Reprinted from BEHAVIORAL SCIENCE
Vol. 15, No. 4, July, 1970
Printed in U.S.A.

RESEARCH REPORTS

SCHIZOPHRENIA: POSSIBLE DEPENDENCE OF ASSOCIATIONAL SPAN, BOWING, AND PRIMACY VS. RECENCY ON SPIKING THRESHOLD

by Stephen Grossberg[1] *and James Pepe, Department of Mathematics, M.I.T.*

INTRODUCTION

THE hypothesis has been advanced that certain schizophrenic patients are in a continual state of overarousal, leading to poor attention, and perhaps to schizophrenic punning (Kornetsky and Eliasson, 1969; Maher, 1968). Physiological factors that can yield overarousal include a pathological reduction in spiking thresholds of cells that send signals to many other cells, or alternatively a reduction in the strength of lateral inhibitory interactions between these cells. Even low energy but persistent peripheral excitation can under these circumstances build up high internal noise levels, because signals between the cells will not be sufficiently damped by the threshold or inhibitory mechanisms.

This note announces the occurrence of analogous phenomena in a rigorously defined learning network having a suggestive psychological, neurophysiological, anatomical, and biochemical interpretation (Grossberg, 1969*a*–*c*). We study in this network the serial learning of a long list of behavioral events as it depends on spiking threshold. In normal subjects, one characteristically finds such phenomena as bowing (middle of the list harder to learn than the ends) and primacy dominating recency (beginning easier to learn than the end) (Grossberg, 1969*d*). Altering the spiking threshold of the network systematically alters these effects.

For example, as the spiking threshold decreases, recency gradually gains strength over primacy until finally recency prevails. This phenomenon is due to the buildup of background noise as a result of the persistent presentation of events in the serial paradigm. As ever more serial events are presented, the associational strengths near the list's beginning are eventually competitively inhibited by incorrect associations that form due to inadequate threshold cut-offs. Corresponding difficulties in "paying attention" to long behavior dependencies also occur, and "punning" based on low-order associations becomes plausible.

A NEURAL LEARNING NETWORK

Let n cell bodies v_i, $i = 1, 2, \cdots, n$, be given. (Alternatively, interpret each v_i as a cell body cluster.) Let each cell v_i send a directed axon e_{ij} to every distinct cell v_j, $j \neq i$. Denote the synaptic knob of e_{ij} by N_{ij}.

Denote the average membrane potential of v_k at time t by $x_k(t)$, and denote the average amount of available excitatory transmitter substance in N_{ki} at time t by $z_{ki}(t)$, $i \neq k$. At every time t, let v_k create a spiking frequency proportional to $[x_k(t) - \Gamma]^+$ in e_{ki}, $i \neq k$, where $[w]^+ = \max(w, 0)$ for every real number w, and Γ is the spiking threshold of e_{ki}. Suppose that this signal reaches N_{ki} at time $t + \tau$, and thereupon causes release of excitatory transmitter from N_{ki} at a rate proportional to $[x_k(t) - \Gamma]^+ z_{ki}(t + \tau)$. Let all such signals from cells v_k, $i \neq k$, combine additively at v_i, and cause a proportional alteration in $x_i(t + \tau)$'s rate of change; that is, in $\dot{x}_i(t + \tau)$. Also let x_i decay at an exponential rate α, and perturb x_i with the experimental input I_i. These rules are equivalent to the system of equations

$$(1) \quad \dot{x}_i(t) = -\alpha x_i(t) + \beta \sum_{k \neq i} [x_k(t-\tau) - \Gamma]^+ z_{ki}(t) + I_i(t), \qquad i = 1, 2, \cdots, n.$$

---

[1] Supported in part by the A. P. Sloan Foundation (71609), the NSF (GP-13778), and the ONR (N00014-67-A-0204-0016).

The amount $z_{ij}(t)$ of available transmitter in N_{ij} is determined jointly by the presynaptic spiking frequency from v_i and the postsynaptic potential of v_j, as in

$$(2) \quad \dot{z}_{ij}(t) = -\gamma z_{ij}(t) + \delta[x_i(t - \tau) - \Gamma]^+ x_j(t), \; i \neq j.$$

Systems of Types 1 and 2, known generically as *embedding fields*, have been mathematically proved capable of discriminating, learning, and performing complicated tasks (Grossberg, 1969*e*, *f*; 1970*a*,*b*). The mathematical success of these systems, and their derivation from simple psychological hypotheses (Grossberg, 1969*a*) lends some weight to the as yet hypothetical Equation 2 for regulation of transmitter production. All the results of this note also depend strongly on joint pre- and post-synaptic control of transmitter production. The results to be announced are rigorously derived elsewhere (Grossberg and Pepe, 1970) and extend previous results for the case $\Gamma = 0$ (Grossberg, 1969*d*).

PSYCHOLOGICAL INTERPRETATION

The variables $x_i(t)$ and $z_{ij}(t)$ have psychological labels which are defined as follows. Let occurrence of the psychological event r_i at time $t = T$ create a brief input pulse at v_i with onset time $t = T$. Then $x_i(t)$ is called the *stimulus trace* of r_i at time t, and $z_{ij}(t)$ is the *associational strength* of the behavioral transition $r_i \to r_j$ at time t. Our results discuss the functions

$$(3) \qquad y_{ij}(t) = z_{ij}(t)[\textstyle\sum_{k \neq i} z_{ik}(t)]^{-1},$$

which measure the strength of the association $r_i \to r_j$ relative to the strength of all competing associations $r_i \to r_k$, $i \neq k \neq j$, through time. $y_{ij}(t)$ is thus a measure of the distinguishability of the association $r_i \to r_j$ during recall trials; strong competing associations $r_i \to r_k$ can annihilate behavioral effects of $r_i \to r_j$ via lateral inhibition if $y_{ij}(t)$ is too small (Grossberg, 1969*b*, 1970*b*).

A network closely related to Equations 1 and 2 will be exposed below to a single trial of a serial learning experiment using the list $L = r_1 r_2 \cdots r_{L-1} r_L$ of behavioral events. Cumulative effects of successive trials have been discussed (Grossberg, 1969*d*). Suppose that r_1 is presented at $t = 0$, r_2 is presented at $t = \tau, \cdots$, and r_L is presented at $t = (L - 1)\tau$. This paradigm defines inputs $I_i(t)$ to v_i which satisfy

$$(4) \quad I_i(t) = \begin{cases} I_{i+1}(t + \tau), \\ \quad \text{if } i = 1, 2, \cdots, L - 1, \\ 0, \text{if } i = L + 1, L + 2, \cdots, n. \end{cases}$$

To simplify notation, let $I_1(t) \equiv J(t)$, where $J(t)$ is positive throughout, and only in, the interval $(0, \lambda)$.

The effects of serial inputs on the relative associational strengths $y_{ij}(t)$ can be directly computed if higher-order nonlinear effects are ignored. The computation is carried out using Equations 2, 3, 4 and

$$(5) \qquad \dot{x}_i(t) = -\alpha x_i(t) + I_i(t)$$

instead of (1), $i = 1, 2, \cdots, n$. This system is called the *bare field* of the network. The computation also uses convenient restrictions on the parameters, which can be removed yielding obvious modifications; namely,

(*a*) The network is initially at rest and in a state of maximal ignorance; that is, all $x_i(t) = 0$ for $-\tau \leqq t \leqq 0$, and all $z_{ij}(0) = \epsilon > 0$, $i \neq j$;
(*b*) $\lambda < \tau$;
(*c*) $J(t)$ is continuous and has a single maximum;
(*d*) $\int_0^\lambda e^{-\alpha(\lambda - v)} J(v)\, dv > \Gamma$; and
(*e*) $\gamma = 0$: the decay rate of associations is small compared to the time scale of the transients to be studied.

ASSOCIATIONAL SPAN

Under these conditions, a precise study is possible of the associational span, primacy vs. recency, and bowing as a function of spiking threshold Γ. The associational span is defined heuristically as the maximum duration during which associations can be formed between a given r_i and other events r_k. Alternatively, it can be defined as the number of r_k with which r_i can form an association. We choose to use the following simple definition. Let $T_1 = \inf\{t : x_1(t) > \Gamma\}$ and $T_2 = \sup\{t : x_1(t) > \Gamma\}$. T_1 and T_2 are finite since by Equation 5 and Condition (*a*),

$$x_1(t) = \int_0^t e^{-\alpha(t-v)} J(v)\, dv,$$

and thus $x_1(\lambda) > \Gamma$ by Condition (*d*). In fact, $x_1(t) > \Gamma$ for $T_1 < t < T_2$ by Condition (*c*). By Equation 2 and Conditions (*a*) and (*e*),

$$z_{ij}(t) = \epsilon + \delta \int_0^t [x_i(v - \tau) - \Gamma]^+ x_j(v)\, dv,$$

which implies

$$z_{ij}(t) = \epsilon + \delta \int_0^t [x_1(v - i\tau) - \Gamma]^+ x_1(v - (j-1)\tau)\, dv$$

for $1 \leq i, j \leq L$, by Equation 4 and Condition (*a*). Thus $z_{ij}(t)$ can grow only during times when $x_1(t - i\tau) > \Gamma$, or only if $T_1 + i\tau < t < T_2 + i\tau$. The interval $(T_1 + i\tau, T_2 + i\tau)$ is therefore called the *associational interval* of r_i, and $S = T_2 - T_1$ is called the *associational span* of r_i.

Note that if $J(t)$ is a rectangular input pulse of intensity J, then

$$S = \lambda + \frac{1}{\alpha} \log \left[\left(\frac{J}{\alpha\Gamma} - 1\right) \cdot (1 - e^{-\alpha\lambda})^{-1}\right], \quad (5)$$

which is monotone decreasing in Γ. The interpretation of this fact is not trivial. For Γ too small, S becomes so large that response interference greatly diminishes the relative strength of correct associations $r_i \rightarrow r_{i+1}$. For Γ too large, however, even though no associations of the form $r_i \rightarrow r_{i+k}$, $k > 1$, can compete with $r_i \rightarrow r_{i+1}$, $[x_i(t) - \Gamma]^+$ is usually zero or small in value, so little learning occurs. Thus there exists an optimal region of threshold choice that reduces response interference and supplies enough energy to form the correct association. Notice in Equation 5 that decreasing J has the same qualitative effects as increasing Γ. The interplay between J and inhibitory interaction strength is, by contrast, often far more subtle (Grossberg, 1970*b*).

BOWING

Bowing cannot occur for all choices of Γ. For example, again choose Γ so large that $[x_i(t) - \Gamma]^+ = 0$ whenever $x_j(t) > 0$ and $j > i + 1$. Then no future associations $r_i \rightarrow r_{i+k}$, $k > 1$, can ever form. Consequently as the list position increases, the major effect on the association $r_i \rightarrow r_{i+1}$ is to increase response interference due to increasing numbers of backward response alternatives. Apart from this degenerate case, however, bowing always occurs in the bare field.

Rigorously expressed, bowing is a property of the function $B(i, \Gamma) \equiv \lim_{t\rightarrow\infty} y_{i,i+1}(t)$, $i = 1, 2, \cdots, L$. Grossberg and Pepe (1970) prove that for any fixed $\Gamma \geq 0$, $B(i, \Gamma)$ either first decreases and then increases as i increases from 1 to L, or the degenerate case occurs in which $B(i, \Gamma)$ is monotone decreasing. By definition, for fixed Γ, the bow occurs at the list position $K(\Gamma)$ for which $B(i, \Gamma)$ is a minimum. In the bare field, $K(\Gamma)$ is a monotone increasing function of Γ. Furthermore, $K(0) = \frac{1}{2}(L - 1)$ if L is odd and $K(0) = \frac{1}{2}L$ if L is even (Grossberg, 1969*d*). In the degenerate case above, $K(\Gamma) = L$ for sufficiently large Γ. Thus maximal difficulty in learning can occur at any list position greater than the list's numerical middle. Since "normal" learning requires a positive Γ, the bow will occur nearer to the end than to the beginning of the list, and the bowed curve will therefore be skewed. This also occurs in vivo (Grossberg, 1969*d*).

The above remarks describe, strictly speaking, "asymptotic" bowing, since $t = \infty$. Letting $B(i, \Gamma, t) \equiv y_{i,i+1}(t)$, suppose $\min_i B(i, \Gamma, t)$ occurs at list position $K(t, \Gamma)$ for every fixed t and Γ. It can be shown that, for every fixed $\Gamma \geq 0$, $K(t, \Gamma)$ ultimately decreases from $K(t, \Gamma) = L$ to $K(t, \Gamma) = K(\Gamma)$ as t increases from the time at which r_L is presented to infinity. This happens because the *non*-occurrence of the events r_{L+1}, $r_{L+2}, \cdots, r_n$ gradually decreases response interference at the end of the list. Thus skewing depends both on Γ and on the intertrial interval (Grossberg, 1969*d*).

PRIMACY VS. RECENCY

The function $f_{1L}(\Gamma) \equiv y_{12}(\infty) y_{L-1,L}^{-1}(\infty)$ measures the relative dominance of primacy over recency at large times t. In the bare field, $f_{1L}(0) < 1$ (recency dominates primacy), $f_{1L}(\Gamma)$ is monotone increasing in $\Gamma \geq 0$, and there exists a critical threshold

$\Gamma = \Gamma_0 < \infty$ such that $f_{1L}(\Gamma) > 1$ if $\Gamma > \Gamma_0$ (primacy dominates recency). Thus as Γ decreases, all but the most recent inputs are lost in background noise as the serial presentation of events proceeds. Consequently, the network cannot "pay attention" to long behavioral dependencies. First order associational connections therefore dominate the behavioral record, as is suggested also in schizophrenic punning.

CONCLUSION

Decreasing spiking threshold increases associational span, thereby passing from a region of slow learning to good learning, and finally to massive response interference and difficulties in "paying attention". Simultaneously, asymptotic learning difficulty moves from the end toward the middle of the list, and primacy eventually loses its battle with recency. Grossberg (1969*e*) discusses related mechanisms for "paying attention" that use explicit inhibitory interactions. Grossberg (1969*d*) also discusses other serial effects in the case $\Gamma = 0$, such as whether the bowed curve is raised or lowered in specific situations.

The dependence of learning phenomena on spiking threshold illustrates that local changes in membrane properties—say due to changes in Ca^{++} binding as a result of massive nutritional deficiencies (Vitamin *D*) (Schacter, 1969)—can in principle yield profound behavioral alterations. Entirely different causes can yield similar behavioral effects. As in Equation 5, creating a persistent source of diffuse overarousal can yield effects similar to reducing thresholds. For example, traumatic behavioral experiences might create such an overarousal by being conditioned to control cells which subserve diffuse anxiety. Effective treatment of these two problems might be quite different, however, since overarousal is not the source of the problem in the latter case. The fact that learning and performance can be profoundly altered by local chemical changes is compatible with the principles of orthomolecular psychiatry (Pauling, 1968).

The above model is clearly a highly idealized representation of possible neural events. Moreover, serial experiments on suitable mentally ill patients might be very hard to perform. Nonetheless, significant challenges can be approached by comparing the general learning behavior of normal and abnormal subjects as it depends on threshold size and inhibitory strength: help abnormal subjects by isolating more of the parameters that cause them learning difficulties; sharpen general categories for pooling serial data by attending to underlying physiological differences; and provide more indirect psychological information concerning whether or not transmitter production is jointly dependent on pre- and post-synaptic influences.

REFERENCES

Grossberg, S., *J. Math. Psych.*, 6 (2), 209 (1969a).
Grossberg, S., *Math. Biosci.*, 4, 255 (1969b).
Grossberg, S., *J. Theoret. Biol.*, 22, 327 (1969c).
Grossberg, S., *Math. Biosci.*, 4, 201 (1969d).
Grossberg, S., *J. Math. and Mech.*, 19(1), 53 (1969e).
Grossberg, S., *J. of Stat. Physics*, 1 (2), 319 (1969f).
Grossberg, S., Some Networks That Can Learn, Remember, and Reproduce Any Number of Complicated Space-Time Patterns (II), *Studies in Applied Math.* (June, 1970a) in press.
Grossberg, S., On Neural Pattern Discrimination, *J. Theoret. Biol.* (1970b) in press.
Grossberg, S., and Pepe, J., Overarousal and Threshold Phenomena in Serial Learning, *J. Stat. Physics*, submitted for publication (1970).
Kornetsky, C., and Eliasson, M., *Science*, 165 (3899), 1273 (1969).
Maher, B. A., *Psychology Today*, 2 (6), 30 (1968).
Pauling, L., *Science*, 160 (3825), 265 (1968).
Schacter, D. in *Biological membranes*, R. M. Dowben, Ed. Boston: Little, Brown, & Co. 1969.

(Manuscript received December 10, 1969)

I am content to see what is, and have never troubled myself to think what ought to be. One of the most formidable obstacles to the real progress of knowledge is this insane rage for presuming, and proceeding to decide upon presumption. It is ridiculous that we, with so limited a knowledge, should pretend to determine the laws of nature.

CHARLES GEORGES LEROY, 1736