–257.

Grossberg, S. Classical and instrumental learning by neural networks. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 3). New York: Academic Press, 1974.

Grossberg, S. Adaptive pattern classification and universal recoding, I. *Biological Cybernetics*, 1976, *23*, 121–134. (a)

Grossberg, S. Adaptive pattern classification and universal recoding, II. *Biological Cybernetics*, 1976, *23*, 187–202. (b)

Grossberg, S. Behavioral contrast in short term memory: Binary serial memory models or continuous parallel memory models. *Journal of Mathematical Psychology*, 1978, *17*, 199–219. (a)

Grossberg, S. Decisions, patterns, and oscillations in nonlinear competitive systems, with applications to Volterra–Lotka systems. *Journal of Theoretical Biology*, 1978, *73*, 101–130. (b)

Grossberg, S. A theory of human memory: Self-organization and performance of sensory–motor codes, maps, and plans. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5). New York: Academic Press, 1978. (c)

Hodgkin, A. L. *The conduction of the nervous impulse*. Springfield, Ill.: Charles C Thomas, 1964.

Katz, B. *Nerve, muscle, synapse*. New York: McGraw-Hill, 1966.

Levine, D. S., & Grossberg, S. Visual illusions in neural networks: Hysteresis, angle expansion, tilt aftereffect. *Journal of Theoretical Biology*, 1976, *61*, 477–504.

Pérez, R., Glass, L., & Shlaer, R. Development of specificity in the cat visual cortex. *Journal of Mathel matical Biology*, 1975, *1*, 275–288.

Ratliff, F. *Mach bands: Quantitative studies of neura-networks in the retina*. San Francisco: Holden-Day, 1965.

Sperling, G. Model of visual adaptation and contrast detection. *Perception & Psychophysics*, 1970, *8*, 143–157.

Von der Malsburg, C. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 1973, *14*, 85–100.

Werblin, F. S. Adaptation in a vertebrate retina: Intracellular recording in Necturus. *Journal of Neurophysiology*, 1971, *34*, 228–241.

Received March 10, 1978 ■