Reprinted from ELECTRICAL STIMULATION OF THE BRAIN, 1961
Printed in United States of America

T-10

CHAPTER 39

Implications for Systematic Studies of Behavior

KARL H. PRIBRAM

THIS VOLUME on the effects of brain stimulation is a landmark. Looking back, we perceive a century of observation and experiment. Looking forward, we perceive that present endeavors are but first steps full of problems and promise. The field was initiated when Fritsch and Hitzig demonstrated that movements and sensations could be obtained from human beings when their brain cortex was electrically excited. This led to decades of experiments during which maps of brain topography and function were made. The fact was established beyond doubt that portions of the brain exercise control over sensation and movement. The anatomical details of the connections whereby this control could take place were teased out in painstaking fashion. Arguments raged furiously, for half a century, on the specific mechanisms represented by these anatomical paths, and the ghosts of the arguments have not even now been completely stilled. But, interestingly, nowhere in this volume have these old battles been re-engaged. Rather, new horizons have been charted. These fall into three closely related categories, which can best be stated as questions: (1) How does the brain function in the learning process? (2) How does the brain control the mechanisms of drive? (3) How does man's brain make reinforcement possible—exclusively through mechanisms of drive reduction and induction or through some more inclusive cognitive and perceptual process?

The freshness of the material is attested by the wealth of technical detail in each of the chapters—detail that is presented by an investigator who is charting new territories and wants very much to communicate with his fellow investigators. He knows that unless he presents every last nuance of his experimental procedures communication will fail and his results cannot be duplicated. Only when a particular area of investigation has reached maturity are the details so much a part of the general scientific endeavor that they need not be voiced.

Another indication of freshness is the spottiness of the systematic detail which is recorded. The effect of brain stimulation on behavior must deal with four separate classes of variables: (1) anatomical locus, (2) parameters of stimulation, (3) environmental situation, and (4) opportunity for behavioral response. It is almost impossible for any investigator, especially in the initial stages of his experiment, to cover all these classes systematically. Yet such systematic exploration may be crucial to the interpretations he must make of his data. Some of the chapters cover one of these classes thoroughly but ignore the other three. Other chapters explore two or even three of the classes in a very systematic fashion, but only rarely is an investigator prepared to treat all four. Therefore some discrepancy in results, more apparent than real, will plague the reader. Some of the discrepancies are of sufficient import to warrant programmatic attack. Others will gradually iron themselves out as more data become available. Part of the excitement of the presentations stems from the discrepancies, for when all the problems in a field of investigation appear to be solved it is time to turn elsewhere. And that certainly is not the case for these reports of the effects of brain stimulation on behavior.

LEARNING

Fittingly the exposition of the neural mechanisms

involved in conditioning begins with a chapter by Loucks, who initiated his studies a quarter of a century ago. In his current contribution Loucks continues an analysis of the manner in which the excitatory processes that accompany conditional and unconditional stimuli are propagated. His experiments establish that propagation proceeds in an afferent-efferent vertical path from subcortical regions to cortex and that it is not mediated transcortically. The results he describes are in consonance with those of Sperry (1947, 1958; Sperry *et al.*, 1955), whose incisional interruptions of transcortical paths failed to produce any effect on discriminative and motor behavior. They are also supported by the experiments of Harrison, who was able to condition an animal to localized cortical stimulation. Harrison found that the locus of the effective signal was restricted to one or two millimeters and that more remote stimulation failed to produce the conditional response (Crosser and Harrison, 1960). The final sentence of the chapter by Loucks is worth repeating here, since it states succinctly the problems to which the investigators of the effects of cortical stimulation on the learning process are addressing themselves: "Before any very specific inferences can be made about the nature of the engram, it may be necessary to find two centers in relatively contiguous regions—one mediating the conditional signal excitation and the other mediating the reinforcing process—which will make possible the successive isolation of crucial units by the implanting of structural barriers."

Gengerelli makes a somewhat different approach to the same problem. He too is interested in the neural mechanisms that mediate the learning process—i.e., engram formation. To this end Gengerelli builds a mathematical model, which he then puts to test in the laboratory. His emphasis is on the coequality of an inhibitory process to counteract the reinforcement of an excitatory process. The precision of the mathematical model is unfortunately not as yet matched by equal precision and systematic inquiry in the behavioral, situational, and anatomical variables that determine the results of his experiments. While parameters of electrical stimulation are explored, only a beginning is made in the examination of situational variables that influence behavior. Anatomical data are practically absent. So, to date, the test of the model is a preliminary one, as is recognized by Gengerelli himself.

The experiments reported by Flynn, MacLean, and Kim must be noted by those who are trying to isolate engrammatic processes at the isocortical level. The units which Loucks and Gengerelli are attempting to study may be units only by virtue of some limbic system mechanism which is necessary to their formation. Flynn and his collaborators present some interesting observations on the role of the hippocampus in engram formation. They show that hippocampal seizures interfere with the learning and performance of conditional responses, and they suggest that, unless traces are allowed to accumulate in the hippocampus, conditioning cannot take place. Their results support those obtained with the ablation techniques: Lesions of the medial and basal limbic structures, but not of isocortex, affect extinction of an acquired conditioned avoidance response (Pribram and Weiskrantz, 1957; Hunt and Diamond, 1957). That the limbic systems are essential to the performance of even some of the simplest conditional responses seems by now to be well established (see Brazier, 1958*b*, 1959*a*).

Doty's experimental approach takes this multiple determination of the engram seriously. He maps the variety of neural mechanisms that contribute to the formation of a single conditional response. To date this mapping procedure has not progressed far enough to allow many generalizations to be made about the nature of the mechanisms involved. But Doty argues effectively that the best explanation for the multitude of neural systems involved in the establishment of any conditional response is that the neural alteration which accounts for the response must be located some place in the effector system. He thus supports Sperry (1952), who has argued that a change in the level of excitability of a particular system would render it prone to be triggered by slight additional sensory input. But a question remains as to the path and the locus of convergence of these various neural mechanisms. Lashley's experiments and those of Sperry make it improbable that convergence takes place transcortically. Subcortical paths and loci are suggested by the experiments of Loucks and of Doty. Then what remains of the functions usually imputed to the motor cortex? It might be, as Lashley suggested and as his and other experimental results have indicated (e.g., those of Roberts, Chapter 37; Pribram *et al.*, 1955-56), that the motor cortex is just one other of the several facilitatory mechanisms that preset the final common path so that it can be more readily triggered either by sensory input or by intrinsic neural activity. Thus the search for the engram ranges wide.

Burns and Mogenson, in a chapter that is thorough and rich in ideas, take this search one step further. They suggest that different neural mechanisms may be involved in acquisition and extinction. Their experiments show that isocortical stimulation during extinction results in selective interference with the animal's performance and not with the rate of extinction. These data can be apposed to those obtained when stimulations and lesions of the limbic systems are made and the rate of extinction is markedly affected. When the "habit" is in the process of being acquired, the isocortical stimulation used in these experiments disrupts not only the response sequences but also the sequences involving the attainment of reward. After the habit is acquired, however, the deficits in performance are less in magnitude. Finally, during extinction the functions involved in the eventual suppression of previously reinforced responses are not interfered with by isocortical stimulation. These data are supported by another series of experiments that have used aluminum hydroxide implantations in various cortical areas and have shown that interference with acquisition of specific problem-solving behavior is selective. Once performance becomes habitual, the implantations have no effect (Pribram, 1951; Obrist *et al.*, 1960; Stamm and Pribram, 1960).

There is available to experimental psychologists a quantitative theoretical approach to the problem of the relations between acquisition and extinction of competing responses. This body of "statistical" learning theory (Bush and Mosteller, 1955; Estes, 1959; Suppes *et al.*, 1960) would usefully provide the background for further studies performed along the lines initiated by Burns and Mogenson. Somewhat more complex tasks are used by those interested in statistical learning theory than were used in the studies reported here. These more complex tasks could be put to good advantage. Quantitative predictions can be made and checked concerning the effects of isocortical (and limbic) stimulation on the factors that determine acquisition, performance, and extinction. Such quantitative studies would not only extend our knowledge of the neural mechanisms involved in the learning process but also add precision to the definition of "stimulus" and "reinforcement"—two concepts now plagued with vagueness in these otherwise precise instruments.

DRIVE

The chapters on drive mechanisms, for the most part, center around "cerebral self-stimulation." An exception is the report by Lindsley, who presents an interesting development of his activation theory. He speaks of behavioral arousal, specific attention, and a feedback control mechanism—all as a function of activation from the brain-stem reticular formation. At the cortex, graded response mechanisms—the cyclic changes in excitability that have been given so much neurophysiological attention recently—are directly affected. On the basis of his experiments, Lindsley suggests that these cyclic changes, especially those recorded as occipital alpha rhythm, function as a neuronic shutter device that controls and regulates sensory input. Despite the limitations that plague reaction-time studies and the controversies that rage with regard to the origin of the alpha rhythm, Lindsley's data and hypotheses cannot be ignored. He himself reviews critically the evidence for and against the various aspects of his notions. Many arguments are sure to follow publication of this presentation, which in itself is sufficient recommendation.

Olds continues his brilliant explorations of what he calls localized reward systems in large regions in the brain. He describes experiments in which the interaction between self-stimulation and food deprivation and between self-stimulation and hormonal effects are measured; also experiments on the effects of self-stimulation. He continues to view his results in hedonistic terms—that is, he states categorically that central excitation can lead to pain and to pleasure and that pleasure is rewarding and pain punishing. His chapter tells a simple, direct story. If his results were consistently confirmed by others, our work would be complete. It may turn out that the story is not as simple as it seems.

Brady presents data somewhat similar to those of Olds, but a greater range of species was used for his studies. Though Brady's work is essentially empirical in nature and relatively unconcerned with theory, his data can be fitted into other, more systematic, reports. This empirical mapping Brady does well. His experiments have the virtue that some new situations (various schedules of reinforcement) are used and that at least several anatomical structures are explored in species other than the rat. And this virtue leads to the recognition of discrepancies. But more of this later.

...Miller reviews his experiments with his usual theoretical sophistication. His experiments use a variety of situations and explore a variety of responses. Anatomically, a fair portion of the brain stem has been covered, and the results amplify

considerably those reported by Olds and by Brady. Situation does make a difference. Excitation of the very same point in the brain may produce effects that can be interpreted as rewarding or punishing, depending on the test situation—and even in the same situation "paradoxical" results may be obtained. But the paradox may be in the interpretation, not in the data. For these data can be added to those where excitation results directly in eating, drinking, and sexual behavior (such as those covered in the excellent chapter by Smith), and to data on learning (such as those reported by Knott and Ingram), to suggest a variety of consistent notions (even theories) about drive mechanisms. Miller presents one scheme for doing this; other schemes have been suggested, for example by Sheer.

Sheer presents a telling and detailed argument against hedonism—an argument based in part on the experimental results obtained in his laboratory. He concentrates on those facts that point to behavioral "emotional" facilitation (as this results from interoceptive stimulation), which, when associated with ongoing response sequences, is reinforcing. His formulations describe one feasible mechanism of reinforcement, a mechanism which, in some important respects, is similar to that proposed by another and very famous neuropsychologist who had also described the neurological mechanisms that could be involved (Freud, 1954). But the problem of reinforcement is not so simply solved. Sheer's formulations include a considerable amount, though not all of the data that deal with drive mechanisms. For instance, Deutsch (1960) has shown that experimentally separable mechanisms are responsible for the initiation and the cessation of drinking. And other views, each based on an extensive amount of evidence, have emphasized either the selective or the homeostatic or (as has Sheer) the facilitatory factors. The experimental background for each of these positions has a long history. A brief survey and attempt at synthesis of the disparate positions might prove useful.

More than a century ago Claude Bernard (1858) initiated a branch of neurophysiology concerned with the regulation of the organism's metabolism and endocrine functions by the central nervous system. His now famous *piqûre* experiments, in which diabetes was produced by making small stab wounds in the brain stem, led him to the conception of a *milieu intérieur* that is still central in the thinking of modern neurophysiologists (e.g., the reviews by Colle *et al.*, 1952). Somewhat less well known are the extensive series of experiments by Karplus and Kreidl (1909, 1910, 1912, 1928), which thoroughly explored the relations between diencephalic centers and the regulation of visceral activities. This branch of neurophysiology has been advanced by Cannon's (1929) formulation of the concept of homeostasis and by the laboratory analyses that occupied him and his collaborators. Another group of experimentalists, directed by Ranson (1937), explored the relations between hypothalamic mechanisms and the maintenance of body temperature and food intake and activity. Contemporary investigations of the thirst mechanisms (Andersson, 1952; Andersson and McCann, 1955, 1956*a*, *b*), of endocrine control (Harris *et al.*, 1958; Hume, 1958), and of the respiratory center (Meyer, 1957) are some of the highlights in this area of investigation.

In spite of a variety of contrary opinions, these investigations have tended to confirm the idea that specific centers exist in the central nervous system to control one or another of the metabolic and endocrine activities of the organism. Most of the evidence against centers has come from studies that deal with parts of the central nervous system other than these centrally located regions in the brain-stem core and with behavior other than that involved in the regulation of the organism's metabolism. What, then, characterizes these structures? What makes them different from other central neuronal aggregates? The most useful way to look at this difference seems to revolve around the specific sensitivities of the centers to one or another physiochemical substance. And this is exactly how receptors are defined in the peripheral nervous system.

The conception that receptor mechanisms may be located around the midline ventricles of the brain stem derives support from two sources. The experiments already alluded to have produced data that are consistent with the concept of homeostasis. Any homeostat must include an element that is especially sensitive to the range of physical (or chemical) events that the homeostat attempts to regulate. In the case of the brain-stem homeostats, this sensitive element could be entirely located in the peripheral mechanism that is afferently connected with the central nervous system, but experiment has demonstrated that at least some of the sensitivity is located centrally. For example, hypertonic saline injected into the third ventricle immediately causes goats to drink voluminously (Andersson, 1952); heat applied to the base of the

anterior extremity of the third ventricle immediately causes changes in the heat regulating mechanisms all over the body (Ranson *et al.*, 1937); and very local changes in the partial pressure of CO_2 in the posterior brain stem dramatically alter the rate and depth of respiration (Meyer, 1957). All these sensitivities are specific and restricted to very small regions, and all are localized in structures fairly near the third and fourth ventricles of the brain stem.

The second source of support for the conception that receptor mechanisms might be located near the midline ventricular system is less direct. Ontogenetically, this median part of the central nervous system is derived from the most dorsal part of the neural crest. Invagination to form the neural tube makes periventricular components the components most akin in origin to the epidermal portions of the ectodermal formations. And it is these portions of the ectoderm that induce some of the more specialized of the receptors, such as the retina. Furthermore, the sensitivities of the periventricular mechanisms are very similar to those of the skin. Temperature change, deformation, and changes in hydration are some of the major categories of stimuli to which both are sensitive.

In summary, then, the work of a century of neurophysiological experiment seems to be leading to the conception that a series of specialized receptors are located near the midline ventricular systems of the brain stem. These specialized receptors are the classical centers for the control of respiration, food intake, etc., that have interested physiologists and biochemists concerned with the neural regulation of the organism's metabolism and endocrine functions. These receptors are conceived to function as sensitive elements of a variety of homeostats concerned with the regulation of appetitive-consummatory processes.

Immediately beyond the limits of the periventricular receptor centers lies a matrix of neural reticulum spotted here and there with neuronal aggregates and coursed only occasionally by long nerve fibers. The anatomy of the brain-stem reticular formation has been detailed by Brodal (1957), the Scheibels (1958), and Russell (see Chapter 15). Its physiology is well documented by Jasper *et al.* (1958), by Magoun (1958*b*), and also in this volume.

Characteristically, the reticular systems are composed of fairly short, fine-fibered neurons with vast dendritic networks. Inputs converge on each nerve cell from many branches of the long classical projection tracts that originate in the various receptor fields of the organism. Each neural element in the system is influenced by a variety of sensory modes as indicated by changes in the electrical activity recorded with microelectrodes. In addition, a reciprocal relation with the rest of the neuraxis exists. For example, the cerebral cortex is activated when the brain-stem reticular formation is electrically excited; and, conversely, cortical stimulation affects the activity of the reticular systems. This convergence of input and diffuseness of interrelations suggests that the most likely action of these systems is to influence the general state of excitability of the nervous system. This suggestion is supported by the finding that cortical rhythms are activated and deactivated by electrical stimulation of the reticular formation and by the fact that lesions and stimulations of these systems have been shown to be related to such behavioral processes as the sleep-wakefulness cycle, alertness, and attention. Furthermore, the anatomical structure of these systems is of a kind that suggests graded response mechanisms rather than signal transmission. Synapses and dendrites are abundant; fibers are, for the most part, short and fine, so that the conduction velocity of an impulse is slow and its amplitude small. Such graded response mechanisms are especially sensitive to changes in their chemical environment. A great number of studies have related the action of neural transmitters and psychopharmacological agents to the functions of these systems (Bradley and Elkes, 1957). The proximity of the reticular systems to the more specialized periventricular receptors is therefore germane to the problem of the "homeostatic" regulation of the organism's *milieu intérieur*, ordinarily referred to by psychologists as "mechanisms of drive."

A considerable literature has developed with regard to the role of the reticular systems and their forebrain extensions in the regulation of drives. Lindsley (1951) and Hebb (1955*a*) have spelled out the details of "activation" theories based on neurophysiological evidence. Some constructive criticism of these theories has come from investigators who have explored the internal core systems and their forebrain extensions. These investigators have been impressed with the selective action of various locations on specific drive mechanisms. The formulations of central excitatory mechanisms or central motive states, as proposed by Beach (1958) and by Morgan (1959), lean in this direction. The

studies of Teitelbaum (1955) and Stellar (1954) on the hypothalamic control of feeding, and those of Miller *et al.* (1950, 1955, 1956) on the regulation of thirst and hunger, make the same point.

Most explicit in opposition to an activation theory, yet somewhat different from the selection notion, have been the statements of Olds, who has interpreted the data from self-stimulation experiments to mean that a central hedonistic mechanism exists. Since animals will work to produce electrical excitation in parts of their brains and will work to stop such stimulation in other parts, a neural "pleasure system" and a neural "displeasure system" are postulated. The "pleasure system" is subdivided into portions that deal differently with sex and hunger. Furthermore, Olds has proposed that the pleasure system works as a positive feedback mechanism, so that an organism's pleasurable activity is stopped only when restraints external to the organism are imposed.

Each of these formulations is based on a particular set of data; each ignores, for the most part, the data on which the alternate proposals are based. All these theories share a view that is essentially nonhomeostatic. Activation, selection, and hedonism all emphasize direction rather than equilibration. Attempts have been made to reconcile direction with equilibration. Miller's (1959) discussions of the cue properties and the drive properties (conceived in need-reduction terms) of stimuli are probably the best known of these attempts.

Taken together, the neurophysiological and behavioral evidence seems to add up to the view that both an equilibratory and a directional component characterize drives, and that selection, activation, and equilibration are all important. However, hedonism need not be invoked, nor need one consider the selective or cue properties of stimuli as the sole directional components of drive. A simpler view can be formulated. Consider the various elements that make up a homeostat. The sensitive element has already been mentioned; a homeostat must contain a receptor that is selectively sensitive to the physical or chemical process that is to be regulated. Such receptors exist in the central nervous system. A homeostat is essentially a feedback unit. As such, it must be so constituted that errors in adjustment are fed back to the sensitive element in time for this element to signal the disparity to the operate mechanism. There is ample neurophysiological and behavioral evidence that negative feedback mechanisms exist. After all, this is the evidence on which the concept of homeostasis rests. But, in addition to these elements of the homeostat, there is another which has been largely ignored. Negative feedback mechanisms are usually equipped with a device by which their bias can be set (e.g., the thermostats in our homes can be set to one or another temperature). The activating mechanisms and electrical self-stimulation can be thought of as changing (by alterations in graded response mechanisms) the biases of the various homeostats with which they are anatomically juxtaposed. The laws that govern the changes in biases would be different from those that govern regulations once the biases have been set. For instance, small increments of change in bias are apt to be accommodated smoothly and to be directional; more abrupt changes are apt to cause marked fluctuations until the regulatory mechanism can re-establish equilibrium. No simple hedonistic rule can be applied; that is, behavior is not always guided toward some pleasurable consequence. Selection of stimuli depends on the state of the receptor. Activation shifts biases. Equilibratory homeostasis in the classical sense (and thus need-reduction) is seen as only one phase, the equilibrational, of a rather more complex process. Basic to this process is the up-to-date neural homeostat constituted of a receptor, negative feedback, and bias. Thus the essential mechanism of drive is conceived to be a biased (i.e., a tuneable) homeostat.

A further refinement can be made of this conception on the basis of the Deutsch (1960) studies. An ordinary homeostat is started and stopped by the *same* stimulus. For example, air at a lower temperature starts, and at a higher temperature stops, the mechanism controlled by a thermostat. Deutsch has shown that a different mechanism is responsible for starting and for stopping eating and drinking behavior. The start mechanism is most likely to work through the sensitive element, as already described. But the stop mechanism is mediated through a variety of inputs by way of afferents whose origin is distributed along a good portion of the upper part of the alimentary system. These afferents could alter bias through the action of the diffusely organized portions of the brain stem, especially those immediately surrounding the core receptors. Or there could be an anatomically distinct central station for this mechanism. Experiments to determine which is the correct alternative are now feasible.

In short, a thoroughgoing experimental analysis

of the neurology of drive is now under way. In addition, as the reported formulations of Sheer and of Lindsley emphasize, there is considerably more to be studied about an organism's motives and attitudes than the neural mechanisms involved in "joy" and "pain." Has Olds the neurophysiologist (Olds, 1959) forgotten what Olds the neuropsychologist (Olds, 1956*d*) contributed in this area?

REINFORCEMENT

Some new insights into these problems and the discrepant data might come from an entirely different source. Verbal reports have been made by persons stimulated through electrodes implanted into the structures involved in these controversial drive mechanisms. The effects on behavior of electrical excitation of the brain began with the study of the *human* brain—and this primary interest continues through the centuries. Roberts beautifully summarizes the reports from the foremost laboratory in this field: Penfield's neurosurgical amphitheater.

King warns of some of the limitations and restrictions that accompany clinical studies but are not present when subhuman mammals are used. His efforts have been addressed in large part to the development of refined psychomotor tests that are sensitive to the slight alterations that cerebral stimulation may produce. He has not, however, neglected verbal behavior, and his description of a film strip of the responses of a patient to stimulation of the amygdala is most interesting. Qualitative changes in mood and affect can be presented in this fashion, and perhaps quantitative techniques can be derived from these initial studies.

Spiegel and Wycis continue their explorations of subcortical mechanisms. The data they present are always of interest to neurologically oriented psychologists. Walker and Marshall also present data of interest primarily to neurologists. Their contribution to the study of the mechanisms of propagation of potential changes in the nervous system—and to the problem of epilepsy in particular—is important and novel in its emphasis on subcortical as well as on cortical genesis of the seizure patterns. Their report that even the gross "epileptic" discharges of neural potentials "are not derived from the summation of potentials of large masses of tissue; instead, they are derived from highly localized fields a few cubic millimeters in size within a volume conductor" must be taken seriously by those interested in engram formation. The results, taken together with those obtained by Loucks and others as reported in this volume, indicate that "irradiation" or "association" or the "formation of cell assemblies" is strongly inhibited once a certain limited size of neuronal aggregates has been involved. Please note, simulators of neuronal networks!

Feindel returns us to the question of the functions of some of the neural structures that Olds considers to be mechanisms of reward. He does not even mention reward, reinforcement, or drive. He is concerned with arrest of awareness and arrest of memory-recording as a result of stimulation of the periamygdaloid region. Two courses are open to us: We can say, "Oh well, people are different from animals. Let's forget the whole business." Or we can search deeply into our understanding (or lack thereof) of what is meant by reward and by punishment, by drive and by attention, by memory processing, and by learning. The results of the neuropsychological experiments presented in this volume make the latter course the more attractive one. The animal studies in and of themselves demand reinterpretation of our notions about fundamental psychological concepts, so why not include data obtained from such observations on man? Verbal reports are behavioral data; they may confuse if treated too literally and without reference to other types of behavioral response. But used judiciously, and with proper control, verbal reports may clarify. Perhaps the reward-memory controversy that rages about the functions of the limbic structures can be resolved by such an analysis. The following examination of some experimental evidence that bears on the relations between limbic structures and the neurobehavioral processes concerned in reinforcement shows the direction such analysis can take.

The history of the problem begins with neurobehavioral studies performed on animals. Ablations and stimulation of any of the various structures that make up the limbic systems interfere with a variety of behaviors. These data have been detailed in Jasper *et al.* (1958), Klüver (1952), Wolstenholme and O'Connor (1958), and in this volume. In order to remain uncommitted with respect to one or another theoretical position and yet have a pedagogically useful categorization, some neutral label that describes this behavioral complex of feeding, fleeing, fighting, and sex had to be invented (by now students everywhere refer to the complex as "the four F's"). Feeding includes such aspects as hoarding; sex includes mating and maternal manifestations. The data have been used to support the concept that the limbic systems serve motivation and emo-

tion. But when this concept is examined carefully, the support is seen to be spurious. If motivation and emotion are conceived to be viscerally determined reactions, the limbic systems ought to be primarily concerned with visceral regulation. Indeed, the limbic systems have been called the "visceral brain" (MacLean, 1949, 1952, 1955) in order to emphasize this relation. And special relations of motivation and emotion with the autonomic nervous system and the viscera do obtain. But these relations are not selective; other parts of the cerebral mantle (e.g., the motor cortex) also control autonomic and visceral activities, and the control which the limbic systems exercise is certainly not limited to viscera or the autonomic nervous system.

In order to account for data in terms of the emotion-motivation concept, an alternative hypothesis has been proposed—namely, that the limbic systems serve instinctual, innate patterns of behavior, phylogenetically and ontogenetically old (J. D. Green, 1958). This alternative has not always been clearly separated from the visceral hypothesis. Often both are invoked (MacLean, 1958)—the one to account for some facts and the other to be used as soon as the first fails. Support for the hypothesis of instinctual behavior comes from comparative neurology, since some of the structures included in the limbic systems are among the oldest to be found in the endbrain. But this hypothesis also fails to be supported upon close scrutiny. Not all the structures in the limbic systems are old; some are accretions as phylogenetically recent as to appear first in primates. And behavior such as fleeing, tested in a conditioned avoidance situation, is learned and highly specific to the situation. Abnormalities of sexual behavior produced by limbic system lesions in cats have been shown to depend not on hypersexuality per se but on the differences between normal and operated cats as to where sexual behavior takes place (i.e., the territorial range). The experiential components that determine this sort of behavior are not to be ignored.

Neither hypothesis is adequate. The limbic systems cannot be conceived as the neural substrata of motivation and emotion if these are thought of exclusively in terms of visceral-autonomic activities or as old, primitive, innately determined processes (Pribram, 1958*a*). The search for an adequate explanation must continue.

The neurosurgical clinic has inadvertently produced another set of data that bear on this problem. By no stretch of imagination can these data fit simply the rubrics of motivation or emotion. Extensive resections of the medial structures of the temporal lobe of the brain of man—the amygdaloid and hippocampal formations of the limbic systems—result in a very peculiar and dramatic syndrome (B. Milner, 1958). Patients with such lesions are able to repeat correctly a series of digits that they have just heard for the first time. On this test of immediate memory they are practically as efficient as they were before the lesion was made. Moreover, their memory for events prior to their surgical operation is apparently normal. And they are capable of reacting emotionally in trying situations. But, if distracted, these patients are unable to carry out a sequence of actions—that is, they are unable to recall what they are supposed to do. If there is an interruption of a test procedure, the patient will not only be unable to continue where he left off but will, in fact, not even recall that there was any task at all. If the physician should be called from the room for a quarter of an hour, the patient will fail to recall that he had ever seen him before. This patient can be directed to a grocery store where he can purchase the items on a written list without having to refer to that list any oftener than would a person with an intact brain. But once he has completed the shopping, he does not recall what he is supposed to do next and is completely incapable of finding his way home.

Memory is a complicated affair. Not only must the engram be recorded and stored, it must also be accessible when it is appropriate to the occasion. The syndrome shown by these patients can be summarized by the statement that they are unable to recall whatever is necessary to execute a sequence of actions. Given an external plan written out on a piece of paper, they can carry on quite well. Where in the memory process the difficulty lies can only be guessed at present; such guesses have been made (see Brazier, 1958*b*, 1959*a*).

On the surface the defect shown by these patients would seem to have little in common with the disturbances noted in the animal experiments. Perhaps a deeper analysis can show that some common elements exist. The element common to the activities of feeding, fleeing, fighting, and sex is that they are all comprised of sequences of acts (Pribram 1958*a*). After amygdalectomy, the threshold for initiating the sequence is high and the behavior runs inappropriately long once it has been initiated (Weiskrantz, 1956). Feeding behavior may be difficult to initiate; once it has started it is difficult to

stop. After median cortex lesions (cingulate), the maternal behavior of rats is peculiar (Stamm, 1955*b*). When a normal mother rat is faced with a situation in which her brood has been strewn around the cage, she will pick up one baby at a time and carry it to the nest, go back to pick up another and return it to the nest, and so forth, until all the youngsters are safely back in the nest. The lesioned mother rat will pick up her baby rat and carry it to the nest, only to remove it on subsequent trips. After half an hour of this the baby rats are still strewn all over the cage, and eventually are left to die. What would happen if the mother rat could read, the babies were labeled, and the mother were given a written list of directions to plan the retrieval of her brood?

The element common to the patients with limbic system resections and animals who show disturbances of the activities of feeding, fleeing, fighting, and sex seems to be in the execution of certain sequences of actions. Deficiencies appear where the execution depends on some fairly complex plan that has to be carried inside the head. Therefore limbic system lesions can be thought to interfere with behavior because of some defect in the planning mechanism and not because of disturbed emotion or motivation, nor primarily because of some global defect in memory. Limbic system function is thus conceived to be related primarily to the mechanism of the execution of complex sequences of action.

Analysis of the neural mechanism that underlies the execution of sequences of actions has just begun. Electrical changes have been recorded from the amygdaloid complex of the limbic systems whenever the organism is exposed to a novel event or one that has meaning in terms of reward and punishment (Brazier, 1958*b*, 1959*a*; John and Killam, 1959, 1960). These electrical changes subside once the organism is familiar with the event unless the hippocampal formation of the limbic systems has been ablated, in which case electrical changes continue to occur when the event takes place. The amygdaloid complex is necessary to the establishment of electrocortical conditioned responses. The suggestion has been made that the hippocampal formation inhibits (perhaps by way of the reticular core of the brain stem) the succession of unrelated inputs to the amygdala that might occur and so allows this structure to maintain the neural activity necessary to the conditioning process. In a conditioning or learning situation, electrical changes are recorded from the hippocampal formation during the initial trials. Later no such changes accompany successful action; they occur only when errors are made (Adey, 1960; Adey *et al.*, 1960).

In somewhat greater detail the story goes something like this: An organism's exposure to a relatively intense novel environmental stimulus is accompanied by generalized desynchronization in the electrical activity recorded from both isocortex and the basal forebrain (e.g., the hippocampus). This startle reaction, or orienting reflex, is accompanied by behavioral arrest of movement except for head and eye (and perhaps bodily) orientation toward the stimulating event. This initial stage may give way to behavioral fight or flight.

Of particular interest here are those occasions where the original stimulus remains or is repeated and the organism becomes familiar with the events. In these instances desynchronization of the isocortically recorded electrical activity continues, but hypersynchronous slow waves are now obtained from the basal forebrain. When such hypersynchrony is experimentally induced in these structures, ongoing problem-solving behavior is interfered with (MacLean *et al.*, 1955–56), much in the same way as when these structures are surgically removed (Pribram and Weiskrantz, 1957; Hunt and Diamond, 1957). As already noted, animals with such lesions are hyperreactive to novel stimulation. The assumption is therefore made that the slow activity in the basal forebrain reflects the cessation of its *usual* gating action on the central effects of novel stimuli. (The gating action is conceived to take place via the brain-stem reticular formation.) This second stage is called the orienting *reaction* (differentiated from the orienting reflex). It is characterized by heightened behavioral orientation and attention to *all* aspects of the environment.

With repetition of the situation, a third stage sets in, when the organism is said to habituate. This stage is again characterized by both isocortical and basal forebrain electrical desynchronization, though the isocortical manifestation is now no longer generalized but relatively restricted, in the experimental situation, to the cortex subserving the sensory mode through which the environmental stimulus has been presented.

Even more impressive evidence for the time course of the neural activity involved in this process has been demonstrated by John and Killam (1959, 1960), who employed electrical "tracers" in the form of visual stimulus frequencies to which the brain's electrical activity becomes locked. Their

tracer frequencies are found generally in recordings made from electrodes implanted in allo- and isocortical formations during the initial stages of the problem-solving behavior of cats. As the experiment proceeds, the tracer frequencies become more limited in distribution until, during error-free performance, they are recorded only from the isocortical systems (geniculostriate) concerned with vision.

Furthermore, when such a problem-solving situation is used, Adey (1960; Adey *et al.*, 1960) has shown, by a beautiful series of experiments, that basal forebrain electrical activity (recorded from the hippocampus) is also characteristically different in the initial "startle" stage of the experiment and the final "conditioned" stage. As habituation proceeds, a shift is recorded in electrical phase of the activity of the several layers of the hippocampal cortex. In the "startle" stage the electrical activity of the layers that are connected with the brain-stem core (including the reticular formation) precedes that recorded from the layers more immediately connected to the isocortex; in the "conditioned" stage the phase relationship is reversed. In this final stage of the problem-solving situation the behavior of the habituated organism is appropriate to the task—that is, performance is approximately errorless. When occasional errors do occur, they are accompanied by the recrudescence of slow activity in the electrical record made from the basal forebrain.

But what is this "habituation" (Sharpless and Jasper, 1956) that thus becomes so all-important? Is it merely the fatiguing-out of the neural mechanism of attention? The indications from human experiments are to the contrary. Sokolov (1960), in an exquisitely designed and performed series of experiments, has demonstrated that habituation in man results to the extent that a neural representation of the stimulus is built up in the nervous system. When the input (e.g., a tone) matches this representation, no characteristic behavioral, autonomic effector, or EEG responses can be recorded. When the input departs from prior inputs (e.g., diminution of the intensity of the tone, or making the tone shorter or longer), it re-evokes the orienting responses (behavioral, autonomic effector, and EEG). This re-evocation is limited to the specific occasion—and only for its duration—during which current input is disparate from prior inputs.

Taken together, the electrophysiological and neuropsychological evidence points to specifiable stages that can be summarized as follows:

SUMMARY

Stage 1. When exposed to a novel event, an organism "takes this in." The stage is accompanied by desynchronization of the electrical activity of both the isocortical and basal allocortical formations of the endbrain. The only behavioral concomitants are "reflex" orientation movements that focus the stimulating event. Lacey (see Brazier, 1959*a*) has noted that this stage corresponds pretty much to "primary attention" as this was defined in introspective psychology.

Stage 2. Should the novel event recur repeatedly, remain unchanged, or change relatively slowly, another process supervenes. It is characterized by continued desynchronization in the electrical activity recorded from the isocortex, but also by a change in the activity recorded from allocortical structures (especially of Ammon's formation). From this neural location, slow waves (i.e., hypersynchrony) can now be recorded. Behaviorally, searching characterizes the activity of the organism. This is the orienting reaction. The organism follows the stimulating event and searches when changes occur, especially once habituation is under way. In many respects the orienting reaction is similar to the "secondary attention" described by the introspectionists.

Stage 3. After repeated exposure to the unchanging or recurrent event, habituation has resulted. The desynchronous electrical activity recorded from isocortex has become restricted to relevant input channels, and slow activity has disappeared from allocortical structures. Here electrical phase has shifted from precedence of brain-stem input to precedence of input from isocortex. And any noted change in the situation is immediately and specifically accompanied by recrudescence of the electrical activities in both the iso- and allocortex characteristic of Stage 2.

During Stage 3 the actions of the organism directed toward the stimulating event may be included in the habituation process. This obtains in conditioning and other problem-solving situations. So, orienting responses re-emerge when the organism's actions result in changes from the recurrent regularities that characterize the total stimulus event—for example, when the oft-obtained food "reinforcing" stimulus fails to appear, or when the for-the-most-part-eliminated shock stimulus fails to be avoided. In other words, the organism now has a mechanism for sensing error or incongruity![1]

[1] The prediction can be made from this formulation that only when a reinforcing event occurs at a perceptibly dif-

IMPLICATIONS FOR SYSTEMATIC STUDIES OF BEHAVIOR

But even more can be stated about the neural mechanism involved in reinforcement: The representation that is set up during habituation is not formed by virtue of the limbic formations only. The isocortex is intimately concerned in the process. We react to environmental patterns as if they were relatively stable configurations. Yet, because of movements made by the organism and changes that take place in the environmental events, receptors are activated by ceaselessly altering patterns of energy. The constant or invariant properties of the proximal stimulus array must therefore be extracted somehow from these changing patterns. It is the function of the projection systems to accomplish the extraction. Concerning the particular way in which this happens, some precise guesses have been made in the form of mathematical models and by the simulation of simple neuronal networks in automata studies (Whitfield, 1960). What is certain, however, is that the projection systems make it possible for an organism to respond to the invariant properties of receptor stimulation. For, as Klüver (1941) has demonstrated so elegantly for vision, lesions of the projection systems lead to changes in behavior that can be described as follows:

Ordinarily, an organism responds selectively to certain properties of the stimulus (e.g., contour, brightness). The range of transformations of these properties (e.g., changing a circle to an ellipse) over which the organism will still make the same response is relatively restricted. After ablations of the occipital cortex, monkeys will respond indiscriminately when the stimulus properties are transformed over a very wide range; only the amount of total luminous flux of the energy that activates the retina is now distinguished. Differences of flux that result in changes of luminance, contrast, and contour have no effect in altering the monkey's reaction. In the absence of the projection system mechanism, the organism cannot extract restricted invariant properties from the retinal excitation. It shows a defect in sensibility, in making existential discriminations.

When the functions of the posterior intrinsic systems are interfered with, existential discriminations remain intact. A monkey with such a defect will catch a gnat in mid-air (Blum *et al.*, 1959). He will react to variations of illumination by varying his rate of response in an operant situation (Ettlinger and Wegener, 1958). But when he is given alternative responses to make to differences in luminance, patterns, etc. (the alternative indicated by a peanut, for example), he fails to make any consistent choice (Chow, 1951, 1952; Mishkin, 1954). It is apparent, however, that even this complex relation between cues, alternatives for response, and the indicators of these alternatives allows some invariant properties to be extracted. Otherwise, intact monkeys and people would not be able to respond consistently. Somehow, through repetition in time, these properties are identified; and, when the functions of the posterior intrinsic systems are interfered with, identification goes awry. The situation becomes unintelligible to the organism; it cannot make the differential discrimination; it does not know what to do; it has an agnosia.

The posterior intrinsic system must, in some way or other, make possible the separation of the invariant properties that characterize the situation in contrast with other, less regular, variables. This can be accomplished if some coded representation of the invariance is established in the posterior intrinsic mechanism. The neuronal patterns that form the representation can be conceived to be of the sort that Hebb (1949) has popularized as "cell assemblies" with additional inhibitory properties such as those used by P. M. Milner (1957) in his Mark II modification. More in accord with the known complexities of the neural net are the neuronal patterns as characterized in the manner suggested by Beurle (1956), who bases his model on the cytological work of Sholl (1956). Beurle has worked out in some mathematical detail a mechanism of cortical function based on interacting waves of excitations or "interference patterns" that were first proposed by Lashley (1952) and would result in the units proposed by Loucks. However the representation may be produced, it must be formed according to some rules established either by the innate structure of the nervous system or through experience. These rules proscribe constraints on the otherwise random properties of the neural network. The rules themselves—that is, the properties of the network and therefore of the representation—have efferent control over the input variables (Pribram, 1960*b*). In other words, they operate by selective attention. What is selected could depend on some kind of match-mismatch process as described by MacKay (1956) and by Bruner (1958).

The selective process would provide the mechanism required by Harlow (1959) in his theory of discrimination learning. This mechanism also serves

ferent level or in a perceptibly different pattern from average prior occurrences, will it guide subsequent behavior.

the purposes needed by J. J. Gibson (1959) to account for his findings that discrimination learning is a process of progressive differentiation, not a process of association. But perhaps the model has its greatest power in the description of what constitutes reinforcement for the organism. The posterior intrinsic mechanism, because of the hierarchical nature of its selective control over its own modification, allows a change in the representation to occur by trial and error. Whenever the error signal is such that the corrective change is not uniquely specified, the representation is modified to include this information, and trials continue. Thus an organism that possesses this mechanism can, given a relatively unchanging or slowly changing environment, search that environment for the additional information that is needed to make the organism fully informed. The neural model would account for the search through negative instances as these are defined in the stimulus-sampling type of stochastic learning theories (Bush and Mosteller, 1955; Estes, 1959; E. J. Green, 1958)—that is, a search by an information-hungry organism reinforced or satisfied only when corrective change of the representation is immediate and can be deduced uniquely from the error signal. When stated in this way, reinforcement becomes one side of the coin of similarity.

The implications of these neuropsychological efforts are, of course, enormous. After a quarter century of relative quiescence, neurobehavioral studies are again *the* avenue that can lead both psychologists and neurophysiologists to sensible choices among disparate alternatives produced by the prodigious stores of data. The ideational drought that extreme positivism and classical behaviorism imposed on systematic studies of behavior is over. The excesses of a romantic approach to the mind-brain problem have withered during that drought—so it was not in vain.