Psychological Review

VOLUME 87 NUMBER 1 JANUARY 1980

How Does a Brain Build a Cognitive Code?

Stephen Grossberg
Department of Mathematics, Boston University

This article indicates how competition between afferent data and learned feedback expectancies can stabilize a developing code by buffering committed populations of detectors against continual erosion by new environmental demands. The gating phenomena that result lead to dynamically maintained critical periods, and to attentional phenomena such as overshadowing in the adult. The functional unit of cognitive coding is suggested to be an adaptive resonance, or amplification and prolongation of neural activity, that occurs when afferent data and efferent expectancies reach consensus through a matching process. The resonant state embodies the perceptual event, or attentional focus, and its amplified and sustained activities are capable of driving slow changes of long-term memory. Mismatch between afferent data and efferent expectancies yields a global suppression of activity and triggers a reset of short-term memory, as well as rapid parallel search and hypothesis testing for uncommitted cells. These mechanisms help to explain and predict, as manifestations of the unified theme of stable code development, positive and negative aftereffects, the McCollough effect, spatial frequency adaptation, monocular rivalry, binocular rivalry and hysteresis, pattern completion, and Gestalt switching; analgesia, partial reinforcement acquisition effect, conditioned reinforcers, underaroused versus overaroused depression; the contingent negative variation, P300, and pontogeniculo-occipital waves; olfactory coding, corticogeniculate feedback, matching of proprioceptive and terminal motor maps, and cerebral dominance. The psychophysiological mechanisms that unify these effects are inherently nonlinear and parallel and are inequivalent to the computer, probabilistic, and linear models currently in use.

How do internal representations of the environment develop through experience? How do these representations achieve an impressive measure of global self-consistency and stability despite the inability of individual nerve cells to discern the behavioral meaning of the representations? How are coding errors corrected, or adaptations to a changing environment effected, if individual nerve cells do not know that these errors or changes have occurred? This article describes how limitations in the types of information available to individual cells can be overcome when the cells act together in suitably designed feedback schemes. The designs that emerge have a natural neural interpretation, and enable us to explain and predict a large variety of psychological and physiological data as manifestations of mechanisms that have evolved

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

Requests for reprints should be sent to Stephen Grossberg, Department of Mathematics, Boston University, Boston, Massachusetts 02215.

to build stable internal representations of a changing environment. In particular, various phenomena that might appear idiosyncratic or counterintuitive when studied in isolation seem plausible and even inevitable when studied as a part of a design for stable coding.

Some of the themes that will arise in our discussion have a long history in psychology. To achieve an exposition of reasonable length, the article is built around a thought experiment that shows us in simple stages how cells can act together to achieve the stable self-organization of evironmentally sensitive codes. If nothing else, the thought experiment is an efficient expository device for sketching how organizational principles, mechanisms, and data are related from the viewpoint of code development, using a minimum of technical preliminaries. On a deeper level, the thought experiment provides hints for a future theory about the types of developmental events that can generate the neural structures in which the codes are formed. It does this by correlating the types of environmental pressures to which the developmental mechanisms are sensitive with the types of neural structures that have evolved to cope with these pressures. References to previous theories and data have been chosen to clarify the thought experiment, to contrast its results with alternative viewpoints, to highlight areas in which more experimentation can sharpen or disconfirm the theory, or to refer to more complete expositions that should be consulted for a thorough understanding of particular results. The thought experiment and its consequences do not, however, depend on these references, and the reader will surely know many other references that can be used to confront and interpret the thought experiment.

1. A Historical Watershed

Some of the themes that will arise were already adumbrated in the work of Helmholtz during the last half of the 19th century (Boring, 1950; Koenigsberger, 1906). Unfortunately, the conceptual and mathematical tools needed to cast these themes as rigorous science were not available until recently. This fact helped to precipitously terminate the productive interdisciplinary activity between physics and psychology that had existed until Helmholtz's time, as illustrated by the perceptual contributions of Mach and Maxwell (Boring, 1950; L. Campbell & Garnett, 1882; Ratliff, 1965) in addition to those of Helmholtz (1866, 1962); to create a schism between psychology and physics that has persisted to the present day; and to unleash a century of controversy and antitheoretical dogma within psychology that led Hilgard and Bower (1975) to write the following first sentence in their excellent review of *Theories of Learning:* "Psychology seems to be constantly in a state of ferment and change, if not of turmoil and revolution" (p. 2).

One illustrative type of psychological data that Helmholtz studied concerned color perception. Newton had noted that white light at a point in space is composed of light of all visible wavelengths in approximately equal measure. Helmholtz realized, however, that the light we perceive to be white tends to be the average color of a whole scene (Beck, 1972). Thus perception at each point is nonlocal; it is due to a psychological process that averages data from many points to define the perceived color at each point. Moreover, this averaging process must be nonlinear, since it is more concerned with relative than absolute light intensities. Unfortunately, most of the mathematical tools that were available to Helmholtz were local and linear.

There is a good evolutionary reason why the light that is perceived to be white tends to be the average color of a scene. We rarely see objects in perfectly white light. Thus our eyes need the ability to average away spurious coloration due to colored light sources, so that we can see the "real" colors of the objects themselves. In other words, we tend to see the "reflectances" of objects, or the relative amounts of light of each wavelength that they reflect, not the total amount of light reaching us from each point. This observation is still a topic of theoretical interest and is the starting point of the modern theory of lightness (Cornsweet, 1970; Grossberg, 1972a; Land 1977).

A more fundamental difficulty faced Helmholtz when he considered the objects of perception. Helmholtz was aware that cognitive factors can dramatically influence our

perceptions and that these factors can evolve or be learned through experience. He referred to all such factors as *unconscious inferences*, and developed his belief that a raw sensory datum, or *perzeption*, is modified by previous experience via a learned imaginal increment, or *vorstellung*, before it becomes a true perception, or *anschauung* (Boring, 1950). In more modern terms, sensory data activate a feedback process whereby a learned template, or expectancy, deforms the sensory data until a consensus is reached between what the data "are" and what we "expect" them to be. Only then do we "perceive" anything.

The struggle between raw data and learned expectations also has an evolutionary rationale. If perceptual and cognitive codes are defined by representations that are spread across many cells, with no single cell knowing the behavioral meaning of the code, then some buffering mechanism is needed to prevent previously established codes from being eroded by the flux of experience. It will be shown below how feedback expectancies establish such a buffer.

Unfortunately, Helmholtz was unable to theoretically represent the nonstationary, or evolutionary, process whereby the expectancy is learned, the feedback process whereby it is read out, or the competitive scheme whereby the afferent data and efferent expectancy struggle to achieve consensus. Helmholtz's conceptual and mathematical tools were linear, local, and stationary.

Section 4 begins to illustrate how nonlinear, nonlocal, and nonstationary concepts can be derived as principles of organization for adapting to a fluctuating environment. The presentation is nontechnical, but it will become apparent as we proceed that without a rigorous mathematical theory as a basis, the heuristic summary would have been impossible, since some of the properties that we will need are not intuitively obvious consequences of their underlying principles, and were derived by mathematical analysis. Furthermore, it will emerge that several design principles for adapting to different aspects of the environment operate together in the same structure. One of the facts that we must face about evolutionary systems is that their simple organizational principles can imply extraordinarily subtle properties. Indeed, part of the dilemma that many students of mind now face is not that they do not know enough facts on which to base a theory, but rather they do not know which facts are principles and which are epiphenomena, and how to derive the multitudinous consequences that occur when a few principles act together. A rigorous theory is indispensable for drawing such conclusions.

The next two sections summarize some familiar experiments whose properties will reappear from a deeper perspective in the thought experiment. These experiments are included to further review one of the themes that Helmholtz confronted, and to prepare the reader for the results of the thought experiment. The sections can be skipped on a first reading.

2. Overshadowing: A Multicomponent Adult Phenomenon With Developmental Implications

Psychological data are often hard to analyze because many processes are going on simultaneously in a given experiment. This point is illustrated below in a classical conditioning paradigm that will be clarified by the theoretical development. Classical conditioning is considered by many to be the most passive type of learning and to be hopelessly inadequate as a basis for cognitive studies. The overshadowing phenomenon illustrates the fact that even classical conditioning is often only one component of a multicomponent process in which attention, expectation, and other "higher order" feedback processes play an important role (Kamin, 1969; Trabasso & Bower, 1968; Wagner, 1969).

Consider the four experiments depicted in Figure 1. Experiment 1 summarizes the simplest form of classical conditioning. An unconditioned stimulus (UCS), such as shock, elicits an unconditioned response (UCR), such as fear, and autonomic signs of fear. The conditioned stimulus (CS), such as a briefly ringing bell, does not initially elicit fear, but after preceding the UCS by a suitable interval on sufficiently many conditioning trials, the CS does elicit a conditioned response (CR) that closely resembles the UCR. In this way, persistently pairing an indifferent cue with a

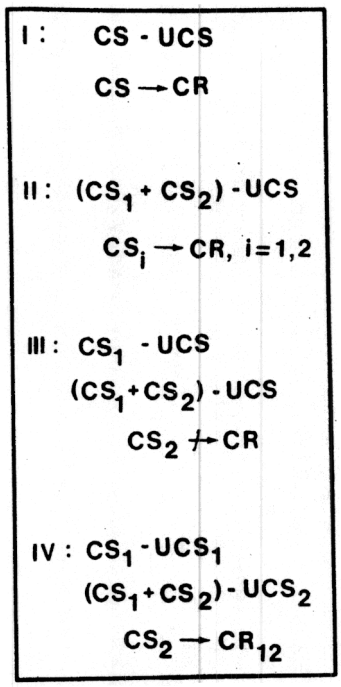

Figure 1. Four experiments illustrate overshadowing. (Experiment I summarizes the standard classical conditioning paradigm: conditioned stimulus–unconditioned stimulus [CS–UCS] pairing enables the CS to elicit a conditioned response (CR). Experiment II shows that joint pairing of two CSs with the UCS can enable each CS separately to elicit a CR. Experiment III shows that prior CS_1–UCS pairing can block later conditioning of CS_2 to the CR. Experiment IV shows that CS_2 can be conditioned if its UCS differs from the one used to condition CS_1. The CR that CS_2 elicits depends on the relationship between both UCSs, hence the notation CR_{12}.)

significant cue can impart some of the effects of the significant cue to the indifferent cue.

In Experiment 2, two CSs, CS_1 and CS_2, occur simultaneously before the UCS on a succession of conditioning trials; for example, a ringing bell and a flashing light both precede shock. It is typical in vivo for many cues to occur simultaneously, or in parallel, and the experimental question is, Is each cue separately conditioned to the fear reaction or is just the entire cue combination conditioned? If the cues are equally salient to the organism and are in other ways matched, then the answer is yes. If either cue CS_1 or CS_2 is presented separately after the conditioning trials, then it can elicit the CR.

Experiment 3 modifies Experiment 2 by performing the conditioning part of Experiment 1 on CS_1 before performing Experiment 2 on CS_1 and CS_2. In other words, first condition CS_1 until it can elicit the CR. Then present CS_1 and CS_2 simultaneously on many trials using the same UCS as was used to condition CS_1. Despite the results of Experiment 2, the CS_2 does not elicit the CR if it is presented after conditioning trials. Somehow prior pairing of CS_1 to the CR "blocks" conditioning of CS_2 to the CR.

The meaning of Experiment 3 is clarified by Experiment 4, which is the same as Experiment 3, with one exception. The UCS that follows CS_1 is not the same UCS that follows the stimulus pair CS_1 and CS_2 taken together. Denote the first UCS by UCS_1 and the second UCS by UCS_2. Suppose, for example, that UCS_1 and UCS_2 are different shock levels. Does CS_2 elicit a CR in this situation? The answer is yes if the two shock levels are sufficiently different. If the shock UCS_2 exceeds UCS_1 by a sufficient amount, then CS_2 elicits fear, or a negative reaction. If, however, the shock level UCS_1 exceeds UCS_2 by a sufficient amount, then CS_2 elicits relief, or a positive reaction.

How can the difference between Experiments 3 and 4 be summarized? In Experiment 3, CS_2 is an irrelevant or uninformative cue, since adding it to CS_1 does not change the expected consequence UCS. In Experiment 4, by contrast, CS_2 is informative because it predicts a change in the UCS. If the change is for the worse, then CS_2 eventually elicits a negative reaction (Bloomfield, 1969). If the change is for the better, then CS_2 eventually elicits a positive reaction (Denny, 1970).

Thus many learners are minimal adaptive predictors. If a given set of cues is followed by expected consequences, then all other cues are treated as irrelevant, as is CS_2 in Experiment 3. Each of us can define a given object using different sets of cues without ever realizing that our private sets are different, so long as the situations in which each of us uses the object always yield expected consequences. By contrast, if unexpected consequences occur, as in Experiment 4, then we somehow enlarge the set of relevant cues to include cues that were erroneously disregarded.

Several important qualitative conclusions can be drawn from these remarks. First, what is conditioned depends on our expectations, and these in turn help to regulate the cues to which we pay attention. Second, cues are conditioned, and indeed codes that interrelate these cues are built up, only if we pay attention to these cues because of their potential informativeness. Third, the mismatch between expected consequences and real events occurs only after attention has been focused on certain cues that thereupon generate the expectancy. Somehow this mismatch "feeds backwards in time" to amplify cues that have previously been overshadowed but that must have contained relevant information that we have erroneously ignored. Fourth, whenever we are faced with unexpected consequences, we do not know which cues have erroneously been ignored. The feedback process must be capable of amplifying all of the cues that are still being stored, albeit in a suppressed state. In other words, the feedback process is nonspecific. Finally, the nonspecific feedback process that is elicited by unexpected events competes with the specific consummatory channels that have focused our attention on the wrong set of cues. This competition between specific and nonspecific mechanisms helps us to reorganize our attentional focus until expected consequences are once again achieved.

This brief discussion reveals several basic processes working together in the overshadowing paradigm:

(a) classical conditioning, (b) attention, (c) learned expectancies, (d) matching between expectancies and sensory data, and (e) a nonspecific system that is activated by unexpected or novel events and competes with the specific consummatory system that focuses attention on prescribed cues.

Thus even classical conditioning is not a passive process when it occurs in realistic behavioral situations. Furthermore, its understanding requires the analysis of such teleological concepts as expectancy and attention. Helmholtz's doctrine of unconscious inference is readily called to mind.

Attention is to many individuals a holistic, if not unscientific, concept that does not mesh well with recent technological advances, say in microelectrode recording from individual nerve cells. Perhaps for this reason the fact that attentional variables can significantly influence what codes will be learned seems to have been ignored by some neurophysiologists who study the development of the visual cortex. For example, Stryker and Sherk (1975) were unable to replicate the Blakemore and Cooper (1970) study of visual code development in kittens. In the Blakemore and Cooper study, kittens were raised in a cylindrical chamber whose walls were painted with vertical black and white bars. The visual cortices of the kittens were reported to possess abnormally small numbers of horizontally tuned feature detectors. Hirsch and Spinelli (1970) performed experiments that did replicate in later experiments. In their experiments, the cats wore goggles, one lens with vertical stripes and the other with horizontal stripes. The corresponding visual cortices were reported to possess abnormally small numbers of feature detectors that were tuned to the orthogonal orientation. The entire controversy focused on such technical details as possible sampling errors due to Blakemore and Cooper's method of placing their electrodes. It is obvious, however, that the two experimental paradigms are attentionally inequivalent. Even perfect experimental technique would not necessarily imply similar experimental results.

3. Parallel Processing and the Persistence of Learned Meanings

The fact that classical conditioning, and for that matter any form of code development or learning, cannot be divorced from feedback processes that are related to attention is also made clear by the example illustrated by Figure 2. In Figure 2a, two classical conditioning experiments are depicted, one in which stimulus S_2 is the UCS for response R_2 and S_1 is its CS, and one in which S_1 is the UCS for R_1 and S_2 is its CS. What would happen if each cue S_1 and S_2 is conditioned to its own response R_1 or R_2, respectively, before a classical conditioning experiment occurs in which S_1 and S_2 are alternately scanned? This is the typical situation in real life, when we scan many cues in parallel, or intermittently, and many of these cues already have their own associations. If classical conditioning were a

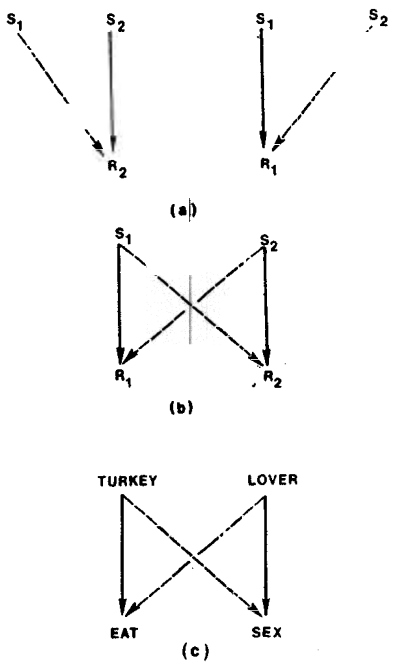

Figure 2. Classical conditioning cannot be a passive feed-forward process during real behavior. (In (a), S_1 acts as a conditioned stimulus (CS) for S_2, whereas S_2 acts as a CS for S_1. In (b), parallel processing of S_1 and S_2, each previously conditioned to responses R_1 and R_2, would yield cross-conditioning. In (c), some of the disastrous consequences of cross-conditioning are illustrated.)

passive feed-forward process, then cross-conditioning from S_1 to R_2 and from S_2 to R_1 would rapidly occur, as in Figure 2b.

However, this is absurd, as the particular example in Figure 2c vividly illustrates. Figure 2c schematizes the situation that would occur due to having a turkey dinner with one's lover. One alternately looks at lover and turkey, with lover associated with sexual responses (among others!) and turkey associated with eating responses. Why do we not come away from dinner wanting to eat our lover and to have sex with turkeys? Somehow the persistence of learned meanings can endure despite the fact that cues that are processed in parallel often generate incompatible responses. This is not always true, however, since if we, for example, consistently use a turkey as a discriminative cue for shock, or even sex, then turkeys might well become associated with fear or sexual arousal. Figure 2 depicts a situation in which the free reorganization of attention, rather than a forced pairing of a CS with a UCS, maintains the learned persistence of meanings. Grossberg (1975) developed a thought experiment in which overcoming the environmentally imposed dilemma of Figure 2 leads to attentional mechanisms that imply the overshadowing phenomena in Figure 1.

Before leaving the subject of overshadowing, we might ask why this adult attentional phenomenon is related to the development of sensory and cognitive codes, even in infants. This article argues that feedback is necessary to stabilize the development of behaviorally meaningful codes in a rich input environment. The feedback processes include attentional mechanisms, and the stabilization of developing codes leads to gating phenomena, or the emergence of critical periods, that are dynamically maintained by the feedback processes.

From this perspective, the structure of an environmentally adaptive tissue is a dynamic scheme whose parameters change very slowly only because of the nature of its maintaining feedback. Death itself is a dramatic example of how seemingly persistent structures can rapidly disintegrate when maintaining feedback is disturbed. When the development of a structure is driven by a particular type of experience, one of the structure's maintaining factors is that variety of experience. A subtle feature of such a developing structure is its ability to selectively amplify those experiences that tend to maintain its structure. Next I

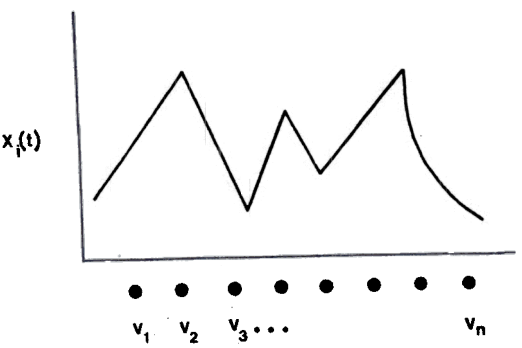

Figure 3. Each cell (or cell population) v_i possess an activity or potential $x_i(t)$, at every time t, $i = 1, 2, \ldots, n$. (The vector $(x_1(t), x_2(t), \ldots, x_n(t))$ of all these activities is a spatial pattern of activity.)

will discuss how feedback expectancies help to accomplish this end.

4. A Thought Experiment: The Need for Learned Feedback Expectancies

We now start to build a framework in which to discuss environmentally driven and behaviorally meaningful code development. Wherever possible, mathematical details will be suppressed, and the *minimal* structure capable of achieving our ends will be defined. This procedure will clarify what mathematical problems have to be solved, what their relationship is to each other, and what types of thematic variations on the minimal structures can be anticipated in different species and different neural locations in the same individual.

The central theoretical theme will be, How can a coding error be corrected if no individual cell knows that one has occurred? The importance of this issue becomes clear when we realize that erroneous cues can accidentally be incorporated into a code when our interactions with the environment are simple and will only become evident when our environmental expectations become more demanding. Even if our code perfectly matched a given environment, we would certainly make errors as the environment itself fluctuates. Furthermore, we never have an absolute criterion of whether our understanding of a fixed environment is faulty, or the environment that we thought we understood is no longer the same. The problem of error correction is fundamental whenever either the environment fluctuates or the individual keeps testing ever-deepening interpretations of the environment using ever sharper criteria of behavioral success.

We begin by introducing the functional elements on which our argument will build. Figure 3 depicts a collection of cells or cell populations, $v_1, v_2, \ldots, v_n$, each of which has an activity, or potential, $x_1(t), x_2(t), \ldots, x_n(t)$ at every time t. The activity $x_i(t)$ or v_i is imagined to be due to inputs $I_i(t)$ to v_i from a prior stage of neural processing, or the external environment, or endogenous sources within v_i itself. At every time t, these activities form a *pattern* $x(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ across the cells $v_1, v_2, \ldots, v_n$, to which we will refer collectively as a *field* of cells F. Henceforth, the time variable t will often be suppressed, since we will always take for granted that we are studying the system at a prescribed time.

Now consider two successive fields $F^{(1)}$ and $F^{(2)}$ of cells. Suppose that a pattern $x^{(1)}$ is active across $F^{(1)}$ (Figure 4). At this point the reader might wish to give $F^{(1)}$ and $F^{(2)}$ a concrete interpretation to help fix ideas. For example, one might think of $F^{(1)}$ as an idealization of the lateral geniculate nucleus (LGN) and $F^{(2)}$ as an idealization of visual cortex. The LGN processes visual data on its way to visual cortex, and it is the way station closest to the visual receptors at which our argument might hold in some species. I emphasize, however, that the results will be generally applicable to all neural stages at which behaviorally meaningful environmental inputs can drive code development. The fact that a significant fraction of visual development seems to be genetically prewired in the geniculocortical pathways of higher mammals like the monkey (Hubel & Wiesel, 1977) will not weaken the general conclusions that we will reach, and in fact various predictions and recent data about LGN, among other structures, will emerge from the analysis.

Suppose that the signal-carrying pathways from $F^{(1)}$ to $F^{(2)}$ act to filter the pattern $x^{(1)}$, and that due to prior developmental experience, this filter "codes" pattern $x^{(1)}$ by eliciting pattern $x^{(2)}$ across $F^{(2)}$. Knowing the detailed structure of this code is unnecessary to make our argument. However, we must be able to show how signal pathways can act as a filter that can be tuned by experience. This is done in Appendix A.

Suppose after the system learns to code $x^{(1)}$ by $x^{(2)}$ that another pattern is presented to $F^{(1)}$ and is erroneously coded at $F^{(2)}$ by $x^{(2)}$. To describe this situation conveniently, I introduce some subscripts. Denote $x^{(1)}$ and $x^{(2)}$ by $x_1^{(1)}$ and $x_1^{(2)}$, respectively, and denote the erroneously coded pattern at $F^{(1)}$ by $x_2^{(1)}$. In Figure 5 we draw the pattern $x_2^{(2)}$ that codes $x_2^{(1)}$ to equal $x_1^{(2)}$. Equality is meant to imply functional equivalence rather than actual identity. We now ask the central question, How can this coding error be corrected if no individual cell knows that an error has occurred?

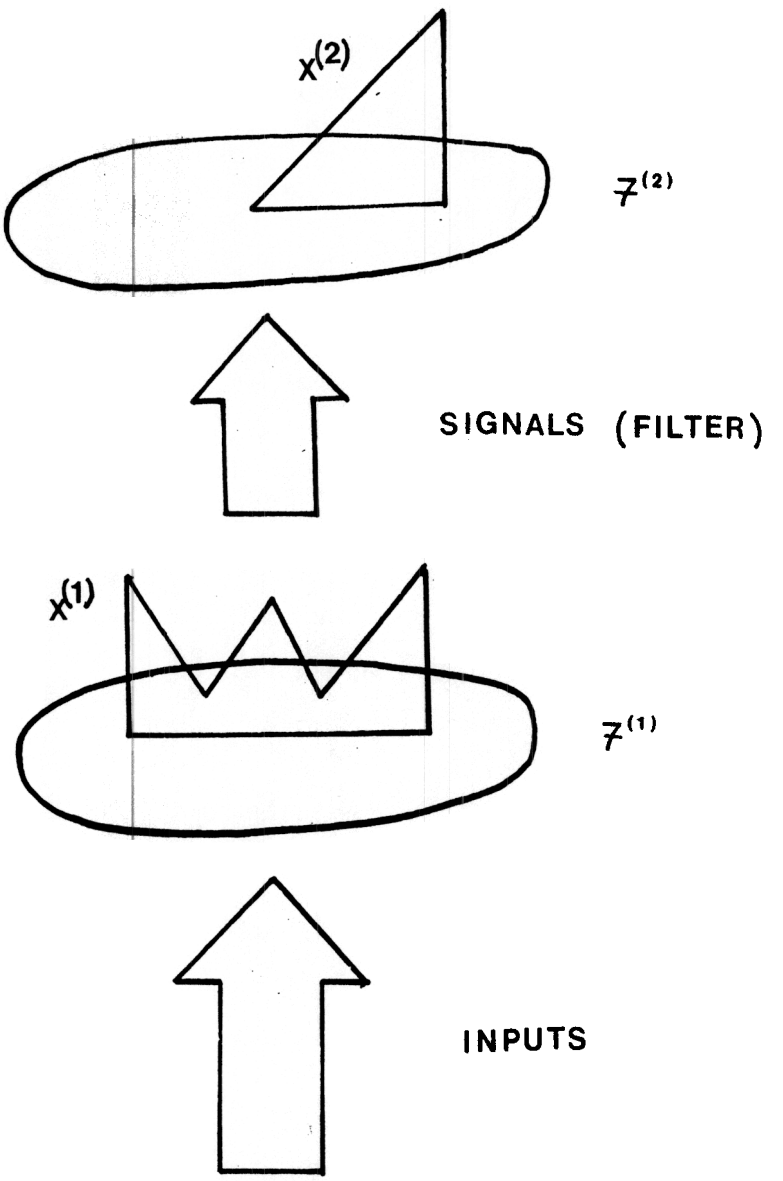

Figure 4. The activity pattern $x^{(1)}$ across $F^{(1)}$ is filtered to elicit a pattern $x^{(2)}$ across $F^{(2)}$.

Our first robust conclusion is now apparent: Whatever the mechanism is that corrects this error, it cannot exist within $F^{(2)}$, since by definition $x_1^{(2)}$ and $x_2^{(2)}$ are functionally equivalent. In principle, $F^{(2)}$ does not have the ability to distinguish the fact that $x_1^{(1)}$, and not $x_2^{(1)}$, should elicit $x_1^{(2)}$, since so far as $F^{(2)}$ knows, $x_1^{(1)}$ *is* active at $F^{(1)}$ rather than $x_2^{(1)}$.

It is important to realize that this argument is independent of coding details. It is based only on the type of information that $F^{(2)}$ cannot, in principle, possess. Much of our argument will be based on similar limitations in the types of information that particular processing stages can, in principle, possess. The robustness of this argument suggests why

the design that overcomes these limitations seems to occur ubiquitously, in one form or another, in so many neural structures.

Where in the network can this error be detected in principle? At the time when $x_2^{(1)}$ elicits $x_1^{(2)}$, there exists no trace within the network that during prior learning trials it was $x_1^{(1)}$ that elicited $x_1^{(2)}$, not $x_2^{(1)}$. Somehow this fact must be represented within the network dynamics. Otherwise, $x_1^{(2)}$ could become associated with $x_2^{(1)}$, just as $x_1^{(1)}$ was on previous developmental trials. The only times that $x_1^{(1)}$ was active in the network were the developmental trials during which the filter from $F^{(1)}$ to $F^{(2)}$ was learning to code $x_1^{(1)}$ by $x_1^{(2)}$. To be, in principle, capable of testing whether the correct pattern $x_1^{(1)}$ elicits $x_1^{(2)}$ on later trials when $x_1^{(1)}$ is not presented, it must be true that during the developmental trials, $x_1^{(2)}$ activates a feedback pathway from $F^{(2)}$ to $F^{(1)}$ that is capable of learning the active pattern $x_1^{(1)}$ at $F^{(1)}$. Then when $x_2^{(1)}$ erroneously activates $x_1^{(2)}$ on future trials, $x_1^{(2)}$ can read out the correct pattern $x_1^{(1)}$ across $F^{(1)}$. When this happens, the two patterns $x_1^{(1)}$ and $x_2^{(1)}$ will be simultaneously active across $F^{(1)}$, and they can be compared, or matched, to test whether or not the correct pattern has activated $x_1^{(2)}$ (Figure 6).

In summary, if in principle it is possible to correct a coding error at $F^{(2)}$, then there must exist learned feedback from $F^{(2)}$ to $F^{(1)}$. This learned feedback represents the pattern that $x_1^{(2)}$ *expects* to be at $F^{(1)}$ due to prior developmental trials. The feed-forward data to $F^{(1)}$ and the learned feedback expectancy, or template, from $F^{(2)}$ to $F^{(1)}$ are thereupon compared at $F^{(1)}$. Figure 7 illustrates this sequence of events as a series of snapshots that can occur at a very fast rate, for example, on the order of hundreds of milliseconds. Helmholtz's doctrine of unconscious inference is readily called to mind.

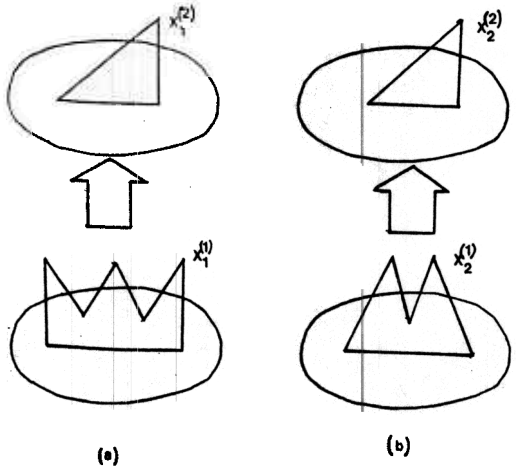

Figure 5. In (a), pattern $x_1^{(1)}$ at $F^{(1)}$ elicits the correct pattern $x_1^{(2)}$ across $F^{(2)}$. In (b), pattern $x_2^{(1)}$ elicits the incorrect pattern $x_2^{(2)}$, which is functionally equivalent to $x_1^{(2)}$ across $F^{(2)}$.

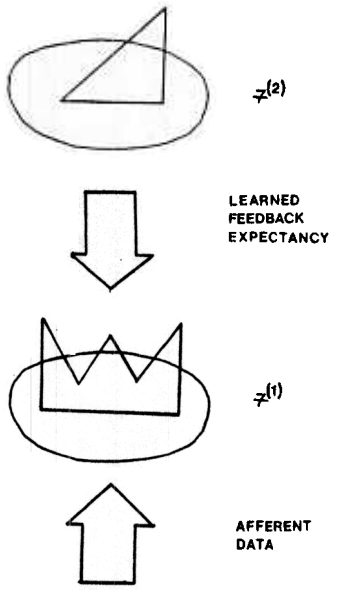

Figure 6. Pattern $x_1^{(2)}$ across $F^{(2)}$ elicits a feedback pattern $x_1^{(1)}$ to $F^{(1)}$, which is the pattern that it sampled across $F^{(1)}$ during previous developmental trials. (Field $F^{(1)}$ becomes an interface where afferent data and learned feedback expectancies are compared.)

The general nature of the preceding argument strongly suggests that feedback pathways will ubiquitously occur from "higher" neural centers to the relay stations that excite them. In fact, reciprocal thalamocortical connections seem to exist in all thalamo-neocortical systems (Macchi & Rinvik, 1976; Tsumoto, Creutzfeldt, & Legéndy, 1978).

At this point, we also recognize two more design problems for mathematics. The first problem is, How do feedback pathways from $F^{(2)}$ learn a pattern of activity across $F^{(1)}$? (See Appendix B for a summary of this mechanism.)

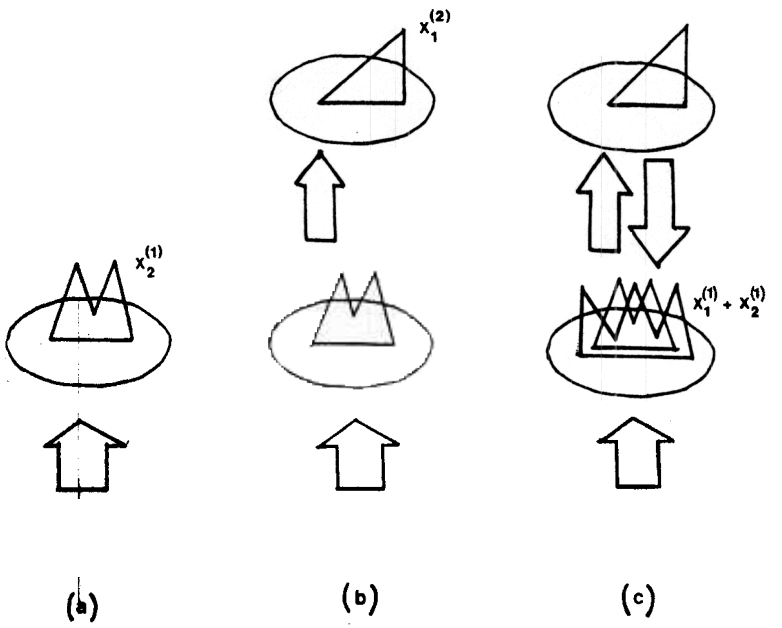

Figure 7. The stages (a), (b), and (c) schematize the rapid sequence of events whereby afferent data is filtered and activates a feedback expectancy that is matched against itself.

5. Noise Suppression, Pattern Matching, and Spatial Frequency Detection

The second design problem that we must face is this: Somehow the mismatch between the patterns $x_1^{(1)}$ and $x_2^{(1)}$ must rapidly shut off activity across $F^{(1)}$. Otherwise, $x_1^{(2)}$ would learn to code $x_2^{(1)}$ much as $x_1^{(2)}$ learned to code $x_1^{(1)}$ on the preceding developmental trials. Pattern $x_1^{(2)}$ must also be rapidly shut off if only to prevent behavioral consequences of $x_1^{(2)}$ from being triggered by further network processing. Moreover, $x_1^{(2)}$ must be shut off in such a fashion that $x_2^{(1)}$ can thereupon be coded by a more suitable pattern across $F^{(2)}$.

The only basis on which these changes can occur is the mismatch of $x_1^{(1)}$ and $x_2^{(1)}$ across $F^{(1)}$. We must therefore ask, How does the mismatch of patterns across a field $F^{(1)}$ of cells inhibit activity across $F^{(1)}$? The mathematical details are summarized in Appendix C. Here, however, it is useful to make the important distinction between mechanisms that develop due to evolutionary pressures and properties that are merely consequences of these mechanisms. One might well worry that the design of a mismatch mechanism is a rather sophisticated evolutionary task. We now indicate that such a mechanism is a consequence of a more basic property, namely noise suppression, and that noise suppression is itself a variation of a basic evolutionary principle. Moreover, other useful properties follow from noise suppression, such as spatial frequency detection and edge enhancement.

The environmental problem out of which the noise suppression property emerges is the *noise-saturation dilemma*. This dilemma has been discussed in detail elsewhere (e.g., Grossberg, 1977, 1978d). The dilemma confronts all noisy cellular systems that process input patterns, as in Figure 3. If the inputs are too small, they can get lost in the noise. If the inputs are amplified to avoid the noisy range, they can saturate all the cells by activating all of their excitable sites, and thereby reduce to zero the cells' sensitivity to differences in the input intensities. Appendix C reviews how competitive interactions among the cells automatically retune their sensitivity to overcome the saturation problem. In a neural context, the competitive interactions are said to be shunting interactions, and they are carried by an on-center off-surround anatomy. The retuning of sensitivity is due to automatic gain control by the inhibitory

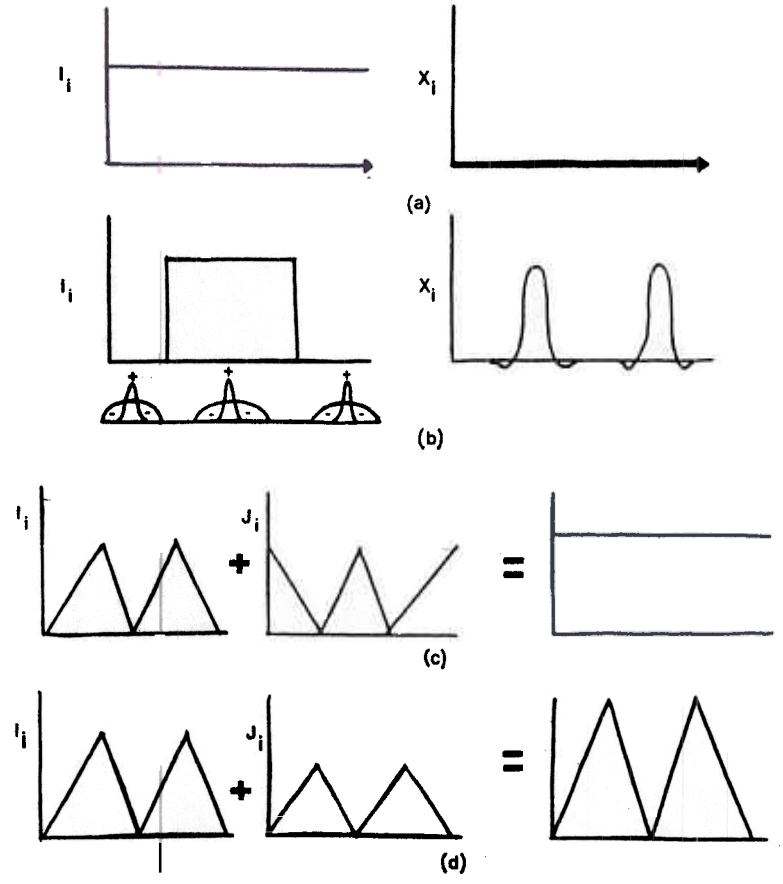

Figure 8. In (a), noise suppression converts a uniform input pattern into a zero activity pattern. In (b), a rectangular input pattern elicits differential activity at its edges because the cells within its interior and beyond its boundary perceive uniform fields. (This is a special case of spatial frequency detection.) In (c), two mismatched patterns add to generate an approximately uniform total input pattern, which will be suppressed by the mechanism of (a). In (d), two matched patterns add to yield a total input pattern that can elicit more vigorous activation than either input pattern taken separately.

off-surround signals. This fundamental property does not exist in additive models of lateral inhibition, such as the Hartline–Ratliff model (Ratliff, 1965). Appendix C shows how the automatic gain control mechanism can inhibit a uniform pattern of inputs, no matter how intense the inputs are. This is the property of noise suppression that we seek.

Figure 8a depicts this noise suppression property. A uniform pattern does not distinguish any cell from any other cell. For example, when the cells are feature detectors of one kind or another, a uniform input pattern contains no information that can distinguish one feature from any other feature. Noise suppression eliminates this irrelevant activity and allows the network to focus on informative discriminations.

Once noise suppression is guaranteed, several consequences automatically follow. For example, Figure 8b shows that such a network responds to the edges of a rectangular input, or to spatial gradients in more general input patterns. This is because cells whose inhibitory surrounds fall outside the rectangle perceive a uniform field, and cells with inhibitory surrounds that are near the center of the rectangle also perceive a uniform field. Both types of cells suppress their inputs. Only cells near the edges of the rectangle do not perceive a uniform pattern. Consequently only the edges of the rectangle elicit large activation.

This argument tacitly supposes that the lateral inhibitory interactions affecting each cell have a prescribed spatial extent, and that the width of the rectangle exceeds this spatial scale. More generally, spatial gradients in an input pattern are matched against the spatial scale of each cell's excitatory and inhibitory interactions. Only those spatial gradients in the input pattern that are nonuniform with respect to the cell's interaction scales generate large activities. By varying the inhibitory scales across cells, one can tune different cells to respond to different spatial frequencies. Thus spatial frequency detectors are a natural consequence of noise suppression properties within cells having a prescribed inhibitory scale. Since all networks in which shunting inhibition occurs have such scales, the existence of spatial frequency detectors should come as no surprise and does not imply that neural networks are Fourier analyzers in the spatial domain (Robson, 1976). Indeed, Fourier analyzers are linear mechanisms. By contrast, shunting networks that are capable of short-term memory contain feedback pathways, and all such networks must be nonlinear to be stable (Grossberg, 1973, 1978d).

Finally, Figure 8c and 8d indicate how a noise suppression mechanism can accomplish pattern matching. Figure 8c supposes that two mismatched patterns feed into $F^{(1)}$, where they add before coupling into the shunting dynamics. Because of the mismatch, the peaks of I_i fill in the troughs of J_i. The total input pattern is approximately uniform and is consequently quenched as noise. By contrast, in Figure 8d, the two patterns match. Their peaks and troughs mutually reinforce each other, so the resultant activities can be amplified beyond the effect of just one pattern. In summary, mismatched input patterns quench activity, whereas matched patterns amplify activity across a field $F^{(1)}$ that is capable of noise suppression.

A subsidiary mathematical question is now evident: How uniform must a pattern be for it to be suppressed? Part of the answer is determined by the choice of structural parameters, such as the strength and spatial distribution of lateral inhibitory coefficients (Appendix C). However, the field $F^{(1)}$ can also be dynamically tuned, or sensitized, by fluctuations in the level of nonspecific arousal that perturbs it through time. An arousal increment can, for example, act by inhibiting the inhibitory interneurons of the network (Ellias & Grossberg, 1975; Grossberg, 1973, 1978e; Grossberg & Levine, 1975). Such a tuning mechanism can simultaneously alter the spatial frequency properties of the network by multiplicatively strengthening or weakening the inhibitory interactions of the cells (Barlow & Levick, 1969a, 1969b). Such mechanisms will arise in a natural fashion as our argument continues.

6. Triggering of Nonspecific Arousal by Unexpected Events

Having suppressed $x_2^{(1)}$ at $F^{(1)}$ due to mismatch with the feedback expectancy $x_1^{(1)}$, we must now use this suppression to inhibit $x_1^{(2)}$ at $F^{(2)}$, since the mismatch at $F^{(1)}$ is the only mechanism in the network that can, in principle, distinguish that an error has occurred at $F^{(2)}$. Moreover, until $x_1^{(2)}$ is quenched, it will continue to read out the template $x_1^{(1)}$ to $F^{(1)}$, which will prevent $x_1^{(2)}$ from eliciting a new signal to $F^{(2)}$.

We were led to the mismatch mechanism at $F^{(1)}$ by noting that $F^{(2)}$ could not discriminate whether an error had occurred. Now we note that $F^{(1)}$'s information is also limited At $F^{(1)}$ it cannot be discerned *which* pattern across $F^{(2)}$ caused the mismatch at $F^{(1)}$. It could have been any pattern whatsoever. All $F^{(1)}$ knows is that a mismatch has occurred. Whatever pattern across $F^{(2)}$ caused the mismatch must be inhibited. Consequently, a mismatch at $F^{(1)}$ must have a *nonspecific* effect on all of $F^{(2)}$, since any of the cells in $F^{(2)}$ might be one of the cells that must be inhibited.

We are therefore led to the following questions: How does mismatch and subsequent quenching of activity across $F^{(1)}$ elicit a nonspecific signal (arousal!) to $F^{(2)}$? Where does the activity that drives this nonspecific arousal pulse come from?

Before answering these questions, we should realize that we have been led to a familiar conclusion: Unexpected or novel events are arousing. (To forcefully remind yourself of this basic fact, test a friend's reaction by unexpectedly slamming your hand on a table).

Now we will consider how such arousal is initiated and how it contributes to attentional processing.

Where does the activity that drives the arousal come from, and why is it released when quenching of activity at $F^{(1)}$ occurs? There are two possible answers to the first part of the question, but only one of them survives closer inspection. The activity is either endogenous (internally and persistently generated) or the activity is elicited by the sensory input. If the activity were endogenous, then arousal would occur whenever $F^{(1)}$ was inactive, whether this inactivity was due to active quenching by mismatched feedback from $F^{(2)}$ or to the absence of sensory inputs. This leads to the unpleasant conclusion that $F^{(2)}$ would be tonically flooded with arousal whenever nothing interesting was happening at $F^{(1)}$ or $F^{(2)}$. Therefore, sensory inputs to $F^{(1)}$ bifurcate before they reach $F^{(1)}$. One pathway is *specific:* It delivers information about the sensory event $F^{(1)}$. The other pathway is *nonspecific:* It activates the arousal mechanism that is capable of nonspecifically influencing $F^{(2)}$. The idea that cues have both informative (specific) and arousal (nonspecific) functions has been empirically known at least since the work of Moruzzi and Magoun on the reticular formation (Hebb, 1955; Moruzzi & Magoun, 1949).

Given that the sensory inputs to $F^{(1)}$ also activate an arousal pathway, what prevents this pathway from being activated except when activity at $F^{(1)}$ is quenched? The answer is now clear: Activity at $F^{(1)}$ inhibits the arousal pathway, and quenching of this activity disinhibits the arousal pathway. Figure 9 schematizes the (very rapid) sequence of events to which we have been led. First, a sensory event elicits a pattern $x_2^{(1)}$ across $F^{(1)}$ as it begins to activate the arousal pathway $\mathcal{A}$. This activation at $\mathcal{A}$ is inhibited by activity from $F^{(1)}$. Simultaneously, pattern $x_2^{(1)}$ activates pathways to $F^{(2)}$ that act as a filter that erroneously activates $x_1^{(2)}$. Pattern $x_1^{(2)}$ reads out the learned feedback expectancy $x_1^{(1)}$ to $F^{(1)}$. Mismatch of $x_1^{(1)}$ and $x_2^{(1)}$ at $F^{(1)}$ quenches activity across $F^{(1)}$. The inhibitory signal from $F^{(1)}$ to $\mathcal{A}$ is also quenched, and the arousal pathway is disinhibited. A nonspecific arousal pulse is hereby unleashed on $F^{(2)}$.

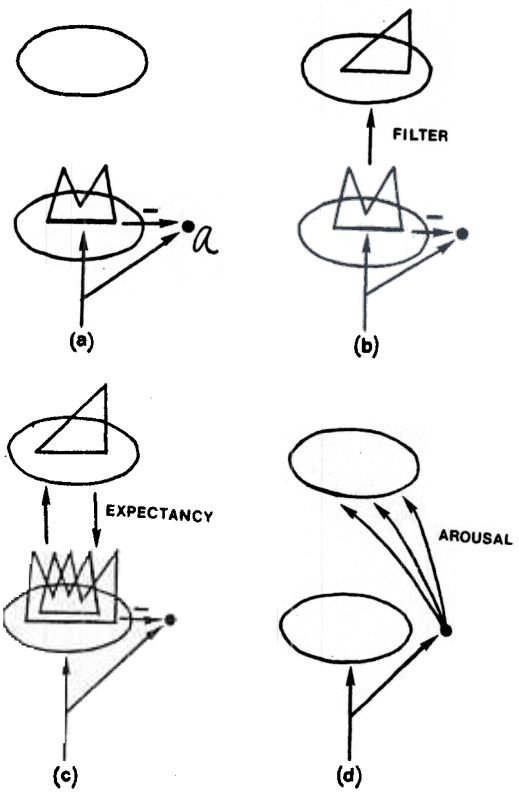

Figure 9. In (a), afferent data elicit activity across $F^{(1)}$ and an input to the arousal source $\mathcal{A}$ that is inhibited by $F^{(1)}$. In (b), the pattern at $F^{(1)}$ maintains inhibition of $\mathcal{A}$ as it is filtered and activates $F^{(2)}$. In (c), the feedback expectancy from $F^{(2)}$ is matched against the pattern at $F^{(1)}$. In d, mismatch attenuates activity across $F^{(1)}$ and thereby disinhibits $\mathcal{A}$, which releases a nonspecific arousal signal to $F^{(2)}$.

7. Parallel Hypothesis Testing in Real Time: The Probabilistic Logic of Complementary Categories

The next design problem is now clearly before us: How does the increment in nonspecific arousal differentially shut off the active cells in $F^{(2)}$? The active cells are the cells that elicited the feedback expectancy to $F^{(1)}$, and since mismatch occurred at $F^{(1)}$, these cells must have been erroneously activated. Consequently, they should be shut off. Furthermore, inactive cells at $F^{(2)}$ should not be inhibited, because these cells must be available for possible coding of $x_2^{(1)}$ during the next time interval. Thus a differential suppression of cells is required: The cells that are most active when arousal occurs should be most

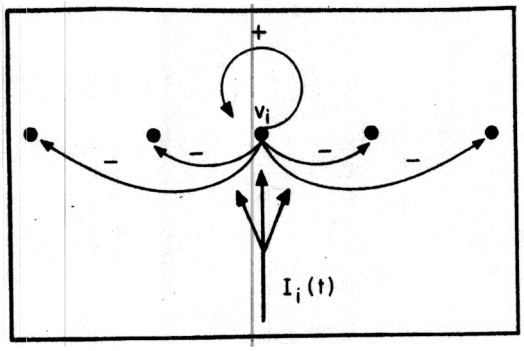

Figure 10. A recurrent shunting on-center off-surround network is capable of contrast-enhancing its input pattern, normalizing its total activity, and storing the contrast-enhanced pattern in short-term memory (STM). (If its feedback signals are properly chosen—e.g., sigmoid, or S-shaped signals—then a quenching threshold exists that defines the activity level below which activity is treated as noise and quenched, and above which activity is contrast enhanced and stored in STM.)

inhibited. This property realizes a kind of probabilistic logic in real time. If activating cell v_i in $F^{(2)}$ to a given degree leads to a certain degree of error or mismatch at $F^{(1)}$, then cell v_i should be inhibited to a degree that is commensurate both with its prior activation and with the size of the arousal increment, or the amount of error. If saying "yes" at v_i leads to error, then change the "yes" to "no," and do it in a graded fashion across the field $F^{(2)}$. Since cells that were only minimally active could have contributed only a small effect to the feedback expectancy, their inhibition will consequently be less, and they can contribute more to the correct coding of $x_2^{(1)}$ during the next time interval.

The arousal-initiated inhibition of cells across $F^{(2)}$ must be enduring as well as selective. Otherwise, as soon as $x_1^{(2)}$ is inhibited, the feedback expectancy, $x_1^{(1)}$, would be shut off, and $x_2^{(1)}$ would be free to reinstate $x_1^{(2)}$ across $F^{(2)}$ once again. The error would perseverate, and the network would be locked into an uncorrectable error. The inhibited cells must therefore stay inhibited long enough for $x_2^{(1)}$ to activate a different pattern across $F^{(2)}$ during the next time interval. The inhibition is therefore slowly varying compared to the time scale of filtering, feedback expectancy, and mismatch.

Once this selective and enduring inhibition is accomplished, the network has a capability for rapid hypothesis testing. By enduringly and selectively inhibiting $x_1^{(2)}$, the network "renormalizes" or "conditionalizes" the field $F^{(2)}$ to respond differently to pattern $x_2^{(1)}$ during the next time interval. If the next pattern elicited by $x_2^{(1)}$ across $F^{(2)}$ also creates a mismatch at $F^{(1)}$, then it will be suppressed, and $F^{(2)}$ will be renormalized again. In this fashion, a sequence of rapid pattern reverberations between $F^{(1)}$ and $F^{(2)}$ can successively conditionalize $F^{(2)}$ until either a match occurs or a set of uncommitted cells is found with which $x_2^{(1)}$ can build a learned filter from $F^{(1)}$ to $F^{(2)}$, and a learned expectancy from $F^{(2)}$ to $F^{(1)}$.

8. The Parallel Dynamics of Recurrent Competitive Networks: Contrast Enhancement, Normalization, Quenching Threshold, Tuning

At this point one can justifiably wonder how $x_2^{(1)}$ elicits a supraliminal pattern across $F^{(2)}$ after $x_1^{(2)}$ is inhibited? If $x_1^{(2)}$ is the pattern that $x_2^{(1)}$ originally excites, and $x_1^{(2)}$ is inhibited, then won't the next pattern elicited by $x_2^{(1)}$ across $F^{(2)}$ have very small activity? In other words, why was the second pattern not also active when $x_1^{(2)}$ was active?

It would have been if the anatomy within $F^{(2)}$ contained only feedforward, or nonrecurrent pathways. Thus we are forced to conclude that the anatomy within $F^{(2)}$ contains feedback, or recurrent pathways. Since all cellular systems face the noise-saturation dilemma, these pathways are distributed in a competitive geometry, or an on-center off-surround anatomy (Figure 10). Mathematical analysis demonstrates that the normalization property holds when recurrent pathways are distributed in a competitive geometry. When these competitive networks are designed to overcome noise amplification and saturation, they enjoy several properties that we need (Appendix D). First, they are capable of contrast-enhancing small differences in initial pattern activities into large and easily discriminable differences that are thereupon stored in short-term memory (STM; see

Figure 11). This property is necessary to build up the codes for $F^{(1)}$ patterns at $F^{(2)}$. Before an $F^{(1)}$ pattern is coded by $F^{(2)}$, it might elicit an almost uniform activity pattern across $F^{(2)}$. The recurrent dynamics within $F^{(2)}$ quickly contrast enhances and stores the contrast-enhanced pattern in STM, where it can be sampled and stored in long-term memory (LTM) by the pathways from $F^{(1)}$ to $F^{(2)}$. When the next occurrence of the same pattern at $F^{(1)}$ occurs, these pathways therefore elicit a more differentiated pattern across $F^{(2)}$, which is again contrast enhanced and stored in STM. The feedback enhancement between STM and LTM continues until the two processes equilibrate other things being equal.

Another property of such a network is its tendency to conserve, or adapt, the total activity that it stores in STM. This is the normalization property that we seek. If certain cells in the network are prevented from sharing the STM activity, say due to arousal-initiated inhibition, then the total activity is renormalized by being distributed to the other cells. Thus after $x_1^{(2)}$ is inhibited across $F^{(2)}$, the network will respond to the signals due to $x_2^{(1)}$ by differentially amplifying them in a way that tends to preserve the total STM activity across $F^{(2)}$. This new STM pattern will inherit much of the STM activity that $x_1^{(2)}$ had before it was suppressed, but the new STM pattern across $F^{(2)}$ will be a quite different pattern than $x_1^{(2)}$, since it is built from $F^{(1)}$ signals that previously fared poorly in the competition for STM activity. The normalization property manifests itself in a large class of psychological data, notably data about behavioral contrast and ratio scales in choice behavior (Grossberg, 1975, 1978a).

These recurrent networks also possess a *quenching threshold* (QT), which is a parameter whose size determines what activities will be suppressed, or quenched, and what activities will be stored in STM (Grossberg, 1973). Activities in populations that start below the QT will be suppressed; activities that exceed the QT will be contrast enhanced and stored in STM. Thus the QT is the cutoff point that defines noise in a recurrent network. All networks that possess a QT can be tuned; that is, by varying the QT, the criterion of

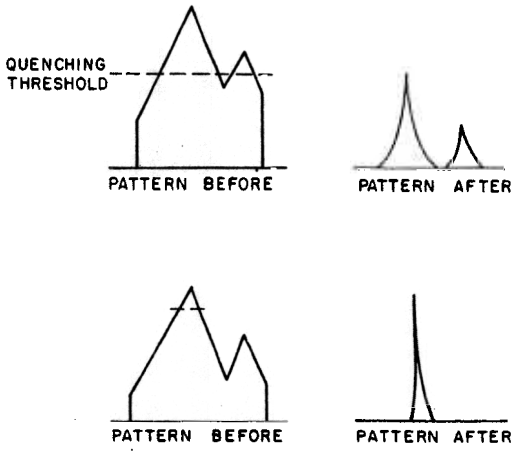

Figure 11. If the quenching threshold is variable, for example, due to shunting signals that nonspecifically control the size of the network's inhibitory feedback signals, then the network's sensitivity can be tuned to alter the ease with which inputs are stored in short-term memory.

which data shall be stored in STM and which data shall be quenched can be altered through time. Several parameters work together to determine QT size, notably the strength of recurrent lateral inhibitory pathways within the network. For example, if a nonspecific arousal pulse multiplicatively inhibits or shunts the inhibitory interneurons of a recurrent network, then its QT will momentarily decrease—the network's inhibitory "gates" will open—to facilitate STM storage.

The normalization property also helps us to understand the relevance of probabilistic models of hypothesis testing to cognitive processing. Normalization plays the role of summing all the probabilities to equal 1. Shunting, or multiplicative network dynamics, plays the role of multiplying the probabilities of independent events. However, probabilistic concepts only approximately describe some aspects of shunting competitive dynamics. A most serious difference is that although the network's hypothesis testing mechanism might produce a serial sequence of renormalizations in time, these operations are performed by parallel, rather than serial mechanisms. Serial mechanisms of hypothesis testing are not equivalent to the parallel theory.

More generally, serial behavioral properties do not imply that the control processes that

subserve them are also serial. In particular, various serial, notably computer, models of memory and cognitive processing have been shown to be fundamentally inequivalent to parallel neural interactions. This inequivalence is noted for the Atkinson and Shiffrin (1968, 1971) theory of free recall in Grossberg (1978a) and for the Schneider and Shiffrin (1976) theory of automatic versus controlled visual information processing in Grossberg (1978e, Section 61). Different predictions of the two types of theory are also described in these articles, and some data that are inexplicable by the serial theories are explained using parallel properties, such as normalization, in a basic way.

9. Antagonistic Rebound Within On-Cell Off-Cell Dipoles

We are now faced with a subtle design problem: How can a nonspecific event, such as arousal, have specific consequences of any kind, let alone generate an exquisitely graded, enduring, and selective suppression of active cells? Here again mathematical analysis was absolutely essential, since the theory could not progress beyond this step had not the answer already been derived (to my own surprise) during work on reinforcement mechanisms (Grossberg, 1972b, 1972c, 1975). In this work, the mechanism helped explain such nontrivial effects as learned helplessness, vicious circle behavior, superconditioning, overshadowing, asymptotically nonchalant avoidance, and peak shift with behavioral contrast when it was joined with suitable conditioning and cognitive mechanisms that were all derived as evolutionary solutions to prescribed environmental pressures. These results extend such popular learning theories as those of Irwin (1971), Kamin (1969), Rescorla and Wagner (1972), and Seligman, Maier, and Solomon (1971) by explicating mechanisms, conceptual distinctions, and predictions in a psychophysiological framework that are invisible to descriptive theories. One might wish to know what reinforcement mechanisms have to do with the development of cognitive codes. The answer is that the property in question occurs whenever optimally designed chemical transducers, or transmitters, occur in competing network channels, or dipoles, whether these channels arise in reinforcement mechanisms, attentional mechanisms (Grossberg, 1975), developmental mechanisms (Grossberg, 1976b), or mechanisms of motor control (Grossberg, 1978e). The property is a robust consequence of a ubiquitous neural design principle, and it guarantees a type of rapid hypothesis testing and error correction wherever this principle is used.

First let us consider some familiar behavioral facts that help to motivate the mechanism. Suppose that I wish to press a lever in response to the offset of a light. If light offset simply turned off the cells that code for light being on, then there would exist no cells whose activity could selectively elicit the lever-press response after the light was turned off. Clearly, offset of the light not only turns off the cells that are turned on by the light, but it also selectively turns on cells that will transiently be active after the light is shut off. The activity of these "off"-cells—namely the cells that are turned on by light offset—can then activate the motor commands leading to the lever press. Let us call the transient activation of the off-cell by cue offset *antagonistic rebound.*

Antagonistic rebound also occurs in a variety of other behavioral situations. For example, shock can unconditionally elicit the emotion of fear and various autonomic consequences of fear (Dunham, 1971; Estes, 1969; Estes & Skinner, 1941). Offset of shock is (other things equal) capable of eliciting relief or a complementary emotional reaction (Denny, 1970; Masterson, 1970; McAllister & McAllister, 1970). In a similar fashion I suggest that when motor command cells are organized in agonist–antagonist pairs, offset of the agonist input can elicit a rebound in the antagonist command cell that acts to rapidly brake the motion in the muscles controlled by the agonist command cell.

When such on-cell off-cell interactions are modeled, one finds examples akin to Figure 12. In Figure 12a, a nonspecific, or adaptation level, input I is delivered equally to both channels, whereas a test input J is delivered to the on-cell channel. These inputs create signals S_1 and S_2 in both channels, and the signals are multiplicatively gated by slowly varying chemical transmitters z_1 and z_2,

respectively. The gated signals S_1z_1 and S_2z_2 thereupon compete and yield the on-cell off-cell responses that are depicted in Figure 12a. Appendix E describes the details that are needed for a better understanding, but the main idea behind antagonistic rebound is easy to describe. Consider Figure 12a. Here the transmitters z_1 and z_2 are depleted by being released at rates proportional to S_1z_1 and S_2z_2, respectively. More depletion of z_1 than z_2 occurs if the signal S_1 exceeds S_2. While the test input J is on, the on-channel receives a larger input than the off-channel, since its total input is J plus the nonspecific input I, whereas the off-cell channel only receives the input I. Consequently, $S_1 > S_2$, so that depletion of transmitter leads to the inequality $z_1 < z_2$. Despite this fact, one can prove that the gated signals satisfy the inequality $S_1z_1 > S_2z_2$. Consequently, the on-channel receives a larger gated signal than the off-channel, so that after competition takes place, there is a net on-reaction.

What happens when the test input is shut off? Both channels receive only the equal nonspecific input I. The signals S_1 and S_2 rapidly equalize until $S_1 = S_2$. However, the transmitters are more slowly varying in time so that the inequality $z_1 < z_2$ continues to hold. The gated signals therefore satisfy $S_1z_1 < S_2z_2$. Now the off-channel receives a larger signal. After competition takes place there is an antagonistic rebound in response to offset of the test input.

Why is the rebound transient in time? The equal signals S_1 and S_2 continue to drive the depletion of the transmitters z_1 and z_2. Gradually the amounts of z_1 and z_2 also equalize so that S_1z_1 and S_2z_2 gradually equalize. As the gated signals equalize, the competition shuts off both the on-channel and the off-channel. These facts are summarized in Table 1.

Table 1
Antagonistic Rebound at Offset of Phasic Input

Test input J is on	Right after offset of J	After dipole equilibrates to offset of J
$I + J > I$	$I = I$	$I = I$
$x_1 > x_2$	$x_1 = x_2$	$x_1 = x_2$
$S_1 > S_2$	$S_1 = S_2$	$S_1 = S_2$
$z_1 < z_2$	$z_1 < z_2$	$z_1 = z_2$
$S_1z_1 > S_2z_2$	$S_1z_1 < S_2z_2$	$S_1z_1 = S_2z_2$
$x_3 > x_4$	$x_3 < x_4$	$x_3 = x_4$
$x_5 > 0 = x_6$	$x_5 = 0 < x_6$	$x_5 = 0 = x_6$

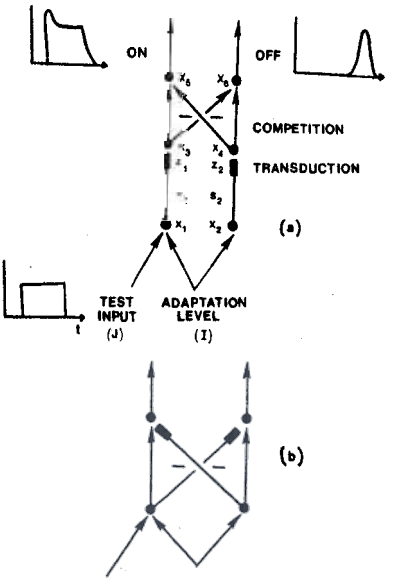

Figure 12. Two examples of on-cell off-cell dipoles. (In (a), the test input J and adaptation level input I add in the on-channel. The adaptation level input perturbs the off-channel. Each input is gated by a slowly varying excitatory transmitter [square synapses]. Then the channels compete before eliciting a net on-response or off-response. In (b), the slowly varying transmitters are inhibitory, and participate in the competition process.)

10. Analgesia, Escape, Partial Reward, and Underaroused Versus Overaroused Depression

In Figure 12a, the two transmitters are excitatory and generate gated signals before competition occurs. Similar effects occur in Figure 12b in which the transmitters are inhibitory and act both as gates and as competing channels. There exist many variations on this theme in vivo. For example, by analyzing more complex learning situations, in particular, experiments on secondary conditioning phenomena, or on transfer between instrumental and classical conditioning, one can show that feedback pathways must exist within the channels that subserve incentive

motivation. These feedback channels lead to meaningful comparisons with psychophysiological data when they are interpreted as a formal analogue of the medial forebrain bundle (Grossberg, 1972c, 1975).

Even the feed-forward networks already have surprising and important properties, however. For example, consider a network in which the on-channel supplies negative incentive motivation ("fear") and the off-channel supplies positive incentive motivation ("relief") in a conditioning paradigm. Choose shock reduction as the experimental manipulation. Let shock excite the on-channel, and suppose that the size of the positive rebound after shock terminates is monotonically related to the rewarding effect of the manipulation. Then one can derive a quantitative formula for rebound size (Grossberg, 1972c) that orders infinitely many possible experiments in terms of how rewarding they will be. In particular, reducing J units of shock to $J/2$ units is less rewarding than reducing $J/2$ units of shock to 0 units, despite the fact that shock reduction equals $J/2$ units in both cases. This analgesic effect is due to intracellular adaptation of the chemical transmitters. Analogous data have been reported by Campbell (1968); B. Campbell and Kraeling (1953); Gardner, Licklider, and Weisz (1961); and Myers (1969). Moreover, it is predicted that three indices should all covary as a function of the reticular formation arousal level, which is interpreted to be a source of nonspecific input to the incentive motivational dipoles. These indices are (a) the rewarding effect due to switching J units to $J/2$ units of shock, (b) the ability of an animal to learn to escape from presentation of a discrete fearful cue, and (c) the relative advantage of partial reward over continuous reward (Grossberg, 1972c).

One also finds that two types of depressed emotional affect exist in the dipole: an underaroused syndrome and an overaroused syndrome. These syndromes are manifestations of the dramatic changes in the net incentive motivation that occur when the arousal level is parametrically changed (Grossberg, 1972c). The two syndromes are the endpoints in an inverted U of net incentive as a function of arousal level. At underaroused levels, the behavioral threshold is abnormally high, but the system is hyperactive after this threshold is exceeded. At overaroused levels, the behavioral threshold is abnormally low, but the system is so hypoactive that little net incentive is ever generated. Parkinson's patients and certain hyperactive children seem to exhibit the underaroused syndrome (Fuxe & Ungerstedt, 1970; Ladisich, Volbehr, & Matussek, 1970; Ricklan, 1973), which is paradoxical because behavioral threshold is inversely related to suprathreshold reactivity. Such underaroused individuals can be brought "down" behaviorally by a drug that acts as an "up"; that is, it raises the adaptation level to the normal range. In Parkinson's patients, this up is L-dopa, and in certain hyperactive children, it is amphetamine.

A general question now presents itself: Do *all* neural dipoles share these properties whether they occur in motivational, sensory, or motor representations? This question is considered for the case of cortical red–green dipole responses to white light in Section 12.

11. Arousal Elicits Antagonistic Rebound: Surprise and Counterconditioning

A surprising feature of the on-cell off-cell dipole is its reaction to rapid temporal fluctuations in arousal, or adaptation level. This reaction allows us to answer the following question posed in Section 9: How can a nonspecific event, such as arousal, selectively suppress active on-cells? Appendix E shows that arousal fluctuations can reset the dipole, despite the fact that they generate equal inputs to the on-cell and off-cell channels. In particular, a sudden increment in arousal can, by itself, cause an antagonistic rebound in the relative activities of the dipole. Moreover, the size of the arousal increment that is needed to cause rebound can be independent of the size of the test input that is driving the on-channel. When this occurs, an arousal increment that is sufficiently large to rebound any dipole will be large enough to rebound all dipoles in a field. In other words, if the mismatch is "wrong" enough to trigger a large arousal increment, then all the errors will be simultaneously corrected. This cannot, in principle, happen in a serial processor. Moreover, the size of the rebound is an increasing

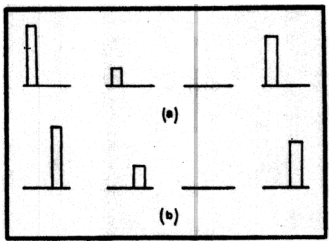

Figure 13. A rebound from on-cell activation to off-cell activation can be elicited by a rapid increment in the arousal or adaptation level of the dipole. (The size of the rebound is determined by the size of the on-cell activation. In (a) are depicted the on-responses of four cells. In (b) are depicted possible rebounds by their off-cells in response to a nonspecific increment.)

function of the size of the on-cell test input (Figure 13). Thus the amount of antagonistic rebound is precisely matched to the amount of on-cell activation that is to be inhibited. Finally, in previously inactive dipoles no rebound occurs, but the arousal increment can sensitize the dipole to future signals by changing by equal amounts the gain, or temporal averaging rate, of the on-cell and off-cell. In summary, the on-cell off-cell dipole is superbly designed to selectively reset $F^{(2)}$, and to do so in an enduring fashion because of the slow fluctuation rate of the transmitter gates.

In a reinforcement context, the rebound due to arousal shows how surprising or unexpected events can reverse net incentive motivation and thereby drive counterconditioning of a behavior's motivational support (Grossberg, 1972b, 1972c). Once the rebound capabilities of surprising events are recognized, one must evaluate with caution such general claims as "the surprising omission of . . . shock . . . can hardly act as a reinforcing event to produce excitatory conditioning" (Dickinson, Hall, & Mackintosh, 1976, p. 321).

The above mechanisms indicate how dynamical critical periods might be laid down by learned feedback expectancies. These expectancies modulate an arousal mechanism that buffers already coded populations by shutting them off so rapidly in response to erroneous STM coding that LTM recording is impossible. In other words, the mechanism helps to stabilize the LTM code against continual erosion by environmental fluctuations.

The thought experiments from which these conclusions follow are purely abstract. One experiment describes how limitations in the types of information available to individual cells can be overcome when the cells act together in suitably designed feedback schemes. Another experiment describes a solution to the noise-saturation dilemma, and yet another experiment describes how to design a chemical transducer and how dipoles formed when such transducers compete in parallel channels can achieve antagonistic rebound. As the thought experiments proceed, however, the resultant network designs take on increasingly neural interpretations. To test the theory by psychophysiological experiments, these empirical connections must be made more explicit. The next three sections discuss three of the major design features in more detail to suggest that some psychophysiological designs are examples of our abstract designs, and to explain and predict some psychophysiological phenomena using formal properties of the abstract designs as a guide. These examples are hardly exhaustive, but they will perhaps be sufficient to enable the reader to continue making new connections. Further details are in the articles of Grossberg (1972b, 1972c, 1975, 1976b, 1978e). The next three sections can be skipped on a first reading if the reader wishes to immediately study Section 15 to find out what happens when the patterns at $F^{(1)}$ and $F^{(2)}$ mutually reinforce each other.

12. Dipole Fields: Positive and Negative Aftereffects, Spatial Frequency Adaptation, Rivalry, and the McCollough Effect

Section 8 noted that $F^{(2)}$ possesses a recurrent on-center off-surround anatomy that is capable of normalizing its total STM activity within its functional channels. Section 9 showed that the cells in this recurrent anatomy are the on-cells of on-cell off-cell dipoles. I therefore conclude that $F^{(2)}$ consists of a field of on-cell off-cell dipoles such that the on-cells interact within a recurrent on-center off-surround anatomy and the off-cells also interact within a recurrent on-center off-surround anatomy. Denote by $F_+^{(2)}$ the recurrent subfield of on-cells, and by $F_-^{(2)}$ the recurrent subfield of off-cells (Figure 14). The

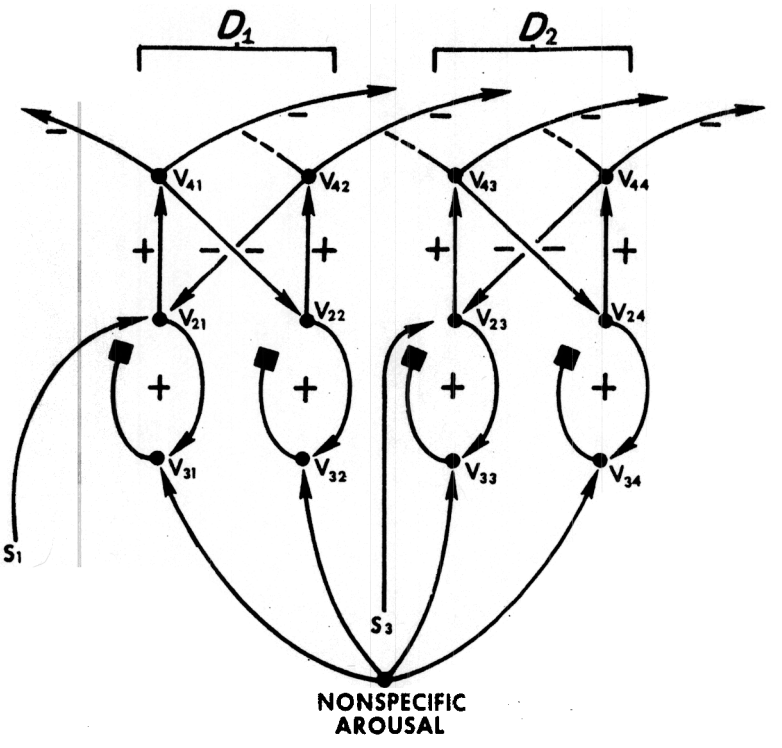

Figure 14. A possible anatomy of two dipoles (D_1 and D_2) is depicted, embedded in recurrent subfields of on-cells and off-cells. (The nonspecific arousal signal is gated by transmitters in the pathways $v_{2i} \rightarrow v_{2i}$, $i = 1, 2, \ldots$. The transmitter gates are depicted by square synapses. The arousal level hereby determines an overall level of transmitter adaptation across the dipole field. The signal S_1 turns on the cell v_{21}, which inhibits its off-cell v_{22} via the inhibitory interneuron v_{41}. Simultaneously, the on-cell v_{21} begins to differentially deplete its transmitter gate via the feedback pathway $v_{21} \rightarrow v_{31} \rightarrow v_{21}$. The interneurons v_{4i}, $i = 1, 2, \ldots$, also activate the recurrent interactions among on-cells and among off-cells that normalize their respective subfields.)

existence of neural, in particular, cortical on-cells and off-cells, and the joining together of nerve cells in on-center off-surround anatomies are familiar neural facts. Moreover, these facts have often been used to explain psychophysiological data (Carterette & Friedman, 1975; Cornsweet, 1970). The present treatment is novel in several respects, however. That a dipole field is a major tool to reset an error and to search for a correct code is, to the best of my knowledge, a new insight. Moreover, the way in which arousal fluctuations interact with slowly varying, competing transmitter gates to cause rebound or a shift in adaptation level, and the way in which shunting interactions define a quenching threshold, normalize field activity, and regulate contrast enhancement also seem to be new insights.

There exists a basic difference between the recurrent inhibition within a subfield and the dipole inhibition between on-cells and their off-cells. Dipole inhibition creates a balance between mutually exclusive categories or features. Intrafield inhibition normalizes and tunes its subfield. For example, suppose that the on-cells in a given field respond to white bars of prescribed orientation on a black field, and their corresponding off-cells respond to black bars of similar orientation on a white field. A continuous shift in the position of a white bar can induce a continuous shift of activity within the on-field, but at each position there can exist either a white bar on a black field or a black bar on a white field, but not both. Next are summarized some of the phenomena that are due to continuous changes within subfields and complementary changes

when dipole rebounds cause a flip between subfields. The goal of this summary is to clarify some of the properties through which dipole fields manifest themselves in perceptual data, and to suggest that these properties are manifestations of code stabilizing mechanisms. The summary will not attempt to describe the global schemata in which these properties are embedded during a live perceptual event, although the article makes clear that interfield signaling processes, such as filtering and expectancy matching, will be important ingredients in the classification of such schemata.

An important property of a dipole field is this: If a test input excites a particular on-cell, then the on-cell inhibits its off-cell. The inhibited off-cell can, in turn, disinhibit a nearby off-cell due to the tonic arousal input and the recurrent anatomy within the off-cell field. The disinhibited off-cell thereupon inhibits its on-cell via dipole interactions. Suppose that the test input is shut off after it has been on long enough to deplete its transmitter gate. (To make this argument quantitative, we must carefully control the duration of experimental inputs relative to the transmitter depletion rate.) Then antagonistic rebound within its dipole can turn on its off-cell, which inhibits the nearby off-cell, whose on-cell is hereby disinhibited and responds by rebounding onward. Negative aftereffects are hereby generated. For example, suppose that the on-cells are orientationally selective such that nearby orientations recurrently excite each other, whereas more distinct orientations inhibit each other (Figure 15a). Then persistent inspection of a field with radial symmetry (Figure 15a) can elicit an aftereffect with circular symmetry (Figure 15c), as MacKay (1957) has reported.

In Section 5 I noted that the noise suppression properties of shunting lateral inhibition also imply spatial frequency properties. Consequently, dipole fields whose subfield inhibition is of shunting type are capable of spatial frequency adaptation. A grating with a sinusoidal luminance profile of prescribed spatial frequency will excite a band of cell types whose inhibitory fields permit maximal excitation by the input. If the input stays on for awhile, the activated transmitter gates will be differentially depleted. Test inputs

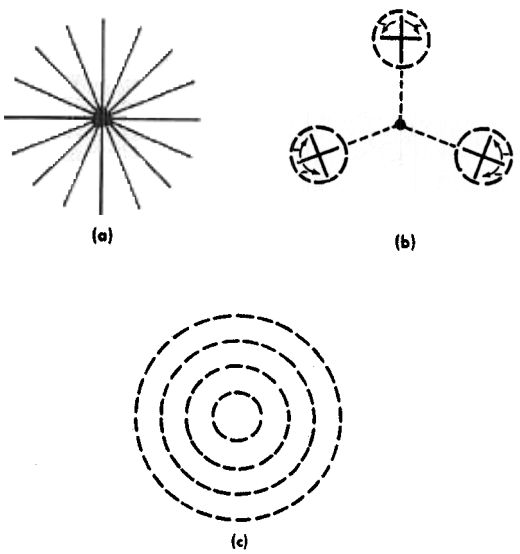

Figure 15. In (a), a pattern with radial symmetry is inspected for a long time. In (b), the net inhibitory interactions among mutually perpendicular orientations at each position are depicted. In (c), offset of the radial pattern elicits antagonistic rebounds across the field that differentially activate the perpendicular orientations.

with similar spatial frequencies share some of these gated pathways, so the overall sensitivity of response to these inputs will be less (Grossberg, 1976b). This view of spatial frequency adaptation contrasts with the view developed by Wilson (1975) that spatial frequency adaptation is due to classical conditioning of an inhibitory transmitter. It is often assumed that a slowly varying effect implies a conditioned change. The alternative notions that "fatigue" and antagonistic activity can yield perceived changes are also very old (see Brown, 1965, for a review).

The present theory refines the latter view by noting how slowly varying changes can follow from dipole adaptation without any conditioning taking place. In particular, even if the adaptational differences decay until they are very similar, contrast enhancement due to fast recurrent competitive interactions can bootstrap these differences into the perceivable range. An interaction between slow transmitters and fast recurrent interactions can hereby create behavioral effects that are much more enduring than the transmitter decay rate would suggest. This suggestion is made

again later for the McCollough effect. The Wilson model differs from the present theory in its STM properties as well as in its description of slow adaptation effects. Wilson used the Wilson–Cowan equations to describe fast intercellular interactions. Among other differences, these equations do not incorporate automatic gain control by lateral inhibitory signals (Grossberg 1973). Consequently, the Wilson–Cowan equations cannot retune their sensitivity in response to shifts in background input intensity, a difficulty that also occurs in all additive models of lateral inhibition.

Pattern-contingent colored aftereffects can also be generated in a dipole field. Suppose that a prescribed field of feature detectors is color coded. Let the on-cells be maximally turned on by red light and the off-cells be maximally turned on by green light for definiteness. Then white light will excite both on-cells and off-cells; that is, white light acts like an adaptation level in this situation. Suppose that a red input whose features are extracted by the field is turned on long enough to substantially deplete its transmitter. What happens if a white input replaces the red input on test trials? The depletion caused by the prior red input now causes the white adaptation level to generate a larger gated signal to the green channel, so a green pattern-contingent aftereffect will be generated.

How enduring will this aftereffect be? Here we must recall that the anatomies of $F_+^{(2)}$ and $F_-^{(2)}$ are recurrent, and that one property of such recurrent anatomies is their ability to contrast enhance small differences in net input into large differences that can then be stored in STM (Section 8). Thus, even if the large initial differences in transmitter depletion within the on-cell off-cell dipoles decay steadily to small differences, the recurrent anatomy can contrast enhance these small differences into a perceptually visible aftereffect when the white test pattern is presented. For this to happen, however, the feature field must be protected from new inputs that can disrupt the pattern of small differences until the test trial occurs. Sleep can hereby prolong the apparent duration of the aftereffect. These properties are familiar ones in the McCollough effect (MacKay & MacKay, 1975; McCollough, 1965).

Various authors have suggested that the long duration of the McCollough effect implicates classical conditioning mechanisms. Montalvo (1976) presented a particularly ingenious application of this idea. This approach seems to trade-off one paradox for another, since the classical conditioning must produce a negative aftereffect during test trials, rather than the positive effect that was experienced during learning trials. Unless one can isolate a large class of phenomena in which classical conditioning reverses the effect on test trials, this explanation is hard to understand from the viewpoint of basic neural design. The present theory points out that slowly varying transmitter gates supplemented by rapid contrast enhancement and STM storage in a recurrent anatomy can also generate long-term effects whose duration is much longer than the transmitter decay rate would suggest. Such long-term effects must unambiguously be ruled out before classical conditioning is invoked as a unitary explanation.

Dipole field structure also helps to explain monocular rivalry (Rauschecker, Campbell, & Atkinson, 1973), whereby two superimposed gratings with the same sinusoidal luminance profile, one vertically oriented and one horizontally oriented, and each illuminated by white light or by different (say complementary) colors, are seen to alternate through time. The tendency toward rivalry can be explained by the recurrent inhibition across orientationally tuned on-cells and across orientationally tuned off-cells; the vertical on-cells tend to inhibit the horizontal on-cells, and conversely. The tendency to alternate can be explained by the fact that persistent STM reverberation of the active vertical on-cells tends to deplete their transmitter gates, thereby weakening their reverberation and providing a relative advantage to the inhibited, and therefore relatively undepleted, horizontal on-cells. When the vertical on-cell depletion reaches a critical value, the horizontal on-cells are sufficiently disinhibited to allow the recurrent dynamics to contrast enhance the horizontally coded inputs into STM. The horizontally coded on-cells thereupon reverberate in STM until the cycle repeats itself. Thus the main effect can be ascribed to combined effects of slow transmitter depletion, recurrent inhibition

across orientations, and the contrast-enhancing capabilities of the recurrent network, even if there are no changes in gaze.

Of particular interest is the fact that the alternation rate depends on the color of the gratings. Two white and black gratings, or two monochromatic gratings, alternate up to three times slower than gratings that are illuminated by complementary colors. This can be discussed in terms of the rebound behavior that occurs between subfields that are orientationally coded and whose dipoles code for complementary colors. When two white and black gratings of sufficient contrast are used, the white inputs can excite both on-cells and off-cells of the color-coded dipoles, thereby inhibiting them. It is therefore assumed that apart from altering their gain, intense black and white gratings cause net excitation primarily in feature fields whose on-cells respond unselectively to light-on and whose off-cells respond to light-off. In such a feature field, the horizontal and vertical white bars excite the same subfield, and the horizontal and vertical black bars excite the complementary subfield. Each subfield tends to adapt or conserve its total STM activity (within its functional channels!) so that there exists a tendency for the horizontal and vertical inputs to compete for STM activity, and to thereby decrease the transmitter depletion rate in active cells.

By contrast, consider what happens in a color-coded dipole field in response to two gratings that use the field's complementary colors, say red-vertical and green-horizontal. Here the red-vertical bars deplete only the red field, and the green-horizontal bars deplete only the green field. There is no direct inhibition within a given subfield between horizontal and vertical orientations. Thus, other things equal, greater STM activation of red-verticals or green-horizontals is possible than in the black–white case because less intrafield competition for STM activity occurs. Greater STM activation implies faster transmitter depletion and faster alternation rates. If this explanation is correct, then it is a special case of a more general phenomenon; namely, that the frequency of perceptual oscillations can be pattern contingent due to the intrafield normalization property.

Other aftereffects provide more direct evidence for the existence of slowing varying transmitter gates. In particular, the effects of changing background illumination, or the *secondary field*, on aftereffects are remarkably similar to the effects of changing arousal level on the rebound. If a secondary field is turned on during the observation of a positive afterimage in darkness, then a rapid transition to a negative afterimage can be generated (Brown, 1965, p. 483; Helmholtz, 1866, 1924). If the secondary field is then turned off, the afterimage can revert in appearance to that of the stage when the secondary field was first turned on. In a dipole, an increase of adaptation level tends to rebound the relative dipole activities. If the arousal level is then decreased, the slowly varying transmitter levels can still be close to their original values, so that the original relative dipole activities are rapidly restored. The higher the luminance of the secondary field, the shorter the afterimage latency, and the more rapidly the afterimage is extinguished (Juhasz, 1920). In a dipole, a higher adaptation level more rapidly equalizes the amounts of transmitter in the two dipole channels by depleting them both at a faster, more uniform rate. When approximately equal levels of transmitter are achieved, the inhibitory interneurons between the dipole's populations kill any relative advantage of one population over the other. The duration of an afterimage increases with an increase in primary stimulus luminance (Brown, 1965, p. 493). In a dipole, increasing the intensity of an input to one population increases the rebound at the other population when the input terminates, much as termination of a more intense shock causes greater relief, other things being equal (Grossberg, 1972c, 1976b).

The preceding considerations lead to some experimental predictions. Some of these concern red and green cortical dipoles. For example, suppose that a red stimulus has activated a red-cell long enough to substantially deplete the transmitter. Does an increment in white light cause a green-cell rebound? Does a decrement in red light from J units to $J/2$ units cause a smaller rebound when white light is on than a decrement from $J/2$ units to 0 units? Is there an inverted U in dipole responsiveness as a function of the arousal

level or the intensity of white light? Does the relative rebound size increase as a function of arousal level size for intermediate levels of arousal? In other words, are visual dipoles designed the same way as motivational dipoles?

Another set of predictions concerns the McCollough effect. For example, how does the McCollough effect depend on the intensity of white light during test trials? A more intense white light should yield an initially larger aftereffect unless the white is so intense that overarousal occurs. Moreover, more intense white should equalize the relative transmitter stores more rapidly than less intense white. This suggests an experiment in which a double test is made. The first test uses prolonged inspection of white bars whose intensity differs across subjects. Before the second test is made, some visual experience should occur to blot out whatever small differences in transmitter storage might still exist after the bright white bars are examined. Then a second test with white bars is given. Subjects who saw less intense bars on the first test should perceive a larger aftereffect.

An experiment concerning spatial frequency adaptation is also suggested. This experiment is analogous to the experiment on aftereffects due to changes in the secondary field. Speaking generally, if spatial frequency adaptation and certain other aftereffects are all due to dipole depletion, albeit in different fields of feature detectors, then they should undergo similar transformations in response to analogous experimental manipulations, other things being equal, notably the persistence with which each feature field is disrupted by uncontrolled inputs. Suppose that when a series of vertical sinusoids drifts horizontally across the visual field, those on-cells and off-cells whose recurrent inhibitory signals collide with visually induced inputs will have their activities suppressed. Consider the on-cells and off-cells that can be activated by the prescribed spatial frequency. What happens as the contrast of the visual pattern is parametrically increased across subjects?

This is a delicate question because more than one dipole field in the coding hierarchy can be activated by such an input. Let us consider what would happen if only one dipole field is activated. In the limit of absolute black and very bright white verticals, both the on-cells and the off-cells would be almost equally excited on the average, albeit at different times, as the light and dark verticals drift over their receptive fields. Neither on-cell nor off-cell would gain a large *relative* advantage, but both would have their transmitter stores significantly depleted by the persistence of the horizontally drifting input. Hence, significant spatial frequency adaptation would occur, but not due to large relative imbalances in the dipoles. What happens as the contrast between the white and black verticals is decreased? Then other things being equal, the off-cells will be depleted more than the on-cells. Hence, a greater relative depletion within the dipoles can be induced at smaller contrast levels than at larger contrast levels. How can this conclusion be tested? Consider two groups of subjects. Let Group 1 be adapted and tested using high contrast gratings. Let Group 2 be adapted on a lower contrast grating and tested using the same higher contrast grating used to test Group 1. The net on-responses at a black–white interface as the test grating slowly drifts across the visual field should be greater in Group 2 than in Group 1. Can such differentially enhanced boundaries between the trailing edge of black and the leading edge of white be perceived? If the answer is yes, then one can properly claim that the effect is a functional analogue within the visual system of the partial reinforcement acquisition effect in the motivational system (Grossberg, 1975).

13. Reset Wave: Reaction Time, P300, and Contingent Negative Variation

The nonspecific arousal that is triggered by unexpected events (or mismatch) selectively and enduringly inhibits active population across $F^{(2)}$. In vivo, do there exist broadly distributed inhibitory waves that are triggered by unexpected events? In average evoked potential experiments, one often finds such a wave, namely the P300 (Rohrbaugh, Donchin, & Eriksen, 1974; Squires, Wickens, Squires, & Donchin, 1976). The theory's relationship to P300 is discussed in Grossberg (1978e), in which the following properties of P300 are

shown to be analogous to properties of the resetting wave: Reaction time is an increasing function of P300 size (Squires et al., 1976); P300 is not the same average evoked potential as the contingent negative variation (CNV) (Donchin, Tueting, Ritter, Kutas, & Heffley, 1975; cf. Section 16); P300 can be elicited in the absence of motor activity (Donchin, Gerbrandt, Leifer, & Tucker, 1972); resetting the STM codes of longer sequences of events can take longer than resetting the STM codes of shorter sequences of events, and due to the relationship between reaction time and P300 size, longer sequences will elicit larger P300s (Remington, 1969; Squires et al., 1976). Moreover, Chapman, McCrary, and Chapman (1978) showed that in a number- and letter-comparison task, there existed an evoked potential component with a poststimulus peak at about 250 msec that is related to the storage of cue-related information in STM. This latency fits well with the idea that STM storage occurs if the feedback expectancy does not create a mismatch. The extra 50 or so msec needed to generate a P300 would also be necessary in the network to trigger the reset wave if a mismatch does occur.

If the P300 is indeed a reset wave of the type that the thought experiment describes, then several types of experiments can be undertaken to test this hypothesis. On the anatomical side, Where does the expectancy matching take place? What pathways subserve the arousal? On the physiological side, Do dipole rebounds cause the inhibition? On the psychophysiological side (e.g., average evoked potential experiments), Is there a more direct experimental paradigm for testing whether P300 directly inhibits STM? In particular, Can a succession of P300s be reliably triggered when information is disconfirmed in successive stages? On a deeper functional level, Does the P300 act to buffer committed cells against continual recording by the flux of experience? If P300 is inhibited, can previously committed cells be recoded? In other words, when we consider cognitive coding, does a chemical switch contribute to code stability, or is code stability entirely dependent on buffering by dynamic reset mechanisms?

As was noted in Section 4, feedback expectancies that trigger STM reset mechanisms should occur in many thalamocortical systems, so that there should exist different reset waves corresponding to each functionally distinct system. In Grossberg (1978e), the preceding scheme is generalized to a variety of examples in which competition occurs between attentional, or consummatory, pathways and novelty, or orienting, pathways. A matching process goes on within the attentional system and computes such information as follows: Are the sensory cues the ones that are expected? Do the proprioceptive motor cues match the terminal motor map that is guiding the limb? If the answer is yes, then goal-oriented arousal systems are activated to support the matching process and its consequences, such as posture. If the answer is no, then complementary arousal systems are activated that support rapid reset and orienting reactions aimed at acquiring new information with which to correct the error. Given that the P300 helps to reset sensory STM in response to unexpected events, does there exist a complementary wave that occurs along with expected events? The CNV would appear to be such a wave (Cohen, 1969), since it is associated with an animal's expectancy, decision (Walter, 1964), motivation (Cant & Bickford, 1967; Irwin, Rebert, McAdam, & Knott, 1966), volition (McAdam, Irwin, Rebert, & Knott, 1966), preparatory set (Low, Borda, Frost, & Kellaway, 1966), and arousal (McAdam, 1969).

If the P300 and the CNV are indeed complementary waves, then experiments should be undertaken to determine the neural loci at which the generators of these waves compete. For example, Section 16 suggests that the hippocampus provides output that contributes to the CNV. Does expectancy mismatch occur within the hippocampus, or in a cell nucleus that activates hippocampus, and thereby release a P300 by disinhibiting its generator?

Having noted the existence of reset and attentional waves that are triggered by sensory events, it is natural to ask whether there exist analogous waves that are triggered by motor events? To answer this question, the next section considers how eye movements can modulate the LGN's sensitivity to afferent visual signals and the related questions of whether the LGN has a dipole field organization and whether feedback from visual cortex

to LGN can selectively attenuate or amplify afferent visual signals. This discussion leads to a reinterpretation of LGN data and to some predictions. These predictions concern the possible existence of a reset motor wave and the timing of certain developmental events relative to the end of the critical period for plasticity in the primary visual cortices.

14. Template Matching and Reset: PGO Wave, Geniculate Dipoles, and Corticogeniculate Feedback

An example of an "attentional" motor wave seems to be the ponto-geniculo-occipital (PGO) wave whose effects on the LGN are admirably reviewed by Singer (1977). Singer (1977) distinguished at least two types of inhibitory interneurons in his discussion of LGN dynamics:

> There apparently are two inhibitory mechanisms with two different functions. One is based on intrinsic interneurons and presumably conveys the retinotopically organized and highly selective inhibitory interactions between adjacent retinocortical channels. . . . This inhibition seems to be mainly of the feed-forward type. . . . The second inhibitory pathway is exclusively of the recurrent type and is relayed via cells in nucleus reticularis thalami. . . . This extrinsic inhibitory loop is probably involved in more global modifications of LGN excitability as they occur during changes in the animal's state of alertness and during orienting responses associated with eye movements. (p. 394)

Singer noted that mesencephalic reticular formation (MRF) stimulation leads to field potentials in the LGN and the visual cortex that closely resemble PGO waves. LGN transmission is facilitated during PGO waves and during the analogous negative field potential that occurs after MRF stimulation.

One mechanism of MRF facilitation is inhibition of the cells in the nucleus reticularis thalami, which are recurrent inhibitory interneutrons between LGN relay cells. From a theoretical viewpoint, this type of disinhibition would be expected to have nonspecific effects like decreasing the quenching threshold of an entire recurrent subfield of cells, and thereby facilitating transmission of signals through these cells (Grossberg, 1973; Grossberg & Levine, 1975). Such an effect seems to occur in LGN. Since MRF stimulation can completely suppress inhibitory postsynaptic potentials elicited from optic nerve or optic radiation, Singer (1977) concluded

> that the intrinsic inhibitory pathways also get inactivated. However, it cannot yet be decided whether the inhibitory interneurons in the main laminae are also subject to direct reticular inhibition as is the case for cells in nucleus reticularis thalami. (p. 409)

Singer went on to suggest that corticogeniculate feedback could partially accomplish the intrinsic cell inhibition.

For present purposes, the main point is Singer's (1977) functional interpretation of the MRF-induced LGN disinhibition. He claimed

> that the brief phase of disinhibition serves to reset the thalamic relay each time the point of fixation is changed. . . . To assure a bias-free initial processing of the pattern viewed after a saccade . . . inhibitory gradients ought to be erased before the eyes come to rest on the new fixation point . . . the concomitant disinhibition occurs only towards the end of the saccade right before the eyes come to rest. (p. 411)

Singer's remarks can be mechanistically interpreted as follows: As the proprioceptive coordinates of the eye muscles approach the terminal motor coordinates that control the saccade, the two sets of coordinates match, a PGO wave is initiated, it disinhibits LGN relay cells, and prepares the LGN to transmit retinal signals to the visual cortex. If the PGO wave is indeed elicited by a matching process between the terminal motor map and proprioceptive coordinates of the eye muscles, then this matching process should be capable of exciting cells that inhibit the LGN interneurons within the nucleus reticularis thalami. In what neural structure does this matching process take place? One component of this structure might already have been discovered by Tsumoto and Suzuki (1976), who report a pathway from the frontal eye fields to the perigeniculate nucleus in which are found the LGN inhibitory interneurons. Electrical stimulation of the frontal eye fields inhibits the perigeniculate cells and facilitates LGN transmission.

Singer (1977) claimed that the PGO wave resets the LGN so that it can respond to retinal signals without bias. However, nonspecifically reducing the quenching threshold is not the type of selective reset that I have discussed earlier. Indeed, Singer's discussion of LGN dynamics emphasizes the wiping away

of all inhibitory gradients as a reset mechanism. But what if excitatory activities already exist in the LGN when this happens? Why do these activities not get amplified and thereupon maximally bias LGN activity in response to the next retinal input volley? I suggest that the LGN reset that is due to the nucleus thalami reticularis occurs while the eye is moving and the extrinsic inhibitory interneurons are active. This extrinsic inhibitory feedback resets the LGN by generating a high quenching threshold and thereby wiping out the LGN's excitatory patterns. As the eye comes to rest at its intended position, I suggest that matching occurs between the terminal and proprioceptive motor maps of the eye muscles, thereby activating the attentional system, in particular the PGO wave, which sensitizes the LGN to retinal and cortical signals.

Even if the preceding interpretation of Singer's argument is correct, it discusses a nonspecific effect on the QT and the sensitivity of visual pattern processing, but not the selective reset that aims at reorganizing attention in response to an error, or other unexpected event.

Is there a wave that is functionally complementary to the PGO wave, that can precede it, and that drives a selective reset of LGN dynamics in response to unexpected events? If such a wave does exist, it would be functionally analogous to the P300. In this regard, Singer (1977) parenthetically mentions the work of Foote, Manciewicz, and Mordes (1974) to explain the inhibition of LGN transmission that sometimes occurs shortly after MRF stimulation but before the facilitatory phase. Foote et al. suggest that this inhibitory pathway is due to serotonergic fibers originating in the dorsal raphe nucleus. Are these fibers the pathway over which selective reset can occur?

For a selective reset wave to exist, it must operate on on-cell off-cell dipoles. Do such dipoles exist in the LGN? Much of the data discussed by Singer was collected in the cat LGN. Singer (1977) reports here

> that reciprocal inhibitory connections exist between adjacent neurons driven by the same eye that have the same receptive field center characteristics; i.e., between on-center cells and between off-center cells, respectively. (p. 390)

These interneurons are analogous to the intrafield lateral inhibition that was postulated within $F_+^{(2)}$ and $F_-^{(2)}$, but which we now recognize as a prerequisite for total activity adaptation and quenching threshold tuning in any recurrent network. In addition, there exist "reciprocal inhibitory interactions between neurons with antagonistic field center characteristics—that is, between on- and off-center units with spatially overlapping receptive fields" (Singer, 1977, p. 390). These cells would appear to form dipoles. If they are dipoles of the type discussed, then the arousal system that triggers their rebounds will feed into them—from the dorsal raphe nucleus—and activating this arousal system will rebound their relative activities.

These hypotheses should be easier to test in the monkey than the cat, because Schiller and Malpeli (1978) have reported that of the four parvocellular layers in the monkey, the two layers committed to the left eye are subdivided into an on-cell layer and an off-cell layer, and the two layers committed to the right eye are also subdivided into an on-cell layer and an off-cell layer. Do dipole interactions occur between the on-cell and off-cell layers of each eye representation? Does a suitable arousal increment rebound the relative activities of these dipoles? If so, we will have found an elegant functional reason for the existence of this structure in the monkey: Each eye has its own dipole field to carry out its selective reset modes. We will also have found an elegant reason for the existence of intrinsic and extrinsic inhibitory systems: Attentional reduction of the quenching threshold is functionally distinct from, and even complementary to, selective reset.

Another important point of Singer's (1977) article concerns the role of corticogeniculate feedback.

> In a highly selective way the cortex permits transmission of binocular information that can be fused and evaluated in terms of disparity depth cues while it leaves it to the intrinsic LGN circuits to cancel transmission of signals that give rise to disturbing double images. (p. 398)

In other words, the corticogeniculate feedback acts as a template that selectively enhances the type of data that the cortex is capable of coding in a globally self-consistent way.

In summary, the LGN seems to enjoy a dipole field structure whose sensitivity to afferent sensory signals is modulated both by corticogeniculate feedback, which acts like a sensory expectancy-matching mechanism, and by MRF arousal, which lowers the LGN QT in response to proprioceptive-terminal map matching within the eye movement system.

If we interpret the geniculocortical relay as an example of our thought experiment, then several experimental predictions arise. These predictions are made with caution, since a significant part of visual development seems to be genetically prewired in the geniculocortical pathways of higher mammals (Hubel & Wiesel, 1977). It is still not clear, however, to what extent corticogeniculate feedback does help to terminate the visual critical period in these animals. Nor is it clear whether the same neural design that is used in some species, or in individual neural relays, to terminate a critical period using feedback is also used in others wherein a chemical switch or other prewired mechanisms are appended. The predictions flow from the observation that if the geniculocortical system is an example of the thought experiment, *albeit* vestigially, then its reset and search mechanisms must develop before the end of the visual critical period. In particular, if lateral inhibition within the LGN is used to help match cortical and retinal data, then these inhibitory connections must develop before the end of the critical period. The dipole field structure of the cortex must also develop before the end of the critical period. Moreover, mismatch within the LGN system should disinhibit an arousal system capable of rebounding the cortical dipoles. There exists a catecholamine arousal system to neocortex, among other structures (Fuxe, Hökfelt, & Ungerstedt, 1970; Ungerstedt, 1971; Jacobowitz, 1973; Lindvall & Björklund, 1974; Stein, 1974). Is this the arousal system being sought? Does it develop before the end of the critical period? Is this arousal system capable of driving antagonistic rebound in cortical dipoles? Is a catecholamine transmitter always used in arousal systems that drive antagonistic rebound, for example, the catecholamine system originating in the dorsal raphe nucleus that was described by Foote et al. (1974)? Finally, is there a structural similarity between all pairs of attentional and selective reset waves, as between CNV and P300, or PGO and its hypothesized complementary reset wave?

15. Adaptive Resonance, Code Stability, and Attention

The preceding sections discuss some of the network events that occur when feedforward data mismatch feedback expectancies. What happens if an approximate match occurs? Then the activity patterns at $F^{(1)}$ and $F^{(2)}$ elicit interfield signals that mutually reinforce each other, and activities at both levels are amplified and locked into STM. Because the STM activities can now persist much longer in time than the passive decay rates of individual cells, the slowly varying feed-forward filters and feedback expectancies have sufficient time to sample the STM patterns and store them in LTM. I call this dynamical state an *adaptive resonance*. The resonant state provides a global interpretation of the afferent data, or a context-dependent code, that explicates in neural terms the idea that the network is paying attention to the data. The resonance idea suggests that many individual neural events, such as cell potentials and axonal signals, are behaviorally irrelevant until they are bound together by resonant feedback. Of special importance is the observation that unless resonance occurs, no coding in LTM can take place. This observation clarifies from a mechanistic viewpoint the psychological fact that a relationship exists between paying attention to an event and coding it in LTM (Craik & Lockhart, 1972; Craik & Tulving, 1975).

The resonant state provides a context-dependent code due to several factors acting together. For example, the pattern of expectancy feedback can alter, through a matching process, the activity that a given feature detector would have experienced if only afferent signals were operative. Similarly, competitive interactions within a subfield can rapidly alter the net input pattern before storing it in STM. Thus when an activity pattern at a field $F^{(1)}$ is projected by interfield signaling to a field $F^{(2)}$, feedback from $F^{(3)}$ to $F^{(2)}$ can deform this pattern before it is further

reorganized by competition within $F^{(2)}$. Because the resonant state provides a context-dependent code whose resultant patterns in STM and LTM depend on all active components of the system, it is impossible to determine the code from the measurements taken by any single microelectrode, no matter how precise its calibration. I claim that adaptive resonances are the functional units of cognitive coding, and that classification of the resonances that occur in prescribed situations is a central problem for cognitive psychology. The structural substrates of these cognitive units are nonlinear feedback modules involving whole fields of cells rather than individual nerve cells.

The technical details needed to rigorously build up the resonance idea are derived in Grossberg (1976a, 1976b), in which a summary of related coding models is also given, and further developed in Grossberg (1978e). In particular, Grossberg (1976a) points out that a coding theory that depends on a feedforward anatomy with any fixed number of cells is faced with a crippling dilemma: Either a chemical switch turns off code development at a prescribed time, but then the code will be behaviorally meaningless with a high likelihood, or the code is unstable through time whenever the number of patterns in the environment significantly exceeds the number of coding cells. It is also proved that a developing code can be stable in a sparse input environment, but this does not address the typical situation in vivo, where a continuous visual flow, and therefore a nondenumerable series of visual patterns, must be dealt with. Computer models of code development missed these basic points because they typically used small numbers of inputs and small numbers of coding cells.

Once the main point was vividly made, one could see that feedback was essential to stabilize a developing code in a rich input environment, and that the types of feedback that were needed resembled attentional mechanisms that had previously been derived from different considerations, namely classical and instrumental conditioning postulates (Grossberg, 1975). The examples in Sections 2 and 3 illustrate these attentional phenomena. Two pleasing conclusions were thereby drawn: Adult mechanisms, in this case attentional mechanisms, are often continuations along a developmental continuum of infantile mechanisms, in this case code development mechanisms; and the rather mysterious rubrics of "paying attention" and "expectancy" could be attached to the more substantial theme of "code stability and consistency," and the establishment of dynamically maintained critical periods.

Anderson, Silverstein, Ritz, & Jones (1977) also recognized the importance of feedback in defining the functional units of neural network. Their model differs, however, from the present theory in several notable respects. The recurrent STM interactions in their model are defined by linear feedback signals. Grossberg (1973, 1978d) shows how linear feedback signals among cells that are capable of saturating create unphysical instabilities such as noise amplification and compression of an input pattern. Furthermore, known neural nonlinearities, such as sigmoid signals between cells, overcome these instabilities and contrast enhance the input pattern (Appendix D). The LTM interactions in the Anderson et al. model are described by summing up a large number of mutually orthogonal LTM vectors

$$z = \sum_{k=1}^{n} z_k.$$

to form the total LTM trace across a field $F^{(1)}$ of cells. When a signal pattern S from $F^{(1)}$ to $F^{(2)}$ is gated by the total LTM trace, as in

$$S \cdot z = \sum_{k=1}^{n} S \cdot z_k$$

(see Appendix A), it might be perpendicular to all but one of the increments, say z_1. Consequently, the net signal is $S \cdot z_1$. This concept gets into difficulty because in vivo the total LTM trace z must be composed of small quantities (e.g., transmitter concentrations). Each of the summands z_k must therefore be a very small quantity unless n is also very small, but then the theory is powerless. If each z_k is very small, then the net signal depends on gating by a sum of very small quantities. This creates an unstable situation. Furthermore, the LTM trace z_{ij} from cell v_i to cell v_j in the

Anderson et al. model is assumed to equal the LTM trace z_{ji} from v_j to v_i. This symmetry assumption is too restrictive for our purposes, since we do not want the filter from $F^{(1)}$ to $F^{(2)}$ to necessarily equal the expectancy from $F^{(2)}$ to $F^{(1)}$; this would limit the tendency to achieve greater abstractness of feature extraction in a hierarchy of fields. A more serious problem for the LTM symmetry assumption is its implication that the signal from every cell be proportional to its STM trace. This follows because the growth rate of LTM trace z_{ij} is proportional to the product of signal S_{ij} from v_i and v_j times STM trace x_j of v_j. To achieve $z_{ij} \equiv z_{ji}$, it is necessary for $S_{ij}x_j \equiv S_{ji}x_i$, which is possible if $S_{ij} = \alpha x_i$ and $S_{ji} = \alpha x_j$. In particular, the recurrent signals must be linear functions of the STM traces, and the usual instabilities that recurrent linear signals generate among cells will be generated.

The next three sections summarize a few resonant schemes and predictions pertaining thereto that suggest the scope of the resonance phenomenon. Other resonances, notably the olfactory resonance that is described by the distinguished work of Freeman (1975), are discussed more completely in Grossberg (1976b, 1978e). Freeman discovered a resonant phenomenon by performing parallel electrode experiments on the cat prepyriform cortex. When the cat smells an expected scent, its cortical potentials are amplified until a synchronized oscillation of activity is elicited across the cortical tissue. The oscillation organizes the cortical activity into a temporal sequence of spatial patterns. The spatial patterns of activity across cortical cells carry the olfactory code. By contrast, when the cat smells an unexpected scent, then the cortical activity is markedly suppressed. Freeman traces the differences in cortical activity after expected versus unexpected scents to gain changes within the cortical tissue. Appendix C shows how a matching mechanism in a shunting network simultaneously changes gains as it amplifies or attenuates network activity. Freeman also notes a tendency for the most active populations to phase-lead less active populations. This also occurs automatically in a shunting network due to the correlation between gain and asymptote. Because the cortex oscillates, Freeman models his data using second-order differential equations whose coefficients are changed by expectations in a manner that is descriptively stated, but not dynamically explained, by his model. I suggest that the oscillations are caused by feedback between cells that obey first-order differential equations whose gains are changed by signals coupled to the shunting mechanism.

Grossberg (1978b) claimed that adaptive resonances also occur in nonneural tissues, where they are suggested to be a basic design principle in a universal developmental code. Syncytium formation during sea urchin gastrulation is identified as a possible adaptive resonance phenomenon—in particular, the law whereby pseudopods from the mesenchymal cells adhere to ectodermal cells to form a syncytium has the same form as the law for an LTM trace—and some predictions are made to test this hypothesis.

16. Are Conditioned Reinforcer Pathways and Conditioned Incentive Motivational Pathways Reciprocal Pathways in an Adaptive Resonance?

Grossberg (1975) has described a psychophysiological theory of attention in which an adaptive resonance occurs. This resonance helps to explain why the dilemma of cross-conditioning that is depicted in Figure 2 does not routinely occur. Figure 16 idealizes this resonance. Speaking intuitively, the internal representations of external cues elicit signals that are, before conditioning takes place, distributed nonspecifically across the various internal drive representations. During conditioning trials, the pattern of reinforcement and drive levels strengthens the LTM traces within certain of these signal pathways and weakens the LTM traces within other pathways. These conditioned changes in the signal pathways endow the external cues with conditioned reinforcer properties. On recall trials, these conditioned signals combine with internal drive inputs to determine whether or not feedback signals will be elicited. The feedback signals play the role of incentive motivation in the network. Incentive motivation is released in a given feedback pathway only if the momentary balance of conditioned reinforcer signals plus drive inputs compete

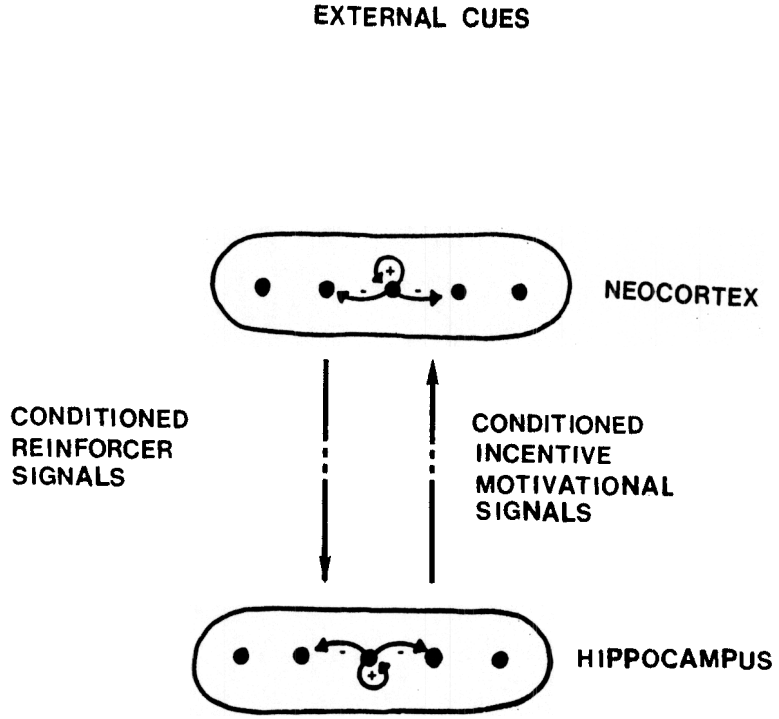

Figure 16. An adaptive resonance between neocortex and hippocampus is suggested to occur when external cues are compatible with internal needs. (The conditionable feedback pathways in this module subserve conditioned reinforcer properties and conditioned incentive motivational properties.)

favorably against these factors within the other feedback pathways.

Before conditioning occurs, each of the incentive motivational channels nonspecifically projects to the external cue representations. As in the case of the conditioned reinforcers, the incentive motivational channels are conditionable, and their LTM traces can be strengthened when their signals are large and contiguous to active external cue representations. Thus after conditioning occurs, an internal drive representation can deliver incentive motivational signals preferentially to those external cue representations with which it was previously associated. In this way, activating a given external cue representation can sensitize an ensemble of motivationally related external cue representations via incentive motivational feedback. A type of subliminal psychological set is hereby formed. Since the external cue representations compete among themselves for storage in STM, the conditioned incentive motivational feedback abets the storage of compatible cues and tends to overshadow the storage of incompatible cues. Thus, during alternate scanning of incompatible cues, attentional switching between resonances compatible with one cue class and then the other class buffers each class against indiscriminate cross-conditioning with incompatible cues. Various data and related theories about reinforcement and attention are analyzed in the light of these concepts in the articles of Grossberg (1972b,

1972c, 1975). Herein I suggest a neural substrate for this resonance and a psychophysiological experiment to test its existence.

In Figure 16, $F^{(1)}$ and $F^{(2)}$ both possess recurrent on-center off-surround interactions, and both the $F^{(1)} \rightarrow F^{(2)}$ and $F^{(2)} \rightarrow F^{(1)}$ pathways are conditionable. Region $F^{(1)}$ contains external cue representations, and region $F^{(2)}$ contains internal drive representations. When this network is embedded into a more complete system of interactions, an interpretation of $F^{(1)}$ as neocortex and of $F^{(2)}$ as hippocampus is suggested. Given this interpretation, the conditioned reinforcer pathways $F^{(1)} \rightarrow F^{(2)}$ should have a final common pathway at hippocampal pyramidal cells, and their LTM traces should be sensitive to the balance of drives and reinforcements through time. Relevant data have been collected by Berger and Thompson (1977), who describe neural plasticity at the hippocampal pyramids during classical conditioning of the rabbit nictitating membrane response.

The conditioned incentive motivational pathways $F^{(2)} \rightarrow F^{(1)}$ should have a final common pathway at neocortical pyramidal cells, and their LTM traces should be sensitive to the balance of motivation and cue saliency through time. The CNV is a conditionable neocortical potential shift that has been associated with an animal's motivational state (Cant & Bickford, 1967; Irwin et al., 1966), and Walter (1964) has hypothesized that the CNV shifts the average baseline of the cortex by depolarizing the apical dendrites of its pyramidal cells. If the conditioned incentive motivational feedback is indeed realized by the CNV and if adaptive resonances between conditioned reinforcers and conditioned incentives do exist, then there should exist neural feedback loops between neocortex and hippocampus such that while conditioned reinforcer properties are being established with the hippocampal pyramid cells as a final common pathway, simultaneously conditioned incentive properties are being conditioned with the apical dendrites of neocortical pyramid cells as a final common pathway. Experiments to test this prediction would require either simultaneous measurement from electrodes in the neocortical and hippocampal loci of the resonant circuit, or correlation of electrode measurements in the hippocampus simultaneous with CNV measurements.

17. Pattern Completion, Hysteresis, and Gestalt Switching

Consider what happens to an adaptive resonance as its afferent data are slowly and continuously deformed through time, say from the letter O to the letter D. By "slowly" I mean slowly relative to the rate with which resonant feedback can be exchanged. Recall that feedback from $F^{(2)}$ to $F^{(1)}$ can deform what "is" perceived into what "is expected to be" perceived. Otherwise expressed, the feedback is a prototype, or higher order Gestalt, that can deform and even complete activity patterns across lower order feature detectors. For example, suppose that a sensory event is coded by an activity pattern across the feature detectors of a field $F^{(1)}$. The $F^{(1)}$ pattern is then coded by certain populations in $F^{(2)}$. If the sensory event has never before been experienced, then the $F^{(2)}$ populations that are chosen are those whose codes most nearly match the sensory event because the pattern at $F^{(1)}$ is projected onto $F^{(2)}$ by the positional gradients in the $F^{(1)} \rightarrow F^{(2)}$ pathways (Appendix A). If no approximate match is possible, then mismatch at $F^{(1)}$ will trigger a reset wave that selectively inhibits $F^{(2)}$ and elicits a search routine. If an approximate match is possible, however, then the feedback signals from $F^{(2)}$ to $F^{(1)}$ will elicit the template of the sensory events that are optimally coded by the $F^{(2)}$ pattern. These feedback signals rapidly deform the $F^{(1)}$ pattern until this STM pattern is a mixture of feedforward codes and feedback templates. Otherwise expressed, $F^{(2)}$ tries to complete the $F^{(1)}$ pattern using the prototype, or template, that its active populations release.

In Grossberg (1978e, Section 40), another completion mechanism is also suggested, namely a normative drift. This mechanism generalizes the line neutralization phenomenon that was described by Gibson (1937). In suitably designed feature fields, STM activity at a particular coding cell can spontaneously drift toward the "highest order" coding cell in its vicinity, due either to the existence of more cell sites, or to larger and spatially more

broadly distributed feedback signals, at the highest order cells. After STM activity drifts to its local norm, the highest order cell can thereupon release its feedback template. It was shown in Levine and Grossberg (1976) that such drifts are a type of lateral masking due to the recurrent interactions within the feature field. I suggest that many Gestalt-like pattern completions are manifestations of intrafield competitive transformations, such as normative drifts, and the deformation by feedback expectancies of lower order STM patterns. Such global dynamical transformations transcend the capabilities of classical pattern discrimination models (e.g., Duda & Hart, 1973).

Two important manifestations of the completion property are hysteresis and Gestalt switching. For example, once an STM resonance is established in response to the letter O, the resonance resists changing its codes when small changes in the sensory event occur—this is hysteresis. Hysteresis occurs because the active $F^{(2)} \rightarrow F^{(1)}$ template keeps trying to deform the shifting $F^{(1)}$ STM pattern back to one that will continue to code the $F^{(2)}$ populations that originally elicited this template.

If, however, the sensory event changes so much that the mismatch of test and template patterns becomes too great, then the arousal-and-reset mechanism is triggered. This event inhibits the old code at $F^{(2)}$ and forces a search for a distinct code. A dramatic switch between global percepts can hereby be effected. The global nature of the switch is due not only to the rapid suppression of the previously active $F^{(2)}$ code but also to the fact that $F^{(2)}$ contains populations that can synthesize data from many feature detectors in $F^{(1)}$, and the feedback templates of these populations can reorganize large segments of the $F^{(1)}$ field. I suggest that an analogous two-stage process of hysteresis and reset is operative in various visual illusions, such as Necker's cube (Graham, 1965). When ambiguous figures are presented, these mechanisms can elicit spontaneous switches of perceptual interpretation due either to shifts of gaze or to the input-induced cyclic rates of transmitter depletion that can occur even if the gaze remains relatively fixed (Section 12).

If Gestalt switching is a two-stage process, then at the moment of switching, a reset wave should occur. Does a P300 occur at the moment of perceived switching? If not, can this paradigm be used to discover what average evoked potential, if any, parallels activation of the reset mechanism?

18. Binocular Resonance and Rivalry

My final example indicates that adaptive resonances need not be hierarchically organized, and points to a class of resonances of particular importance. Hysteresis can occur between two reciprocally connected fields even if they are not hierarchically organized, since the individual cells do not know whether they are in a hierarchy. For example, suppose that each eye activates a field of monocularly coded feature detectors. Suppose that each monocular field is endowed with a recurrent on-center off-surround anatomy, indeed with a recurrent dipole field of on-cell and off-cells. Let the on-cells in each monocular field be capable of exciting corresponding on-cells in the other monocular field. In other words, signals from a given monocular field act as a template for the other monocular field. It does not matter what features are coded by these detectors to draw the following conclusion. Once a resonance is established between the two monocular fields, hysteresis will prevent small and slow changes in the input patterns from changing the coded activity. Julesz (1971) introduced a field of physical dipoles to model the binocular hysteresis that he and Fender described (Fender & Julesz, 1967). Resonance between two recurrent on-center off-surround anatomies undergoing shunting dynamics provides a neural model of the phenomenon. Such a binocular resonance will generate properties of binocular rivalry, since competition within each subfield of the recurrent networks will inhibit feature detectors that do not participate in the resonance.

The construction of monocular representations whose binocular resonances code globally self-consistent invariants of stereopsis is presently being undertaken. Although this construction is not yet complete, some observations can be made in broad strokes to guide

the reader who is interested in pursuing the elucidation of perceptual and motor resonances.

Before the two eyes can fixate on the scenes that will drive binocular development, there must already exist enough prewired visual feature detectors to direct the eye movement system to lock in the fixation process. Thus the existence of prewired visual feature fields does not argue against the need for visual tuning by experience. Such tuning seems necessary to achieve accurate stereopsis in the face of significant variations in bodily parameters due to individual differences and growth (Daniels & Pettigrew, 1976). An effort should be made to correlate individual and species differences in the motor mechanisms that are used to accumulate visual data and to act on the visual environment with corresponding differences in prewired sensory feature detectors and the ultimate feature fields that can be synthesized (Arbib, 1972; Creutzfeldt & Nothdurft, 1978).

Even if feature development can continue in the absence of visual experience, this does not imply that visual experience does not alter visual development. Just as imprinting can be driven by endogenous drive sources that are later supplanted by environmentally reactive drive sources (Sluckin, 1964), an effort should be made to test whether visual development is driven by endogenous arousal sources before these sources are supplanted by visual experience, in particular by visually reactive arousal sources.

Binocular visual resonances seem to be a special case of bilateral resonances that are due to the bilateral organization of the body, for example, binaural auditory resonances. As in the case of binocular corticogeniculate feedback (Singer, 1977), bilateral interactions at each of several anatomical stages help to select the activity patterns that elicit and are modulated by hierarchical signals. The hierarchical signals are supplemented by environmental feedback signals to complete the sensorimotor loops that control the circular reactions of a developing individual (Piaget, 1963). An effort should be made to correlate the structures that emit the environmental signals with those that receive them, for example, the algebraic properties of motor speech commands with the corresponding properties of auditory feedback patterns (Grossberg, 1978e).

19. Symmetry and Symmetry Breaking in Sensory and Motor Systems

An important theme in the design of adaptive resonances will be the analysis of their symmetry and symmetry-breaking properties. This theme is unavoidable when sensory resonances are studied side by side with their motor counterparts, as Section 18 suggests. For example, the system schematized in Figure 9 shows a manifest asymmetry in the construction of its arousal and pattern analysis components. However, this system forms only one part of a larger system that enjoys a much more symmetric structure in which two subsystems compete, namely an attentional and an orienting subsystem (Lynn, 1966). The component in Figure 9 is part of the attentional system, which also includes incentive motivational and CNV components (Grossberg, 1975). This subsystem focuses attention on cues that are expected to generate prescribed consequences of behavior. It can overshadow irrelevant cues, as in Section 2, by selectively amplifying certain patterns at the expense of others. The complementary orienting system is also capable of selectively amplifying patterns, but these are not the patterns that code for sensory or cognitive events. They are, rather, the motor maps that are capable of directing the subject toward sources of unexpected environmental events.

The dichotomous but interdependent nature of these subsystems is illustrated by the existence of X-cells and Y-cells in mammalian retinas and by the neural pathways that these cells excite. The X-cells project primarily to the LGN, where their signals are processed as visual data, whereas the Y-cells have axons that bifurcate to send branches both to the LGN and the superior colliculus (Fukuda & Stone, 1974; Robson, 1976), which has been identified as an area in which a visuomotor map for eye movements is elaborated (Wurtz & Goldberg, 1972). The competitive nature of these two subsystems is illustrated by considering how different our motor reactions can be when a loud sound to the left is unexpected versus when it is a learned discriminative

cue for rapid button pushing that will be highly rewarded if it is sufficiently rapid. In the former case, our eyes and head rotate rapidly to the left. In the latter case, rotation can be inhibited and supplanted by a rapid button push.

Competition between attentional and orienting subsystems may clarify certain paradoxes about mental illness. As just summarized, the attentional system focuses attention on cues that are expected to generate prescribed consequences of behavior and can thereby overshadow irrelevant cues. The competing system is triggered by unexpected events (novelty) and allows the network to redefine the set of relevant cues to avoid unexpected consequences. Overarousal of either subsystem can yield attentional deficits (Grossberg, 1972c; Grossberg & Pepe, 1970, 1971), but the exact nature of the deficit and its proper treatment depends on the particular subsystem that is overaroused. For example, a schizophreniclike syndrome of reduced attentional span and contextual collapse can be elicited by overarousal of the incentive-motivational system, but would not necessarily be cured by a depressant that acted differentially on the novelty (reticular formation) system. In fact, depressing the wrong arousal system can cause a paradoxical deterioration of a syndrome by disinhibiting the hyperactive competing arousal system that caused the syndrome. Complicating the situation further is the inverted-U in responsiveness that can be caused by parametrically exciting either of the arousal systems separately (Section 10).

Alternation between attentional and orienting reactions seems also to occur, and in a cyclic fashion, within the motor system during the performance of a familiar sequence of skilled movements. Grossberg (1978e, Sections 48–54) used a thought experiment concerning the information available to a behaving infant to derive a minimal network for the learning of circular reactions. A central mechanism in this network is the matching or mismatching of a terminal motor map, or where the end organ expects to go, and a proprioceptive motor map, or where the end organ now is. Proprioceptive-terminal map matching is the analogue within motor systems of expectancy matching in sensory systems (Tanji & Evarts, 1976). Proprioceptive-terminal map matching means that the end organ has reached the location where it expects to be. I suggest that such matching is capable of eliciting signals that not only support the motor postures and perceptual sensitivity needed to pay attention—reflected in the PGO wave—but also release from STM the next motor command in a goal-directed motor sequence. The new motor command instates a new terminal motor map that mismatches the current proprioceptive map, thereby inhibiting the attentional arousal and releasing the new orienting reaction. Thus the matching process seems to cyclically sow the seeds of its own destruction, at least until the entire motor plan is executed. An effort should be made to test whether proprioceptive-terminal matching does indeed elicit signals that reset motor commands in a goal-directed motor plan.

The minimal dimension of the symmetry that is needed to design bilateral hierarchical resonances between competing subsystems is a 16-fold symmetry, since each subsystem contains at least two levels capable of matching their patterns, and each level contains a pair of dipole fields to compute a bilateral resonance.

Despite the greater symmetry that manifests itself by studying competing subsystems side by side, it is inevitable that neural system design will exhibit substantial symmetry breaking. In addition to the asymmetry between excitatory and inhibitory configurations that supports neural development and evolution (e.g., on-center off-surround anatomy), such environmental asymmetries as between light versus dark and between up versus down must be reflected in the neural machinery that has adapted to them. Some insights concerning this neural machinery are suggested in terms of the preceding discussion. For example, if certain off-cells are tonically on in darkness, and if offset of a light triggers a transient output signal from the corresponding off-cell, then why does the tonic activity of this off-cell in the dark not drive a tonic output signal? If the off-field is normalized, then when all the off-cells are on in the dark, none of them is sufficiently active to exceed the output threshold, which is chosen higher than the quenching threshold. After a light is turned off, a particular off-cell's activity is

differentially rebounded for a short time during which its activity exceeds the output threshold. Tonic activity and transient outputs are hereby reconciled. This example illustrates the importance of carefully tuning the relative levels of overall network activity and output threshold.

By contrast, suppose that the output threshold is lowered by disinhibiting the output cells' axon hillocks, or that the overall network activity is enhanced by lowering its quenching threshold—perhaps as in the nucleus reticularis thalami. Then the off-field can deliver tonic output signals to its target cells. If, for example, the target cells control the contraction of muscles, then the tonic muscle signals can maintain a posture that resists the effects of gravity, for example, standing. In this situation, periodic phasic inputs to the on-cells, whether due to external sources or to feedback signals from the off-cells, can cause an oscillatory motor reaction during every cycle of which agonist contraction is followed by an antagonist rebound, for example, walking. Thus, differential tuning of output threshold and normalized activity can convert transient off-cell output signals, as in phasic sensory responses, into tonic off-cell output signals that either balance a persistent asymmetry in environmental influences, as in standing, or energize rhythmic output bursts, as in walking.

20. Cerebral Dominance: The Anatomy of Temporal Versus Spatial Encoding

A more profound type of symmetry breaking occurs between the attentional and orienting subsystems, due to the different nature of cognitive and motor data, and within the attentional subsystem itself, due to the different processing of data about space and time. A pattern of activity across a field of populations at a given time is inherently ambiguous. Does the pattern code a single event in time, such as the features in a visual scene, or does it code the order information in a series of events? Because of this fundamental ambiguity, it is suggested in Grossberg (1978e) that different STM reset mechanisms are needed to reset spatial versus temporal data. The spatial reset mechanism is a matching mechanism such as I have just discussed. The temporal reset mechanism is derived from a study of free recall and serial learning. The output signal from a population in a temporal processor is suggested to activate a self-destructive inhibitory feedback signal. This feedback inhibition prevents perseverative performance of the same item, and conditionalizes the order information among the populations that remain active, with the most active population performed first, since its reaction time for generating an output signal is smallest. The readout of order information from a field of active populations is suggested to be accomplished by either a nonspecific decrease in all the output thresholds or a nonspecific amplification of the total STM activity in the field. Again the relative size of these two levels is a crucial parameter in determining network performance. Thus, the readout of sensory order information is suggested to be mechanistically analogous to the activation of a sequential motor program. By this scheme, a list of items can be performed in a perfect serial ordering despite the fact that all the mechanisms in the network are parallel mechanisms. Serial properties do not imply serial processes.

I suggest that the cortical microanatomy that subserves spatial versus temporal processing will be found to exhibit these different STM reset mechanisms. Consequently, to unambiguously decode temporal versus spatial data, somehow the populations that code the different types of data must be spatially segregated so that they can be endowed with their disparate STM reset mechanisms. The ambiguous meaning of spatial patterns hereby suggests the need to spatially segregate the processing of sequential, including language-like, codes from codes concerning themselves with spatial integration. This dichotomy might be one reason for the emergence of cerebral dominance (Gazzaniga, 1970, chap. 8), despite the fact that a typical speech act can include both spatial and temporal coding elements, and thus requires cyclic resetting of both types of codes. Visual and auditory processing are sensory prototypes of higher codes that emphasize spatial and temporal processing, respectively. Since visual and auditory representations are bilateral, the trend toward segregation of spatial versus temporal process-

ing in separate hemispheres can be viewed as a symmetry-breaking operation with a drift of visuallike processing into the nondominant hemisphere and auditorylike processing into the dominant hemisphere. The symmetry between bilateral resonances in these regions should be correspondingly broken, leading to a generalized avalanche or command structure between the two hemispheres to coordinate the temporal unfolding of spatial representations. An effort should be made to test whether the cortical microanatomy in spatial versus temporal processors exhibits traces of different reset mechanisms in the anatomy of inhibitory feedback interneurons.

21. Conclusion: How to Understand Evolutionary Data?

The thought experiment in this article illustrates a general method for discovering the mechanisms behind psychological data. Many psychological phenomena are facets of the evolutionary process—variously called chunking, unitization, or automation—whereby behavioral fragments are grouped into new control units that become the fragments of still higher behavioral units in a continuing process of hierarchical organization and command synthesis. By its very nature, this evolutionary process hides the mechanistic substrate on which it is built, so that we can behave in a world of percepts, feelings, and plans rather than of cells, signals, and transmitters. Because our brains are these evolutionary devices, we have immediate introspective evidence about basic psychological processes, and can consensually define concepts like reward, punishment, frustration, expectation, memory, and plan even without a scientific understanding of their mechanistic substrates. To represent these consensual concepts in our scientific work by processes that mirror their introspective properties is, however, a fundamental mistake. Then the consensual impression of events blinds us to their functional representation.

For example, language processes whose properties seem discrete and serial are often realized by continuous and parallel control processes (Grossberg, 1978a, 1978e). The two types of representation are not fundamentally equivalent and generate different predictions. Similarly, behavioral properties that seem linear are often controlled by nonlinear processes (Grossberg, 1978d). Again the two types of description are fundamentally not equivalent. When a theory is erroneously built on consensual properties, it soon meets data that it finds paradoxical. Then the theory either collapses or is decorated with a succession of ad hoc hypotheses. Theoretical epicycles soon crowd the scientific landscape, and theory gets a bad name even though we cannot live without it.

An alternative procedure is to respect the wisdom of evolution by trying to imitate it. To do this, at each stage of theory construction, prescribed environmental pressures are identified that force adaptive designs on the behaving brain. Most of us know these pressures; they are familiar precisely because they are among the constraints to which we have successfully adapted. Thus the theory is grounded on a firm basis. By contrast with the consensual method, these pressures are properties of the environment rather than of our behavior. The thought experiments show how these environmental constraints generate explicit minimal mechanisms for coping with them. Such experiments include information that eludes experimental techniques for several reasons. For example, they show how many system components work together, and they compress into a unified description environmental pressures that act over long, or at least nonsimultaneous, times. Most importantly, the thought experiments explicate design constraints that are needed to adapt in a real-time setting. These real-time constraints are often the most crucial ones, and they are invisible to descriptive or purely formal theories.

Once the minimal mechanisms that realize several environmental pressures are constructed, mathematical analysis shows how they work together to generate data and predictions whose complexity and subtlety transcend the apparent simplicity of the environmental pressures, as well as unaided intuition. This procedure defines new conceptual categories into which to divide the data, and also points to important environmental pressures that have been overlooked,

by clearly delineating what the mechanisms can and cannot do. In this way, a small number of principles and mechanisms is organized in an evolutionary progression, and large bodies of data are hierarchically grouped as manifestations of these principles.

In the present article a thought experiment shows how limitations in the types of information available to individual cells can be overcome when the cells act together in suitably designed feedback schemes. The explication of these schemes in a rigorous setting (see the appendices) forces us to study a series of general design problems whose complete solution includes many examples that go beyond the thought experiment; for example, competitive systems (their decision schemes, self-tuning, adaptation, fast pattern transformations, and STM), nonstationary prediction systems (their filtering, pattern learning, and LTM), dipole systems (their transduction and rebound properties), and resonant systems (their hysteresis, deformation, and reset properties). This thought experiment is just one in a series that has helped to unravel psychological mechanisms and to generate as yet untested predictions.

An early thought experiment used the simplest classical conditioning postulates, interpreted in real time (see Grossberg, 1974, for a review), to derive explicit neural networks. When, for example, these networks are exposed serially to long lists, a variety of serial learning properties automatically occur, such as bowing, skewing, anchoring, primacy dominating recency, anticipatory and perseverative generalization gradients, and response oscillation (Grossberg, 1969b; Grossberg & Pepe, 1970, 1971). In addition, mathematical analysis unexpectedly showed how overarousal can cause an attentional deficit with reduced attentional span and collapsed contextual constraints. This overaroused syndrome includes a change toward less skewing of the bowed error curve and toward recency dominating primacy. These formal properties have not yet been empirically tested.

Using these results on classical conditioning, another thought experiment about classical conditioning became necessary. The time intervals between CS and UCS presentations on successive learning trials are not always the same. In a real-time theory, this trivial fact creates a severe synchronization problem whose solution unexpectedly led to explicit mechanisms of instrumental conditioning (Grossberg, 1971a, 1972b, 1972c). Many insights about instrumental mechanisms and their relationship to Pavlovian mechanisms were hereby derived. One of them is especially pertinent to this article. A dipole mechanism was forced on the theory to control net incentive motivation through time. Mathematical analysis of the dipole revealed several unexpected properties (Sections 10 and 11) including the ability of arousal, and hence of unexpected events, to adapt or rebound the dipole. The detailed understanding of dipole dynamics helped to clarify many novelty-related phenomena, such as learned helplessness, superconditioning, and vicious circle behavior. It also forced on the theory the realization that cognitive events, via expectancy matching, can directly influence reinforcement, via the dipole. In summary, a simple environmental pressure concerning a real-time synchronization problem in classical conditioning was solved by mechanisms of instrumental conditioning and led to a role for cognitive processing in the direct evaluation of reinforcement.

With these results in hand, a thought experiment about feature fields came into view. The parallel activation of many cells by external cues can easily destroy decision rules that regulate the balance of net incentive through time. The minimal solution of this difficulty is to impose a normalization property at the processing stages where cues are stored in STM (Grossberg, 1972c). This normalization property had already been noticed as a property by which competitive shunting networks solve the saturation problem (Grossberg, 1970). These results from reinforcement theory made it clear that further progress concerning feature extraction and related perceptual phenomena required a frontal attack on the mathematics of competitive systems. The early results in this direction (Grossberg, 1973) eventually led to many surprising properties, the most general being that every competitive system induces a decision scheme that can be used to predict its behavior through time (Grossberg, 1978c). For present

purposes, the normalization and quenching threshold properties are particularly important, since they show how arousal can tune STM, and thereby help to control what cues are overshadowed vs. what cues are processed. Another role for cognitive events, again acting on arousal via expectancies, was hereby discerned.

Once the normalization and quenching threshold properties were discovered, a thought experiment was suggested that joins together facets of perceptual and motivational processing: How can cues with incompatible motivational consequences be processed in parallel without causing chaotic cross-conditioning (Figure 2)? This thought experiment showed how incentive motivational feedback can influence STM storage to yield stable self-consistent coding and, as side benefits, explanations of attentional data such as overshadowing and discrimination-learning data such as peak shift and behavioral contrast (Grossberg, 1975). Several other theoretical stages then followed as the attentional phenomena were recognized to be special cases of the resonance idea. It became possible to build a theory of stable code development (Grossberg, 1976a, 1976b), which, in turn, suggested a psychophysiological foundation for cognitive theory (Grossberg, 1978e), one of whose facets is heuristically summarized by the present thought experiment.

The evolutionary procedure thus embodies a program of real-time theory construction in psychological studies that underscores the need to understand the collective properties of hierarchically organized nonlinear neural networks. Because the rigorous analysis of such networks is well under way, we can anticipate an emergent resonance between experimental psychology and psychophysiological theory during our generation.

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.

Arbib, M. A. *The metaphorical brain*. New York: Wiley, 1972.

Atkinson, R. C., & Shiffrin, R. M. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *Advances in the psychology of learning and motivation research and theory* (Vol. 2). New York: Academic Press, 1968.

Atkinson, R. C., & Shiffrin, R. M. The control of short-term memory. *Scientific American*, 1971, *225*, 82–90.

Barlow, H. B., & Levick, W. R. Changes in the maintained discharge with adaptation level in the cat retina. *Journal of Physiology*, 1969, *202*, 699–718. (a)

Barlow, H. B., & Levick, W. R. Three factors limiting the reliable detection of light by retinal ganglion cells of the cat. *Journal of Physiology*, 1969, *200*, 1–24. (b)

Beck, J. *Surface color perception*. Ithaca, N.Y.: Cornell University Press, 1972.

Berger, T. W., & Thompson, R. F. Limbic system interrelations: Functional division among hippocampal–septal connections. *Science*, 1977, *197*, 587–589.

Blakemore, C., & Cooper, G. F. Development of the brain depends on the visual environment. *Nature*, 1970, *228*, 477–478.

Bloomfield, T. M. Behavioral contrast and the peak shift. In R. M. Gilbert & N. S. Sutherland (Eds.), *Animal discrimination learning*. New York: Academic Press, 1969.

Boring, E. G. *A history of experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts, 1950.

Brown, J. L. Afterimages. In C. H. Graham (Ed.), *Vision and visual perception*. New York: Wiley, 1965.

Campbell, B. A. Interaction of aversive stimuli: Summation or inhibition? *Journal of Experimental Psychology*, 1968, *78*, 181–190.

Campbell, B. A., & Kraeling, D. Response strength as a function of drive level and amount of drive reduction. *Journal of Experimental Psychology*, 1953, *45*, 97–101.

Campbell, F. W., & Howell, E. R. Monocular alternation: A method for the investigation of pattern vision. *Journal of Physiology*, 1972, *225*, 19–21.

Campbell, L., & Garnett, W. *The life of James Clerk Maxwell*. London: Macmillan, 1882.

Cant, B. R., & Bickford, R. G. The effect of motivation on the contingent negative variation (CNV). *Electroencephalography and Clinical Neurophysiology*, 1967, *23*, 594.

Carterette, E. C., & Friedman, M. P. (Eds.). *Handbook of perception: Seeing* (Vol. 5). New York: Academic Press, 1975.

Chapman, R. M., McCrary, J. W., & Chapman, J. A. Short-term memory: The "storage" component of human brain response predicts recall. *Science*, 1978, *202*, 1211–1214.

Cohen, J. Very slow brain potentials relating to expectancy: The CNV. In E. Donchin & D. B. Lindsley (Eds.), *Average evoked potentials*. Washington, D.C.: National Aeronautics and Space Administration, 1969.

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

Craik, F. I. M. & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 671–684.

Craik, F. I. M., & Tulving, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 1975, *104*, 268–294.

Creutzfeldt, O. D., & Northdurft, H. C. Representation of complex visual stimuli in the brain. *Naturwissenschaften*, 1978, *65*, 307–318.

Daniels, J. D., & Pettigrew, J. D. Development of neuronal responses in the visual system of cats. In G. Gottlieb (Ed.), *Neural and behavioral specificity* (Vol. 3). New York: Academic Press, 1976.

Denny, M. R. Relaxation theory and experiments. In F. R. Brush (Ed.), *Aversive conditioning and learning*. Academic Press: New York, 1970.

Dickinson, A., Hall, G., & Mackintosh, N. J. Surprise and the attenuation of blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, 1976, *4*, 313–322.

Donchin, E., Gerbrandt, L. A., Leifer, L., & Tucker, L. Is the contingent negative variation contingent on a motor response? *Psychophysiology*, 1972, *9*, 178–188.

Donchin, E., Tueting, P., Ritter, W., Kutas, M., & Heffley, E. *Electroencephalography and Clinical Neurophysiology*, 1975, *38*, 1–13.

Duda, R. O., & Hart, P. E. *Pattern classification and scene analysis*. New York: Wiley, 1973.

Dunham, P. J. Punishment: Method and theory. *Psychological Review*, 1971, *78*, 58–70.

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.

Estes, W. K. Outline of a theory of punishment. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts, 1969.

Estes, W. K., & Skinner, B. F. Some quantitative properties of anxiety. *Journal of Experimental Psychology*, 1941, *29*, 390–400.

Fender, D., & Julesz, B. Extension of Panum's fusional area in binocularly stabilized vision. *Journal of the Optical Society of America*, 1967, *57*, 819–830.

Foote, W. E., Manciewicz, R. J., & Mordes, J. P. Effect of midbrain raphe and lateral mesencephalic stimulation on spontaneous and evoked activity in the lateral geniculate of the cat. *Experimental Brain Research*, 1974, *19*, 124–130.

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

Fukuda, Y., & Stone, J. Retinal distribution and central projections of X-, Y-, and W-cells of the cat's retina. *Journal of Neurophysiology*, 1974, *37*, 749–772.

Fuxe, K., Hökfelt, T., & Ungerstedt, U. Morphological and functional aspects of central monoamine neurons. *International Review of Neurobiology*, 1970, *13*, 93–126.

Fuxe, K., & Ungerstedt, U. Histochemical, biochemical, and functional studies on central monoamine neurons after acute and chronic amphetamine administration. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.

Gardner, W. J., Licklider, J. C. R., & Weisz, A. Z. Suppression of pain by sound. *Science*, 1961, *132*, 32–33.

Gazzaniga, M. S. *The bisected brain*. New York: Appleton-Century-Crofts, 1970.

Gibson, J. J. Adaptation with negative aftereffect. *Psychological Review*, 1937, *44*, 222–244.

Graham, C. H. Visual form perception. In C. H. Graham (Ed.), *Vision and visual perception*. New York: Wiley, 1965.

Grossberg, S. *The theory of embedding fields with applications to psychology and neurophysiology*. New York: Rockefeller Institute for Medical Research, 1964.

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

Grossberg, S. On the serial learning of lists. *Mathematical Biosciences*, 1969, *4*, 201–253. (b)

Grossberg, S. On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, 1969, *22*, 325–364. (c)

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

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

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

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. Qualitative theory. *Mathematical Biosciences*, 1972, *15*, 39–67. (b)

Grossberg, S. A neural theory of punishment and avoidance, II. Quantitative theory. *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. A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology*, 1975, *18*, 263–327.

Grossberg, S. Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, 1976, *23*, 121–134. (a)

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

Grossberg, S. Pattern formation by the global limits of a nonlinear competitive interaction in n dimensions. *Journal of Mathematical Biology*, 1977, *4*, 237–256.

Grossberg, S. Behavioral contrast in short-term memory: Serial binary memory models or parallel continuous memory models? *Journal of Mathematical Psychology*, 1978, *17*, 199-219. (a)

Grossberg, S. Communication, memory, and development. In R. Rosen & F. Snell (Eds.), *Progress in theoretical biology* (Vol. 5). New York: Academic Press, 1978. (b)

Grossberg, S. Decisions, patterns, and oscillations in the dynamics of competitive systems with applications to Volterra-Lotka systems. *Journal of Theoretical Biology*, 1978, *73*, 101-130. (c)

Grossberg, S. Do all neural models really look alike? A comment on Anderson, Silverstein, Ritz, and Jones. *Psychological Review*, 1978, *85*, 592-596. (d)

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. (e)

Grossberg, S., & Levine, D. S. Some developmental and attentional biases in the contrast enhancement and short term memory of recurrent neural networks. *Journal of Theoretical Biology*, 1975, *53*, 341-380.

Grossberg, S., and Pepe, J. Schizophrenia: Possible dependence of associational span, bowing, and primacy vs. recency on spiking threshold. *Behavioral Science*, 1970, *15*, 359-362.

Grossberg, S., & Pepe, J. Spiking threshold and overarousal effects on serial learning. *Journal of Statistical Physics*, 1971, *3*, 95-125.

Hebb, D. O. Drives and the CNS (conceptual nervous system). *Psychological Review*, 1955, *62*, 243-254.

Helmholtz, H. von. *Handbuch der physiologischen optik* (1st. ed.). Hamburg, Leipzig: Voss, 1866.

Helmholtz, H. von. *Physiological optics* (Vol. 2) (J. P. C. Southall, Ed.). New York: Dover, 1962.

Hilgard, E. R., & Bower, G. H. *Theories of learning* (4th ed.). Englewood Cliffs, N.J.: Prentice-Hall, 1975.

Hirsch, H. V. B., & Spinelli, D. N. Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cats. *Science*, 1970, *168*, 869-871.

Hubel, D. H., & Wiesel, T. N. Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London* (B), 1977, *198*, 1-59.

Irwin, F. W. *Intentional behavior and motivation: A cognitive theory*. Philadelphia, Pa.: Lippincott, 1971.

Irwin, D. A., Rebert, C. S., McAdam, D. W., & Knott, J. R. Slow potential change (CNV) in the human EEG as a function of motivational variables. *Electroencephalography and Clinical Neurophysiology*, 1966, *21*, 412-413.

Jacobowitz, D. M. Effects of 6-hydroxydopa. In E. Usdin & H. S. Snyder (Eds.), *Frontiers in catecholamine research*. New York: Pergamon Press, 1973.

Juhasz, A. Über die komplementärge-färbten nachbilder. *Zeitschrift fur Psychologie*, 1920, *51*, 233-263.

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

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

Koenigsberger, L. *Hermann von Helmholtz*. (F. A. Welby, trans.). Oxford, England: Clarendon, 1906.

Ladisich, W., Volbehr, H., & Matussek, N. Paradoxical effect of amphetamine on hyperactive states in correlation with catecholamine metabolism in brain. In E. Costa & S. Garattini (Eds.), *Amphetamines and related compounds*. New York: Raven Press, 1970.

Land, E. H. The retinex theory of color vision. *Scientific American*, 1977, *237*, 108-128.

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

Lindvall, O., & Björklund, A. The organization of the ascending catecholamine neuron systems in the rat brain as revealed by the glyoxylic acid fluorescence method. *Acta Physiologia Scandinavia Supplement*, 1974, *412*, 1-48.

Low, M. D., Borda, R. P., Frost, J. D., & Kellaway, P. Surface negative slow potential shift associated with conditioning in man. *Neurology*, 1966, *16*, 711-782.

Lynn, R. *Attention, arousal, and the orientation reaction*. New York: Pergamon Press, 1966.

Macchi, G., & Rinvik, E. Thalamo-telencephalic circuits: A neuroanatomical survey. In A. Rémond (Ed.), *Handbook of electroencephalography and clinical neurophysiology* (Vol. 2, Pt. A). Amsterdam: Elsevier, 1976.

MacKay, D. M. Moving visual images produced by regular stationary patterns. *Nature*, 1957, *180*, 849-850.

MacKay, D. M., & MacKay, V. What causes decay of pattern-contingent chromatic aftereffects? *Vision Research*, 1975, *15*, 462-464.

Masterson, F. A. Is termination of a warning signal an effective reward for the rat? *Journal of Comparative and Physiological Psychology*, 1970, *72*, 471-475.

McAdam, D. W. Increases in CNS excitability during negative cortical slow potentials in man. *Electroencephalography and Clinical Neurophysiology*, 1969, *26*, 216-219.

McAdam, D. W., Irwin, D. A., Rebert, C. S., & Knott, J. R. Conative control of the contingent negative variation. *Electroencephalography and Clinical Neurophysiology*, 1966, *21*, 194-195.

McAllister, W. R., & McAllister, D. E. Behavioral measurement of conditioned fear. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press, 1970.

McCollough, C. Color adaptation of edge-detectors in the human visual system. *Science*, 1965, *149*, 1115-1116.

Montalvo, F. S. A neural network model of the McCollough effect. *Biological Cybernetics*, 1976, *25*, 49-56.

Moruzzi, G., & Magoun, H. W. Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, 1949, *1*, 455-473.

Myers, A. K. Effects of continuous loud noise during instrumental shock-escape conditioning. *Journal of*

Comparative and Physiological Psychology, 1969, *68*, 617–622.

Piaget, J. *The origins of intelligence in children*. New York: Norton, 1963.

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

Rauschecker, J. P. J., Campbell, F. W., & Atkinson, J. Colour opponent neurones in the human visual system. *Nature*, 1973, *245*, 42–45.

Remington, R. J. Analysis of sequential effects in choice reaction times. *Journal of Experimental Psychology*, 1969, *2*, 250–257.

Rescorla, R. A., & Wagner, A. R. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. F. Prokasy (Eds.), *Classical conditioning II*. New York: Appleton-Century-Crofts, 1972.

Ricklan, M. *L-dopa and parkinsonism: A psychological assessment*. Springfield, Ill.: Charles C Thomas, 1973.

Robson, J. G. Receptive fields: Neural representation of the spatial and intensive attributes of the visual image. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 5). New York: Academic Press, 1976.

Rohrbaugh, J., Donchin, E., & Eriksen, C. Decision making and the P300 component of the cortical evoked response. *Perception & Psychophysics*, 1974, *15*, 368–374.

Schiller, P. H., & Malpeli, J. G. Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. *Journal of Neurophysiology*, 1978, *41*, 788–797.

Schneider, W., & Shiffrin, R. M. Automatic and controlled information processing in vision. In D. LaBarge & S. J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension*. Hillsdale, N.J.: Erlbaum, 1976.

Seligman, M. E. P., Maier, S. F., & Solomon, R. L. Unpredictable and uncontrollable aversive events. In F. R. Brush (Ed.), *Aversive conditioning and learning*. New York: Academic Press, 1971.

Singer, W. Control of thalamic transmission by corticofugal and ascending reticular pathways in the visual system. *Physiological Review*, 1977, *57*, 386–420.

Sluckin, W. *Imprinting and early learning*. London: Methuen, 1964.

Squires, K., Wickens, C., Squires, N., & Donchin, E. The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 1976, *193*, 1142–1146.

Stein, L. Norepinephrine reward pathways: Role in self-stimulation, memory consolidation, and schizophrenia. In J. K. Cole & T. B. Sonderegger (Eds.), *Nebraska Symposium on Motivation* (Vol. 22). Lincoln: University of Nebraska Press, 1974.

Stryker, M., & Sherk, H. Modification of cortical orientation selectivity in the cat by restricted visual experience: A reexamination. *Science*, 1975, *190*, 904–905.

Tanji, J., & Evarts, E. V. Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. *Journal of Neurophysiology*, 1976, *39*, 1062–1068.

Thomas, G. B., Jr. *Calculus and analytic geometry*. Reading, Mass.: Addison-Wesley, 1968.

Trabasso, T., & Bower, G. H. *Attention in learning: Theory and research*. New York: Wiley, 1968.

Tsumoto, T., Creutzfeldt, O. D., & Legéndy, C. R. Functional organization of the corticofugal system from visual cortex to lateral geniculate nucleus in the cat. *Experimental Brain Research*, 1978, *32*, 345–364.

Tsumoto, T., & Suzuki, D. A. Effects of frontal eye field stimulation upon activities of the lateral geniculate body of the cat. *Experimental Brain Research*, 1976, *25*, 291–306.

Ungerstedt, U. Stereotoxic mapping of the monoamine pathways in the rat brain. *Acta Physiologica Scandinavia*, 1971, *82* (Supplement 367), 1–48.

Wagner, A. R. Frustrative nonreward: A variety of punishment. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior*. New York: Appleton-Century-Crofts, 1969.

Walter, W. G. Slow potential waves in the human brain associated with expectancy, attention, and decision. *Arch. Psychiat. Nervenkr.*, 1964, *206*, 309–322.

Wilson, H. A synaptic model for spatial frequency adaptation. *Journal of Theoretical Biology*, 1975, *50*, 327–352.

Wurtz, R. H., & Goldberg, M. E. The role of the superior colliculus in visually evoked eye movement. In J. Dichgans & E. Bizzi (Eds.), *Cerebral control of eye movement and motion perception*. Basel, Switzerland: Karger, 1972.

Appendix A

This section summarizes some of the mechanisms whereby an activity pattern across $F^{(1)}$ elicits signals to $F^{(2)}$ that filter the pattern. The filtered signal pattern is then rapidly transformed by recurrent competitive interactions within $F^{(2)}$ before the resultant pattern is stored in STM. The STM pattern endures long enough to alter the interfield path strengths that define the filter. This is an LTM change. Then the process repeats itself until STM and LTM equilibrate.

Filter

Denote the cells of $F^{(1)}$ by $v_i^{(1)}$, $i = 1, 2, \ldots, n$, and the cells of $F^{(2)}$ by $v_j^{(2)}$, $j = 1, 2, \ldots, N$. Let the activity of $v_i^{(1)}$ at time t be $x_i^{(1)}(t)$. Suppose that the activity $x_i^{(1)}(t)$ elicits a signal $S_i = S_i(x_i^{(1)}(t))$ in the pathways from $v_i^{(1)}$ to $F^{(2)}$. Let the net signal from $v_i^{(1)}$ to $v_j^{(2)}$ be $S_i z_{ij}$, where z_{ij} provides a measure of the efficiency of the pathway e_{ij} from $v_i^{(1)}$ to $v_j^{(2)}$. In other words, z_{ij} *gates* signal S_i on its way to $v_j^{(2)}$. Then the *total* signal from $F^{(1)}$ to $v_j^{(2)}$ is

$$T_j = \sum_{i=1}^{n} S_i z_{ij}.$$

This equation for T_j has an informative geometrical interpretation in terms of the vectors $S = (S_1, S_2, \ldots, S_n)$ of signals and $z_j = (z_{1j}, z_{2j}, \ldots, z_{nj})$ of path strengths from $F^{(1)}$ to $v_j^{(2)}$. Namely, T_j is the dot product, or inner product, of S and z_j, which is written $T_j = S \cdot z_j$ (Thomas, 1968). The dot product can be evaluated in terms of the vector lengths

$$\|S\| = \sqrt{\sum_{i=1}^{n} S_i^2}$$

and

$$\|z_j\| = \sqrt{\sum_{i=1}^{n} z_{ij}^2},$$

and the cosine of the angle between S and z_j by the formula $T_j = \|S\| \, \|z_j\| \cos(S, z_j)$. In particular, if all $\|z_j\|$ are equal, then the cell $v_j^{(2)}$ in $F^{(2)}$ that receives the largest signal is the cell whose $\cos(S, z_j)$ is maximal. The cosine can be increased by choosing the coefficients of z_j more proportional, or parallel, to S, and can be decreased by choosing the coefficients more perpendicular, or orthogonal, to S. Thus each z_j *filters* S by producing a net signal T_j whose size depends on how parallel z_j is to S. Otherwise expressed, z_j *projects* S onto $v_j^{(2)}$. The pattern $T = (T_1, T_2, \ldots, T_N)$ of inputs to $F^{(2)}$ represents pattern S by projecting S onto all the cells $F^{(2)}$ with relative input sizes that depend on the choice of all the vectors $z_1, z_2, \ldots, z_N$.

Contrast Enhancement

When S is first presented to $F^{(1)}$, the pattern T of inputs that it elicits across $F^{(2)}$ might be approximately uniform. Recurrent on-center off-surround interactions within $F^{(2)}$ rapidly contrast enhance this input pattern in order to produce a sharper pattern of STM activities across $F^{(2)}$ (Grossberg, 1976a, 1976b). I illustrate this concept with the simplest case: Suppose that $F^{(2)}$ can choose the cell whose initial input is maximal for storage in STM. Denoting the activity of $v_j^{(2)}$ by $x_j^{(2)}$, this law says that

$$\begin{aligned} x_j^{(2)} &= 1 \quad \text{if} \quad T_j > \max\{\epsilon, T_k, k \neq j\} \\ x_j^{(2)} &= 0 \quad \text{if} \quad T_j < \max\{\epsilon, T_k, k \neq j\}. \end{aligned} \tag{A1}$$

The coefficient ϵ designates a quenching threshold that must be exceeded before any STM storage is possible. Suppose for definiteness that $T_1 > \max\{\epsilon, T_k, k \neq 1\}$. Then the activity of $v_1^{(2)}$ is rapidly contrast enhanced and stored in STM, whereas all other activities across $F^{(2)}$ are suppressed.

Coding

The path strengths z_{ij} are LTM traces that can slowly adapt to the signal pattern S from $F^{(1)}$ and the STM pattern across $F^{(2)}$. In the simplest case, z_{ij} changes only if $x_j^{(2)} > 0$. Then

$$\frac{d}{dt} z_{ij} = (-z_{ij} + S_i) x_j^{(2)}. \tag{A2}$$

For example, if $x_1^{(2)} = 1$ and all $x_j^{(2)} = 0$, $j \neq 1$, then this LTM law causes the signal $T_1 = S \cdot z_1$ to be maximized as S is practiced by making z_1 become parallel to S. In this way, presentation of S at $F^{(1)}$ can induce a differentiated STM pattern across $F^{(2)}$ by changing the LTM vectors $z_1, z_2, \ldots, z_N$.

Grossberg's (1976a, 1976b, 1978b, 1978e) articles describe these mechanisms in greater detail. They also show how to generalize the mechanisms to include more complex STM and LTM interactions. Despite these generalizations, the mechanisms are there shown to be unstable in a complex input environment. A precise understanding of this difficulty forced the use of learned feedback expectancies.

Appendix B

This section indicates how a single cell population in $F^{(2)}$ can learn a spectial pattern of activity across $F^{(1)}$. Analogous arguments then show how many simultaneously active cell populations across $F^{(2)}$ can learn a spatial pattern across $F^{(1)}$, albeit not necessarily the same spatial pattern that would have excited a single cell in $F^{(2)}$ by interfield signaling from $F^{(1)}$ to $F^{(2)}$.

Associative Learning

Our laws for associative learning appeared in a monograph by Grossberg (1964), and were mathematically analyzed in a series of articles, leading to a universal theorem of associative learning in Grossberg (1969a, 1971b, 1972d). The universal theorem proves that if these associative learning laws were invented at a prescribed time during the evolutionary process, then they could be used to guarantee unbiased associative learning in essentially any later evolutionary specialization. That is, the laws are capable of learning arbitrary spatial patterns in arbitrarily many, simultaneously active sampling channels that are activated by arbitrary continuous data preprocessing in an essentially arbitrary anatomy. Learning of arbitrary space-time patterns is also guaranteed given modest requirements on the temporal regularity of stimulus sampling. (See Grossberg, 1974, for a review.) Herein I summarize the fact that the unit of LTM is a *spatial pattern*. This is done by considering the *minimal* anatomy that is capable of classical conditioning.

STM and LTM Laws That Factor Pattern from Activity

Let presentation of a CS create an input $I_0(t)$ that activates the cell population v_0. Let the UCS create an input pattern $(I_1(t), I_2(t), \ldots, I_n(t))$ that activates the cell populations $v_1, v_2, \ldots, v_n$, whose outputs elicit the UCR. Let the STM trace of v_i be $x_i(t)$, $j = 0, 1, \ldots, n$, and let the LTM trace of the axon pathway e_{0i} from v_0 to v_i be $z_{0i}(t)$, $i = 1, 2, \ldots, n$ (Figure A1). Suppose that the STM and LTM traces obey the laws

$$\frac{d}{dt}x_0 = -A_0x_0 + I_0(t) \qquad \text{(A3)}$$

$$\frac{d}{dt}x_i = -Ax_i + Bz_{0i} + I_i(t), \qquad \text{(A4)}$$

and

$$\frac{d}{dt}z_{0i} = -Cz_{0i} + Dx_i, \qquad \text{(A5)}$$

$i = 1, 2, \ldots, n$. The terms A_0 and A are STM decay rates. The term C is the LTM decay rate. The terms B and D are signals from v_0 along all the pathways e_{0i}, $i = 1, 2, \ldots, n$; for example, $B(t) = f(x_0(t - \tau))$, where $f(w)$ is a sigmoid function of w. The LTM trace z_{0i} is computed at the interface of the synaptic knob S_{0i} (at the end of e_{0i}) and the postsynaptic cell v_i—that is, at the synaptic knob and/or postsynaptic membrane —where it can gate the signals B on their way to v_i, as in term Bz_{0i} of Equation A4, and simultaneously time-average (term $-Cz_{0i}$) the product of signals D and postsynaptic STM trace x_i (term Dx_i), in A5. In particular, A2 is a special case of A5.

A *spatial pattern* is a UCS whose *relative* activities remain fixed, even though their absolute activities can fluctuate through time,

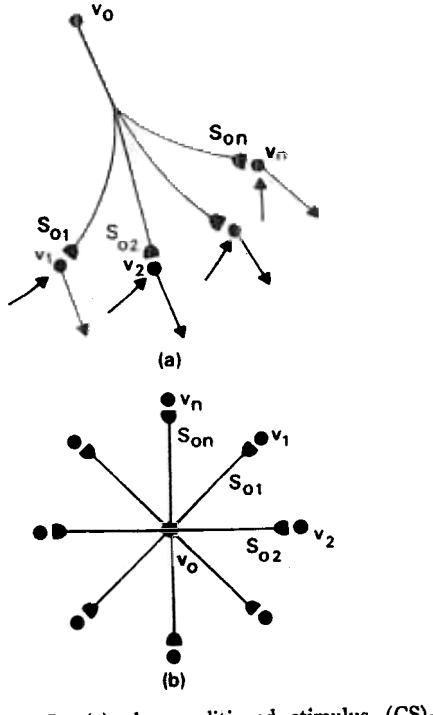

Figure A1. In (a) the conditioned stimulus (CS)-activated population v_0 samples the unconditioned stimulus (UCS)-activated populations $v_1, v_2, \ldots, v_n$; in (b) the *outstar* is the minimal network capable of classical conditioning.

namely, $I_i(t) = \theta_i I(t)$, $i = 1, 2, \ldots, n$, where θ_i is the fixed relative activity and $I(t)$ is the total UCS activity. The convention

$$\sum_{i=1}^{n} \theta_i = 1$$

guarantees the normalization

$$I(t) = \sum_{i=1}^{n} I_i(t).$$

The relative values $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ are like generalized "reflectances" that carry the information in the UCS pattern, whereas $I(t)$ provides the UCS activity that drives system changes in response to θ. It is shown below how this system, which I call an *outstar* (Figure A1), can factorize pattern information θ from information about total activity $I(t)$. This property has many important implications. For example, θ is a probability distribution, since each $\theta_i \geq 0$ and

$$\sum_{k=1}^{n} \theta_k = 1$$

The system learns probabilities despite the fact that it can generate deterministic behavior. There exists a type of "wave-particle" dualism in these systems that helps to explain the partial successes of statistical learning models, and provides an interesting vantage point from which to think about the wave-particle dualism of quantum theory. Also, since there is no evolutionary advantage in perceptually discriminating data that cannot, in principle, be learned, we can expect the neural perceptual apparatus also to process spatial patterns. The brightness and hue constancies of vision illustrate this fact. These observations clarified how perceptual and learning mechanisms are matched to each other, and suggested study of the minimal neural networks that are capable of discriminating a spatial pattern θ; that is, reflectances. Some of these networks were constructed in Grossberg (1970, 1972a) and, not surprisingly, have an anatomy that is remarkably retinal.

System A3–A5 *factorizes* θ and $I(t)$ in the following sense. Equation A3 can be explicitly solved for $x_0(t)$ by integration, and the result used to solve for $B(t)$ and $D(t)$ as functions of time t. Then A4 and A5 can be rewritten in terms of the relative STM traces

$$X_i = x_i\left(\sum_{k=1}^{n} x_k\right)^{-1}$$

and relative LTM traces

$$Z_i = z_{0i}\left(\sum_{k=1}^{n} z_{0k}\right)^{-1}$$

as follows:

$$\frac{d}{dt}X_i = E(Z_i - X_i) + F(\theta_i - X_i) \quad \text{(A6)}$$

and

$$\frac{d}{dt}Z_i = G(X_i - Z_i). \quad \text{(A7)}$$

The coefficients E, F, and G depend only on $I(t)$, on the total STM activity

$$x = \sum_{k=1}^{n} x_k,$$

and on the total LTM activity

$$z = \sum_{k=1}^{n} z_{0k}.$$

By A4 and A5,

$$\frac{d}{dt}x = -Ax + Bz + I \quad \text{(A8)}$$

and

$$\frac{d}{dt}z = -Cz + Dx. \quad \text{(A9)}$$

Equations A8 and A9 are independent of θ; they depend only on the total activity $I(t)$. These equations *decouple* total activity data (I, x, z) from pattern data (θ, X, Z), where $X = (X_1, X_2, \ldots, X_n)$ and $Z = (Z_1, Z_2, \ldots, Z_n)$ are also probability distributions. The total activity data influence the pattern data only via the coefficients E, F, and G, which are always nonnegative. No matter how wildly the CS input $I_0(t)$ and the UCS input $I(t)$ oscillate through time, these coefficients influence only the *rates* with which X and Z are influenced by θ, but not the *directions* in which X and Z can change in response to θ. It is this property that generalizes to yield the universal theorem cited above.

In particular, term $F(\theta_i - X_i)$ in A6 says that X_i approaches θ_i as learning proceeds (UCS read into STM). Term $E(Z_i - X_i)$ in A6 says X_i approaches Z_i (readout of LTM into STM). The net effect of these two terms shows how present demands of the UCS, expressed via θ, and past memories, expressed via Z, compete to change STM via X. Equation A7 shows that Z_i approaches X_i (transfer from STM to LTM). As X approaches θ, and Z approaches X, Z learns the spatial

pattern θ. On later performance trials, a CS input to v_0 activates x_0, which in turn activates the signal B. Signal B reads the pattern Z into STM via the terms Bz_{0i} in A4. Since $Z \cong \theta$, A4 shows that the x_is that are activated in this fashion are proportional to the θ_is, as desired.

Many aspects of associative learning can be understood using these STM and LTM laws in more complex anatomies. In particular, the Z_is are stimulus sampling probabilities whose properties explain in a neural setting the partial successes of statistical learning models. The distributions of STM and LTM traces also mimic and predict various data about serial learning, paired associate learning, and free recall experiments. See Grossberg (1974, 1978a, 1978e) for additional discussion.

Appendix C

This section summarizes how feedforward competitive interactions solve the saturation problem using automatic gain control by inhibitory signals, and how properties such as noise suppression, pattern matching, edge enhancement, and spatial frequency sensitivity follow as special cases.

Noise-Saturation Dilemma

All cellular systems face the following dilemma. If their inputs are too small, they can get lost in noise. If the inputs are too large, they can turn on all excitable sites, thereby saturating the system and rendering it insensitive to input differences across the cells. For example, suppose that the ith cell v_i receives an input I_i that can turn on some of its B excitable sites by mass action. Let $x_i(t)$ be the number of excited sites and $B - x_i(t)$ be the number of unexcited sites at time t. The simplest mass action law for turning on unexcited sites and letting excited sites spontaneously turn off is

$$\frac{d}{dt}x_i = -Ax_i + (B - x_i)I_i, \qquad \text{(A10)}$$

$i = 1, 2, \ldots, n$. Term $(B - x_i)I_i$ says that the input I_i turns on unexcited sites $B - x_i$ by mass action. Term $-Ax_i$ says that excited sites spontaneously becomes unexcited by mass action at rate A. Hence, when $I_i \equiv 0$, x_i can decay to the equilibrium point 0.

System A10 is inadequate for the following reason: Let the inputs form a spatial pattern $I_i = \theta_i I$. Given a fixed pattern $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$, choose a background intensity I and let the system reach equilibrium. This equilibrium is found by setting $(d/dt)x_i = 0$ and solving for x_i:

$$x_i = \frac{B\theta_i I}{A + \theta_i I}. \qquad \text{(A11)}$$

Now keep θ fixed and increase I. That is, process the same pattern with different background activity. Then all x_i in A11 approach B even if the relative input intensity θ_i is small. This is saturation. How can the system preserve its sensitivity to θ even as I increases? In other words, how does the ith cell v_i compute its "reflectance" θ_i in response to a spatial pattern $I_i = \theta_i I$, $i = 1, 2, \ldots, n$, of inputs? Since

$$\theta_i = I_i I^{-1} = I_i\Big(\sum_{k=1}^{n} I_k\Big)^{-1},$$

cell v_i needs to know what all the inputs $I_1, I_2, \ldots, I_n$ are in order to compute θ_i. Since

$$\theta_i = I_i\Big(I_i + \sum_{k \neq i} I_k\Big)^{-1},$$

increasing the ith input I_i "excites" v_i (increases θ_i), whereas increasing any input I_k, $k \neq i$, "inhibits" v_i (decreases θ_i). When this intuition is most simply modeled by a cellular mass action network, we find the system

$$\frac{d}{dt}x_i = -Ax_i + (B - x_i)I_i - x_i \sum_{k \neq i} I_k, \qquad \text{(A12)}$$

$i = 1, 2, \ldots, n$. In Equation A12, I_i excites v_i via term $(B - x_i)I_i$, just as in A10. The new term

$$-x_i \sum_{k \neq i} I_k$$

describes how the inputs I_k, $k \neq i$, inhibit (note the minus sign) the excited sites of v_i (which number x_i) by mass action. The *gain* of x_i is its decay rate. This is found by grouping together all the terms that multiply x_i. The sum of these terms is $A + I$, where

$$I = \sum_{k=1}^{n} I_k.$$

Thus the inputs automatically change the gain of x_i. In A10 the gain of x_i is $A + I_i$. The two gains differ by the sum

$$\sum_{k \neq i} I_k$$

of inhibitory signals. We now note how automatic gain control by the inhibitory signals overcomes the saturation problem.

Present a spatial pattern $I_i = \theta_i I$ to A12 and let each x_i reach equilibrium. Setting $(d/dt)x_i = 0$, we find

$$x_i = \theta_i \frac{BI}{A + I}. \tag{A13}$$

In A13, x_i remains proportional to θ_i no matter how intense I is, and $BI(A + I)^{-1}$ has the form of a Weber-Fechner law. The saturation problem is hereby overcome using automatic gain control by inhibitory signals.

Noise Suppression

In A12, the passive equilibrium point, due to term $-Ax_i$, and the inhibitory saturation point, due to term

$$-x_i \sum_{k \neq i} I_k,$$

are both zero. This is not always true in vivo, where a cell potential can sometimes be actively inhibited below the passive equilibrium point. How does this fact alter pattern processing? Consider the system

$$\frac{d}{dt}x_i = -Ax_i + (B - x_i)I_i - (x_i + C)\sum_{k \neq i} I_k, \tag{A14}$$

which differs from A12 only in that x_i can fluctuate between B and $-C$, rather than B and 0, where $-C < 0$. Often in vivo B represents the saturation point of a Na^+ channel, $-C$ represents the saturation point of a K^+ channel, and B is much larger than C.

To see how the inhibitory saturation point C influences pattern processing, let A14 equilibrate to the spatial pattern $I_i = \theta_i I$. Setting $(d/dt)x_i = 0$, we find the equilibrium activities

$$x_i = \frac{(B + C)I}{A + I}\left(\theta_i - \frac{C}{B + C}\right). \tag{A15}$$

By A15, $x_i > 0$ only if $\theta_i > C(B + C)^{-1}$. The constant $C(B + C)^{-1}$ is an *adaptation level* that θ_i must exceed in order to excite x_i. For simplicity, suppose that the ratio CB^{-1} matches the ratio of the number of cells excited by each I_i, namely 1, to the number of cells inhibited by I_i, namely $(n - 1)$. If $CB^{-1} = (n - 1)^{-1}$, then $C(B + C)^{-1} = 1/n$. Since, in response to a uniform spatial pattern of inputs, all $\theta_i = 1/n$, no matter how intense I is, it then follows by A15 that all $x_i = 0$. This is noise suppression in its simplest form. It is due to a matched symmetry-breaking between the intracelluar excitatory versus inhibitory parameters (B, C) and the intercellular spread of off-surround versus on-center pathways.

Edge Enhancement, Spatial Frequency Detection, and Pattern Matching

The noise suppression property generalizes to systems whose excitatory and inhibitory interactions can depend on intercellular distances, as in

$$\frac{d}{dt}x_i = -Ax_i + (B - x_i)\sum_{k=1}^{n} I_k C_{ki} - (x_i + D)\sum_{k=1}^{n} I_k E_{ki}, \tag{A16}$$

where C_{ki} (E_{ki}) is the excitatory (inhibitory) coefficient from v_k to v_i. Noise suppression at v_i (i.e., $x_i \leq 0$) occurs in response to a uniform pattern (all $\theta_i = 1/n$) in A16 if

$$B\sum_{k=1}^{n} C_{ki} \leq D\sum_{k=1}^{n} E_{ki}, \tag{A17}$$

which generalizes $CB^{-1} = (n - 1)^{-1}$ in A15. If a rectangular pattern perturbs such a network, then a cell's activity x_i will be suppressed either if its interactions fall so far outside the rectangle or so far inside it that the pattern looks uniform to its interaction coefficients C_{ki} and E_{ki}. Consequently, only activities near the edge of the rectangle will be enhanced. More generally, the spatial gradients of activity in any input pattern are matched against the spatial gradients in each cell's interaction coefficients to enhance the activity of only those cells to whom the input pattern looks nonuniform. In recurrent networks, this property is supplemented by active contrast-enhancing, disinhibitory, and STM processes that can join together cells with similar interaction gradients into a dynamically coherent subfield that is sensitive

to a band of spatial frequencies in the input patterns.

Pattern matching is illustrated as follows. Suppose in A14 that each input I_i is a sum of two inputs J_i and K_i whose patterns $J = (J_1, J_2, \ldots, J_n)$ and $K = (K_1, K_2, \ldots, K_n)$ are to be matched. If J and K mismatch each other's peaks and troughs to form an almost uniform total pattern $I = (I_1, I_2, \ldots, I_n)$, then by A15 all x_i will be inhibited if $CB^{-1} \geq (n-1)^{-1}$. By contrast, if the two patterns reinforce each other, say $J_i = \alpha K_i$, then by (A15),

$$x_i = \frac{(B + C)(1 + \alpha)\bar{K}}{A + (1 + \alpha)\bar{K}}\left[\theta_i - \frac{C}{B + C}\right]$$

where

$$\bar{K} = \sum_{i=1}^{n} K$$

and $\theta_i = K_i(\bar{K})^{-1}$. In other words, matching J and K amplifies each x_i without changing the pattern θ_i.

Appendix D

This section summarizes some properties of recurrent on-center off-surround networks, including normalization, contrast enhancement, quenching threshold, and STM properties.

To see how recurrent networks normalize their STM activity, we first note by Appendix C that these networks need competitive interactions to solve the noise-saturation dilemma. The simplest recurrent on-center off-surround network is defined by

$$\frac{d}{dt}x_i = -Ax_i + (B - x_i)[f(x_i) + I_i] - x_i\Big[\sum_{k \neq i} f(x_k) + J_i\Big], \quad \text{(A18)}$$

$i = 1, 2, \ldots, n$. As usual, x_i is the STM activity of v_i, term $(B - x_i)f(x_i)$ describes the self-excitation of v_i via a positive feedback signal $f(x_i)$—the recurrent on-center—and term

$$-x_i \sum_{k \neq i} f(x_k)$$

describes the inhibition of v_i via negative feedback signals $f(x_k)$, $k \neq i$—the recurrent off-surround. Term I_i is the ith excitatory input, and term J_i is the ith inhibitory input, for example,

$$J_i = \sum_{k \neq i} I_k$$

in A12.

Contrast Enhancement, Normalization, and Quenching Threshold

An important problem in system A18 is to choose the feedback signal function $f(w)$ as a function of activity level w in such a way as to suppress noise but contrast enhance and store in STM behaviorally important patterns. This problem was solved in Grossberg (1973). The solution is reviewed in Grossberg (1978e, Sections 14 and 15).

To understand the simplest STM properties, A18 is transformed into pattern variables $X_i = x_i x^{-1}$ and total activity variables

$$x = \sum_{k=1}^{n} x_k$$

using the notation $g(w) = w^{-1}f(w)$ and supposing that all $I_i = J_i = 0$. Then

$$\frac{d}{dt}X_i = BX_i \sum_{k=1}^{n} X_k[g(X_i x) - g(X_k x)] \quad \text{(A19)}$$

and

$$\frac{d}{dt}x = -Ax + (B - x)\sum_{k=1}^{n} f(X_k x). \quad \text{(A20)}$$

For example, if $f(w)$ is linear, namely, $f(w) = Cw$, then $g(w) = C$ and all $(d/dt)X_i = 0$ in A19. In other words, A19 can perfectly remember *any* initial pattern of reflectances. However, by A20 if $A \geq B$, then $x(t)$ approaches zero as $t \to \infty$, whereas if $B > A$, then $x(t)$ approaches $B - A$ as $t \to \infty$, whether or not a prior input pattern occurs. Thus if STM storage is ever possible, then $B > A$, and consequently noise will be amplified as vigorously as inputs. A linear signal amplifies noise, and is therefore inadequate despite its perfect memory of reflectances.

A slower-than-linear signal $f(w)$, for example, $f(w) = Cw(D + w)^{-1}$ or more generally, any $f(w)$ such that $g(w)$ is monotone decreasing, is even worse. By A19, if $X_i > X_k$, $k \neq i$, then $(d/dt)X_i < 0$ and if $X_i < X_k$, $k \neq i$, then $(d/dt)X_i > 0$. All differences in reflectances are hereby erased by the reverberation, and noise amplification also occurs. The whole network experiences a type of seizure.

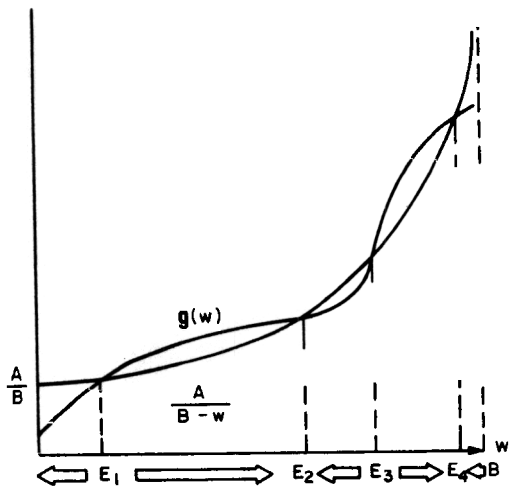

Figure A2. The even solutions E_0, E_2, ... of $g(w) = A(B - w)^{-1}$ are stable equilibrium points of $x(\infty) = \lim_{t\to\infty} x(t)$. (Since $g(w) = w^{-1}f(w)$, these points are solutions of $f(w) = Aw(B - w)^{-1}$. If $x(0) < E_1$, then $x(\infty) = 0$; thus E_1 defines the level below which $x(t)$ is treated as noise and quenched. All equilibrium points satisfy $E_i \leq B$; hence, short-term memory is normalized.)

If $f(w)$ is faster than linear, then the situation is better; for example, $f(w) = Cw^n$, $n > 1$, or more generally any $f(w)$ such that $g(w)$ is monotone increasing. In this case, if $X_i > X_k$, $k \neq i$, then $(d/dt)X_i > 0$, and if $X_i < X_k$, $k \neq i$, then $(d/dt)X_i < 0$. Consequently, this network chooses the population with the initial maximum in activity and totally suppresses activity in all other populations. This network behaves like a finite state, or binary choice machine. The same is true for total activity, since as $t \to \infty$, A20 becomes approximately

$$(d/dt)x \cong x[-A + (B - x)g(x)]. \quad \text{(A21)}$$

Thus the equilibrium points of $x(t)$ as $t \to \infty$ are $E_0 = 0$ and all the solutions of the equation

$$g(x) = A(B - x)^{-1}. \quad \text{(A22)}$$

If $g(0) < A/B$, then the smallest solution E_1 of A22 is unstable (Figure A2) so that small activities $x(t)$ are suppressed as $t \to \infty$. This is noise suppression due to recurrent competition. Every other solution E_2, E_4, ... of A22 is a stable equilibrium point of $x(t)$ as $t \to \infty$ (total activity quantization) and all equilibria are smaller than B (normalization).

The faster-than-linear signal contrast enhances the pattern so violently that the good property of noise suppression is joined to the extreme property of binary choice. This latter property is weakened by contructing a hybrid signal function that is chosen faster than linear at small activities to achieve noise suppression, but which levels off at high activities if only because all signal functions must be bounded. In the simplest case, $f(w)$ is a sigmoid, or **S**-shaped signal function. Then there exists a quenching threshold (QT). If v_is initial activity $x_i(0)$ falls below the QT, then its STM activity is quenched, or laterally masked: $x_i(\infty) = 0$. All the $x_i(0)$s that exceed the QT are contrast enhanced and stored in STM. Simultaneously, the total STM activity is normalized. Speaking intuitively, the QT exists because the faster-than-linear range starts to contrast enhance the pattern. Simultaneously, normalization shifts the activities into the intermediate linear range that stores any pattern, in particular the partially contrast-enchanced pattern. Because a QT exists, the network is a tunable filter. For example, a nonspecific arousal signal that multiplicatively inhibits all the recurrent inhibitory interneurons will lower the QT and facilitate storage of inputs in STM. Grossberg and Levine (1975) mathematically studied how such attentional shunts alter the resultant STM pattern by differentially sensitizing prescribed subfields of feature detectors that are joined together by competitive feedback interactions. The privileged subfields mask the activities in less sensitive subfields.

Such examples, either taken separately or linked together by feedback, provide insight into how interactions between continuously fluctuating quantities can sometimes generate discrete collective properties of the system as a whole. More generally, Grossberg (1978c) proves that every competitive system induces a decision scheme that can be used to globally characterize its pattern transformations as time goes on.

Appendix E

This section summarizes how the simplest transduction law realizable by a depletable chemical generates properties of antagonistic rebound due to specific cue offset and to nonspecific arousal onset when two parallel transduction pathways compete.

Transmitters as Gates

The transmitter law that we need can be derived in two ways. Originally, it was derived as the minimal law that was compatible with psychological postulates of classical conditioning (Grossberg, 1969c, Section 20; Grossberg, 1972c, Section 2). I now show that the law is the simplest transduction rule that can be computed using a depletable chemical transducer.

The simplest transduction rule converts an input I into a proportional signal S, namely,

$$S = BI, \tag{A23}$$

where $B > 0$ is some proportionality constant. Equation A23 says that I is *gated* by B to yield S. If we interpret B as the amount of transducer and BI as the rate with which transducer is released to create signal S, then A23 says that the input I activates the transducer B in a statistically independent, or mass action, way.

When the transducer is released to activate another cell, there must exist a mechanism whereby it can be replenished, so that A23 can be maintained, at least approximately, through time.

Let $z(t)$ be the amount of transducer at time t. How can we keep $z(t) \cong B$ for all $t \geq 0$ so that the transduction rule

$$S = Iz(t) \tag{A24}$$

approximately agrees with A23? This question leads to the following law for the temporal evolution of the amount $z(t)$ of available transducer

$$\frac{dz}{dt} = A(B - z) - Iz. \tag{A25}$$

The term $A(B - z)$ in A25 says that $z(t)$ accumulates until it attains level B. The term does this by accumulating transducer at rate AB, that is proportional to B, and by feedback inhibition of the production rate at a rate $-Az(t)$ that is proportional to $z(t)$. The term $-Iz(t)$ in A25 indicates that transducer is depleted at a rate proportional to its rate of elimination, which is due to gating of I by $z(t)$. When $z(t) \cong B$, term $-Iz$ is proportional to $-BI$, as required by A23. Thus A25 is the law that "corresponds" to the law $S = BI$ when depletion of transducer can occur. It describes four effects working together: production, feedback inhibition, gating, and depletion.

Rebound Due to Cue Offset

Suppose that the adaptation level is I and that the cue input is J. Consider the simplest case in which the total signal in the on-channel is $S_1 = I + J$ and in the off-channel is $S_2 = I$. Let the transmitter z_1 in the on-channel satisfy the equation

$$\frac{d}{dt}z_1 = A(B - z_1) - S_1z_1, \tag{A26}$$

and the transmitter z_2 in the off-channel satisfy the equation

$$\frac{d}{dt}z_2 = A(B - z_2) - S_2z_2. \tag{A27}$$

After z_1 and z_2 equilibrate to S_1 and S_2, $(d/dt)z_1 = (d/dt)z_2 = 0$. Thus by A26 and A27,

$$z_1 = \frac{AB}{A + S_1} \tag{A28}$$

and

$$z_2 = \frac{AB}{A + S_2}. \tag{A29}$$

Since $S_1 > S_2$, it follows that $z_1 < z_2$; that is, z_1 is depleted more than z_2. However, the gated signal in the on-channel is S_1z_1, and the gated signal in the off-channel is S_2z_2. Since

$$S_1z_1 = \frac{ABS_1}{A + S_1} \tag{A30}$$

and

$$S_2z_2 = \frac{ABS_2}{A + S_2}, \tag{A31}$$

it follows from $S_1 > S_2$ that $S_1z_1 > S_2z_2$ despite the fact that $z_1 < z_2$. Thus the on-channel gets a bigger signal than the off-channel. After the two channels compete, the cue input J produces a sustained on-response whose size is proportional to

$$S_1z_1 - S_2z_2 = \frac{A^2BJ}{(A + I + J)(A + I)}. \tag{A32}$$

Now shut J off. Then the cell potentials rapidly adjust until new signal values $S_1^* = I$ and $S_2^* = I$ obtain. However, the transmitters z_1 and z_2 change much more slowly, so that A28 and A29 are approximately valid in a time interval that follows J offset. Thus the net signals are approximately

$$S_1^*z_1 = \frac{ABI}{A + S_1} \tag{A33}$$

and

$$S_2^*z_2 \cong \frac{ABI}{A + S_2}. \tag{A34}$$

Since $S_1 > S_2$, $S_1^* z_1 < S_2^* z_2$. The off-channel now gets the bigger signal, so an antagonistic rebound occurs whose size is approximately

$$S_2^* z_2 - S_1^* z_1 = \frac{ABIJ}{(A + I + J)(A + I)}. \quad \text{(A35)}$$

The rebound is transient because the equal signals $S_1^* = S_2^* = I$ gradually equalize the z_1 and z_2 levels until they both approach $AB(A + S_1^*)^{-1}$. Then $S_1^* z_1 - S_2^* z_2$ approaches zero, so the competition between channels shuts off both of their outputs.

Rebound due to Arousal Onset

Suppose that the on-channel and off-channel have equilibrated to the input levels I and J. Now increase I to I^*, thereby changing the signals to $S_1^* = I^* + J$ and $S_2^* = I^*$. The transmitters z_1 and z_2 continue to obey A28 and A29 for awhile, with $S_1 = I + J$ and $S_2 = J$. A rebound occurs if $S_2^* z_2 > S_1^* z_1$. This inequality is true if

$$I^* > I + A, \quad \text{(A36)}$$

since

$$S_2^* z_2 - S_1^* z_1 = \frac{ABJ(I^* - I - A)}{(A + I)(A + I + J)}. \quad \text{(A37)}$$

In particular, a rebound will occur if I^* exceeds $I + A$ no matter how J is chosen. In other words, if the mismatch is great enough to increment the adaptation level by more than amount A, then all dipoles will simultaneously rebound, and by an amount that increases as a function of J, as in Equation A37. This is not true in all versions of the dipole model, since the signals S_i, $i = 1, 2$, are not always linear functions of their inputs. There exist examples in which the most active dipoles can be rebounded even though less intensely activated dipoles are amplified without being rebounded. Moreover, if the signals are sigmoid functions of input size, then inverted-U effects occur in both the on- and off-responses to cue and arousal increments (Grossberg, 1972b, 1972c, 1975).

Received February 27, 1979
Revision received August 15, 1979 ■