Switchboard versus Statistical Theories of Learning and Memory

Coherent patterns of neural activity reflect the release of memories and may mediate subjective experience.

E. Roy John

Early theories of learning were based largely upon studies of conditioning. In such experiments, techniques were often used that depended basically on the establishment of stimulus substitution between an arbitrarily selected conditioned stimulus (CS) and an unconditioned stimulus (US) which caused some measurable response.

The reasoning of many early workers (especially American behaviorists) might be paraphrased as: "There is an input region which receives the CS and an output region which produces the response to the US. During learning, this input somehow comes to produce that output as a conditioned response. Therefore, a new connection must be established between the CS and US regions; some kind of pathway is built." The conviction that training wore a "groove" of increased excitability along specific neural paths from sensory input to motor output, and that the existence of that groove of cells was the memory of the experience, launched an avalanche of studies of lesions in which investigators sought the locus of the postulated connection. Because the critical event in learning is envisaged as the formation or facilitation of specific connections, such theories are here referred to as switchboard theories. An essential feature of most such theories is that "remembering" requires the discharge of those particular cells which constitute the new line, and of those cells to which the line is directed. In that sense, such theories are also place theories, in which it is assumed that a memory is localized in a discrete set of cells reserved for that function.

Switchboard Theories: Pros and Cons

The switchboard model of learning has long seemed intuitively plausible (*1*). Because the studies of workers such as Lashley (*2*) pinpointed the inadequacy of theories based on simple connection formation between an input and output region, switchboard theories were modified to explain memory as the result of connections between multiple pathways in various regions of the brain (*3*). Although some workers (*4*) used the term "connection" in a functional sense, cautiously defining it in terms of relationships while explicitly disclaiming any literal connotation of a new anatomical junction, many subsequent theorists have not been equally careful. More recently, new responses have commonly been assumed to depend upon the establishment or facilitation of a new pathway (or pathways) of synaptic connections between input and output (*5, 6*).

Most of these switchboard formulations are explicitly localizationist. For example, Ungar (*6*) asserts that the basic elements of acquired information are coded in terms of "labeled" lines or reserved pathways, and that the decision as to whether or not a specific coded event has taken place depends on the pathway in which a neural impulse occurs. Further, he suggests that identical or homologous pathways are responsible for the same patterns of learned behaviors in different individuals.

In view of the absence of any convincing neurophysiological evidence that permanent changes in synaptic efficiency actually occur or that new connections are actually established in learning, it seems incongruous that these ideas have come to be so widely accepted (*7*).

The idea, developed by analogy with models of the conditioned response, that learning establishes a new pathway and that remembering consists of firing in that pathway, is suspect on logical, psychophysiological, and neurophysiological grounds.

The conditioned response (CR) represents a limited class of learning situations. Survival would be improbable if learning in nature required the lengthy repetition characteristic of most conditioning procedures. Not only does extremely rapid learning occur in man, but it has been amply shown that rats, cats, monkeys, and other animals can acquire complex new behaviors without rehearsal, merely by watching other animals perform (*8*). Integrative processes of this type are particularly difficult to reconcile with most popular switchboard theories.

Psychophysiological Considerations

Countless ablation experiments in which various parts of the brain have been destroyed or removed have failed to locate the site of the connection of any pathways responsible for memory (*2*). Detailed analyses of losses of retention reported to be long lasting have usually revealed the deficits to be results of interference with performance caused by changes in general factors such as sensory deficit, motivation, attention, or set, rather than loss of a specific memory. Lesions reported to block retention of a CR often block acquisition as well. Upon retraining, such animals usually reach criterion in substantially fewer trials than initially required, and recovery can occur as a result of other learning (*9*) or even mere passage of time. Further, in many studies including quite recent work (*10*), it has been demonstrated that deficits in retention or acquisition following bilateral ablations or lesions performed during the same operation are markedly diminished or absent if the same lesions are inflicted in multiple operations. It may well be that the brain compensates for the loss of information about one aspect of a complex stimulus, severely disturbed by a localized lesion, by gradually utilizing other aspects of the complex stimulus whose representation resisted the brain insult. It must be conceded that certain functions, especially those based upon complex sensory integrations, can be severely perturbed by localized brain damage. However, it is difficult to point to any unequivocal demonstration that

---

Dr. John is professor of psychiatry, professor of physiology and director of the Brain Research Laboratories in the Department of Psychiatry at New York Medical College, New York 10029.

memory for some specific events was localized in a particular restricted volume of brain tissue.

Lashley (*2*) was finally led to formulate his well known laws of mass action and equipotentiality, asserting that the loss in learned response following ablation would be a function of the volume of brain tissue removed rather than its location. Yet, even if a specific region were conclusively implicated as the locus of a memory, this would not establish that the memory was mediated by the occurrence of discharge in a unique set of cells constituting a specific pathway within that region. Switchboard and localizationist theories of memory cannot be validated merely by invalidating the laws of mass action and equipotentiality.

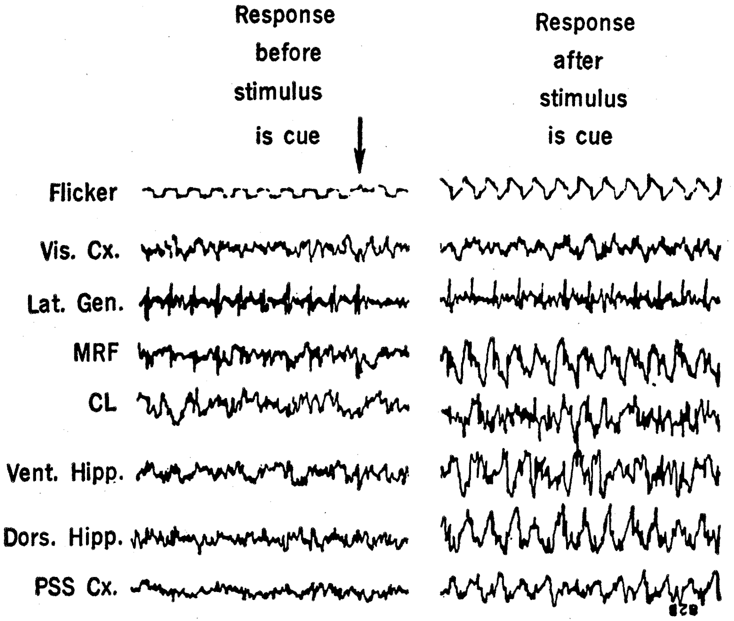
Data which are perhaps the most difficult to reconcile with switchboard theories have resulted from other recent studies of lesions (*11–13*). In these studies, the visual cortex was completely removed or extensive multiple lesions were made in the optic nerve and at other levels of the visual system so that point-to-point, retinocortical pathways were severely or completely disrupted. Little disruption of pattern vision was observed.

The remarkable preservation of visual discriminations, in spite of the massive rerouting of afferent input undoubtedly achieved by these lesions, argues strongly against the belief that these learned behaviors depend upon isomorphic pathways. These results are reminiscent of Lashley's early failure to find deficits in learned responses when animals were forced to perform with limbs which had been immobilized during training (see *2*). Further, the striking absence of behavioral deficit after such severe disruption of the retinocortical system must raise serious doubts as to whether the "feature extractor" properties and hierarchial organization revealed for cells in this topological matrix by many studies are necessary for pattern vision.

Neurophysiological Considerations

Prominent anatomical pathways transmit sensory information centrally from peripheral receptors, while other pathways mediate the activation of specific effectors. Great precision and discreteness have been demonstrated in the central connections of these input and output systems (*14*). Such considerations have led many workers to infer that certain cells in the brain must serve as "feature extractors," reliably reporting the occurrence of a specific event in the environment by a neural discharge. The inference that particular percepts, or even concepts, could be represented by localized firing of a cell served to reinforce beliefs in switchboard theories.

A number of recent studies with microelectrodes reveal great variability and nonspecificity in neuronal responses to stimuli, which are extremely relevant to the evaluation of place theories. Burns and Smith (*15*) reached two im-



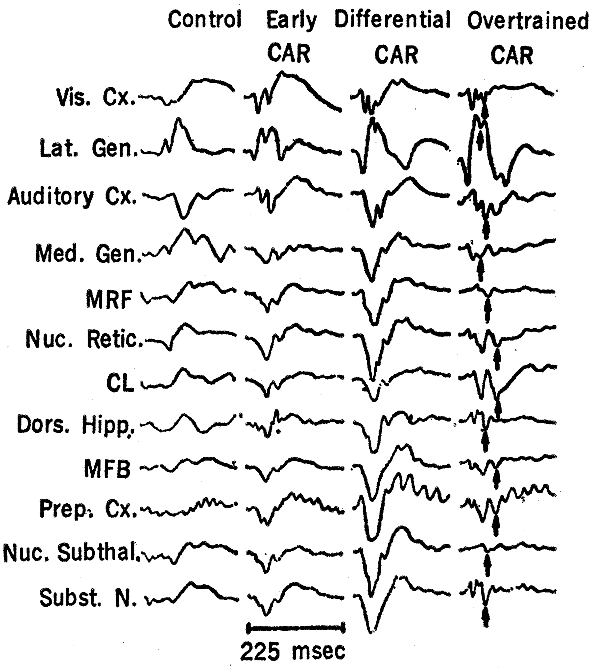


Fig. 1 (left). *Response before stimulus is cue* shows the effect of photic (6-hz flicker) stimulation on a cat after it had learned that milk could be obtained whenever a lever was pressed. The flicker had no signal value at this stage. Little "labeled" activity was elicited by the flicker except in the lateral geniculate (*Lat. Gen.*). The response disappeared in the lateral geniculate because of the internal inhibition which occurred as the cat pressed the lever and waited for milk. *Response after stimulus is cue* shows records during presentation of flicker (6 hz) after the cat had learned a frequency discrimination. The cat was responding correctly by not pressing the lever down. There is marked enhancement of labeled responses at the stimulus frequency (*Flicker*, 6-hz photic stimulation; *Vis. Cx.*, visual cortex; *MRF*, mesencephalic reticular formation; *CL*, nucleus centralis lateralis; *Vent. Hipp.*, ventral hippocampus; *Dors. Hipp.*, dorsal hippocampus; *PSS Cx.*, posterior suprasylvian cortex). [Data from John (*26*)] Fig. 2 (right). Evolution of visual evoked responses. *Control*, average responses evoked in different brain regions of a naive cat by presentation of a novel flicker stimulus. Several regions show little or no response, and different regions display differing types of response. *Early CAR*, responses to the same stimulus shortly after elaboration of a simple conditioned avoidance response (CAR). A definite response with similar features can now be discerned in most regions. *Differential CAR*, changes in the response evoked by the flicker CS shortly after establishment of differential approach-avoidance responses to flicker at two different frequencies. As usual, discrimination training has greatly enhanced the response amplitude, and the similarity between responses in different structures has become more marked. *Overtrained CAR*, after many months of overtraining on the differentiation task, the waveshapes undergo further changes. The arrows point to a component usually absent or markedly smaller in behavioral trials on which this animal failed to perform (*Nuc. Retic.*, nucleus reticularis; *MFB*, median forebrain bundle; *Prep. Cx.*, prepyriform cortex; *Nuc. Subthal.*, nucleus subthalamus; *Subst. N.*, substantia nigra).

portant conclusions: first, that all single neurons observed were in incessant activity. Second, that every cell which was monitored could be demonstrated to alter its firing pattern as a consequence of any arbitrary stimulus which was presented. Gerstein and Kiang (*16*) studied the temporal firing patterns of units in the auditory cortex as a function of stimulus parameters. Although a given unit might respond with a statistically repeatable pattern to many presentations of a particular stimulus, great variability in response to individual presentations could be observed. Further, a particular unit might respond to various stimuli with different response patterns. Doty (*11*) has provided similar data about the ambiguity of the information provided by maximal response rates in "feature extractors" in the visual system.

Evaluating such data, Gerstein (*17*) has commented that the firing pattern of a single neuron cannot carry enough information to account for the discriminative or behavioral abilities of the animal, suggesting that the processing of sensory information, and other complex integrative functions, must be mediated by simultaneous activity in ensembles of neurons, organized into functional groups with boundaries and relationships which may vary from task to task and from time to time.

Cells responsive to inputs from multiple sensory modalities have been reported to range from 18 to 92 percent, in various brain regions. O'Brien and Fox (*18*) have reported that about 80 percent of the cells which they studied in the motor cortex responded to stimuli in more than one sensory modality. Different response patterns were displayed to different stimuli. In the visual cortex, Morrell has found that about 92 percent of the cells with particular stimulus "preferences" also responded to stimuli in other sensory modalities (*19*). Again, each effective stimulus elicited a different response pattern. Numerous workers (*20*) have presented evidence that the supposed specificity of response of "feature extractors" is grossly impaired by altering variables in the environment or the context of the stimulus. Analogous findings show that the details of motor performance cannot be accurately predicted from the firing patterns of single neurons in the motor output regions (*21*). Study of these findings suggests that temporal relations between the firing patterns of an ensemble of cells might be a more important variable in the control of movement than the

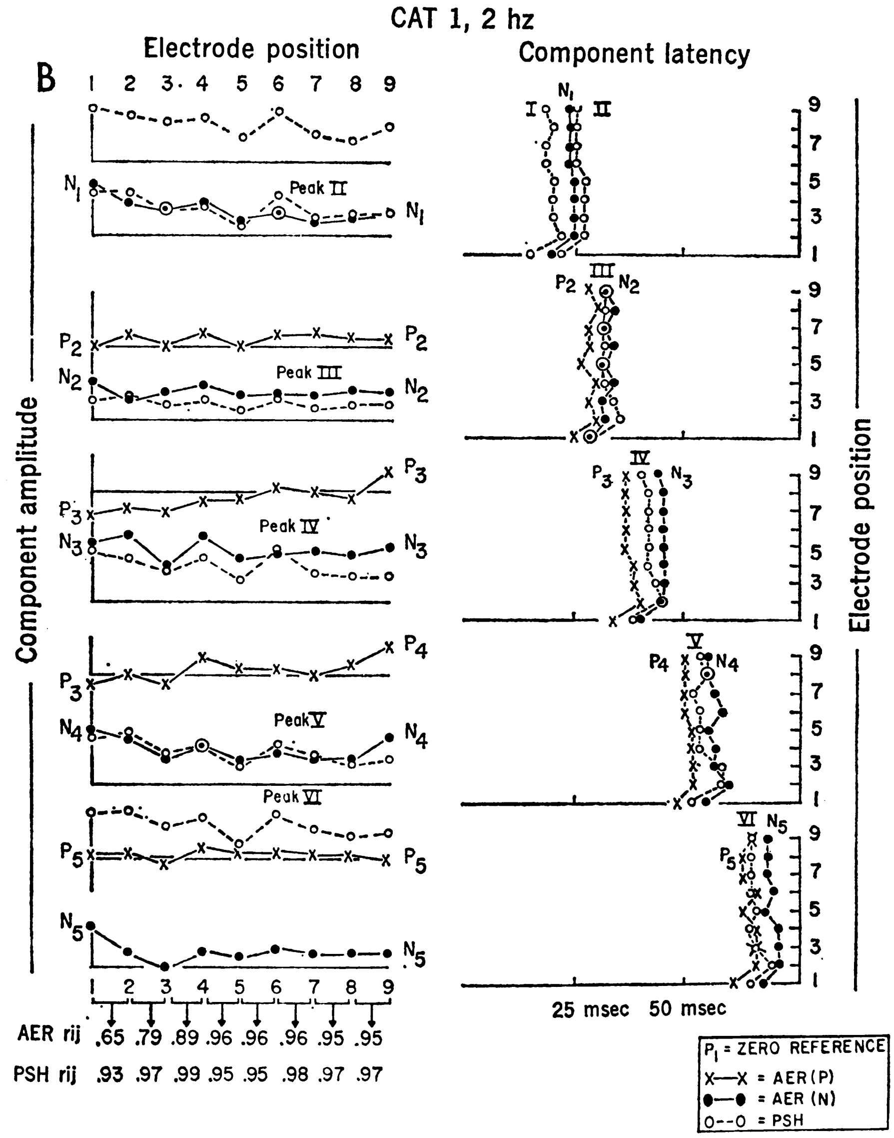


Fig. 3. (A) "Multidrive" instrument containing five independently movable microelectrodes mounted on head of cat. (B) Graphs on the left illustrate the amplitude gradients of positive (*P*) and negative (*N*) AER components and PSH peaks (see below). Gradients computed from data recorded in response to a 2-hz flicker CS presented to a differentially trained cat, which pressed the left lever on a work panel to get milk whenever this stimulus occurred. Average evoked responses (AER) and post-stimulus histograms (PSH) at each electrode position were based upon 500 stimuli presented in multiple behavioral trials. Successively later components are depicted from top to bottom graph. In each graph, amplitude of the component is plotted as vertical displacement, while depth of electrode penetration is plotted as horizontal displacement. At the bottom are presented the correlation coefficients (*rij*) between response waveshapes at adjacent electrode positions. Graphs on the right show the latency distribution of the positive and negative AER component and the PSH peak whose amplitude gradients are found in the graph at the same level on the left side of the figure. Component latency is plotted along the abscissa, while depth of penetration is represented along the ordinate. [Data from E. R. John and P. P. Morgades (*32*)]

particular type of firing pattern exhibited by any single neuron in the motor cortex.

Thus, although central sensory cells may respond most to certain "preferred" stimuli, they fire in response to a variety of stimuli as well as spontaneously and may display maximum response rates to stimuli with different features. Their response to a specific stimulus is not invariant, but depends on a variety of conditions. Similarly, the firing of a motor cell is not related to a movement in an invariant manner. Activity in a particular cell, per se, cannot be uniquely interpreted. Whether concerned with sensory report or motor control, the single cell is unreliable and ambiguous.

Such findings raise perplexing questions. If nerve cells in the central nervous system fire spontaneously with high incidence, if they respond in a variable way to any given stimulus, if they respond to multiple stimuli in various sensory modalities, how is the brain to evaluate the significance of a discharge by a specific cell in a particular pathway? What is to distinguish between neural events related to spontaneous discharge, the input of information about a novel stimulus, and input about a familiar stimulus? How can bursts of activity deviating from a lower baseline rate indicate detection of a "preferred" stimulus if the cell produces maximum firing rates to different combinations of stimulus parameters? The invocation of the notion that information is coded by activity in multiple parallel pathways does not solve this problem, if the multiple pathways are envisaged to converge upon a cell rather than an ensemble. Further, of what avail would such a pathway be if the motor cell to which it led could not effectively control the motor response?

It is unrealistic to think in terms of a pathway between sensory and motor regions in which neurons rest or fire at a baseline rate until a specific event occurs, whereupon they burst into activity which represents that event and affects some feature extractor or event detector which then causes the appropriate response. It seems highly unlikely that single cells or sets of cells become labeled by experience so that their subsequent discharge will uniquely represent that experience. The "filing card" concept of memory, in which discharge or the burst of activity of a cell represents the selection of the card as well as the information written on it, does not seem plausible. Place theories oversimplify the problem of how the central nervous system extracts information about stimuli and their significance from the mass of neuronal activity—relevant, irrelevant, redundant, or random.

As we realize that information is probably not uniquely represented by the discharge of any specifiable cell, the idea that stored or acquired information is represented by discharge in a newly connected pathway becomes suspect. Switchboard theories are dependent upon localizationist assumptions.

Studies of Single Cells in Learning

There has been an increasing flow of information about cellular events in learning, yielding evidence that the response of single cells is plastic, under circumstances permitting assessment and analysis of mechanisms mediating this plasticity (*22*). While one may recognize the value of such studies, certain reservations must be voiced: (i) the changes thus far described have been of relatively brief duration, lacking the persistence of learning; (ii) mechanisms of neural modification in simple preparations of the type often studied are not necessarily relevant to more complex systems; and (iii) most important, the fact that cells modify their activity in learning situations does not suffice to validate switchboard theories. No one assumes that learning occurs magically, without any change in cellular activity. The critical issue is whether the cellular change in learning labels the discharge of the cell to represent a unique event. Let us hope that the relevant changes are of the sort now being studied by investigators working with the single cell. However, the fact that cellular changes occur in learning does not prove that learning forms connections leading to neural activity in specific new pathways, nor that localized firing in such pathways stands for a previous experience.

Investigators of the effects of conditioning on single cells in higher organisms are unanimous in stating that in numerous brain regions the response of a large proportion of cells (10 to 70 percent) is changed during conditioning (*23–25*). The process of learning seems to influence enormous numbers of cells in many brain regions, even with the simple tasks studied. These percentages are so high that the activity of any specific neuron must be modified by a large number of experiences. The profuseness of changes makes it even harder to imagine the mediation of learning by discrete pathways. How could a new pathway be protected when every new experience affects so many cells?

Single cell studies of conditioning in higher organisms have provided no support for the notion that conditioning establishes new connections. On the contrary, it is those polysensory cells which originally responded to both the CS and the US that seem most likely to alter response during conditioning (*26*). Such observations led O'Brien and Fox (*24*) to the following conclusion:

> Cells which do not initially respond to the CS might have been expected to develop some response during a conditioning series, thus "forming new connections." The data give little support to this viewpoint since cells which did not originally respond to the CS usually did not develop any response during the conditioning series. This suggests that learning behavior occurs via modification of an already functional structural organization. No new pathways appear to be formed, but there is, rather, simply modification of existing circuits.

In studies relevant to those conclusions, Morrell *et al.* (*25*) showed that cells responsive to two different stimuli displayed an altered pattern of activity when the two stimuli occurred together. Subsequently, presentation of one of the stimuli alone elicited the pattern seen during pairing. The neural network has the capability of storing and releasing a particular firing pattern which has repeatedly occurred.

Statistical Configuration Theory

These considerations and related data led me to propose (*26*) an alternative to switchboard theories: the statistical configuration theory. The critical event in learning is envisaged as the establishment of representational systems of large numbers of neurons in different parts of the brain, whose activity has been affected in a coordinated way by the spatiotemporal characteristics of the stimuli present during a learning experience. The coherent pattern of discharge of neurons in these regions spreads to numerous other regions of the brain. Sustained transactions of activity between participating cells permit rapid interaction among all regions affected by the incoming sequence of stimuli as well as the subsequent spread. This initiates the development of a *common mode of activity*, a temporal pattern which is coherent across those various regions and specific for that stimulus complex. As this common mode of activity is sustained, certain changes are presumed to take place in the participating neuronal populations, which are thereby established as a representational system. Whether such changes are alterations of "synaptic efficiency" or not, it is assumed that the critical feature of these changes is to increase the probability of recurrence of that coherent pattern in the network. Certain types of preexisting neuronal *transactions* become more probable, but no new connections are assumed to be formed.

This theory is statistical, in that the informational significance of an event is represented by the average behavior of a responsive neural ensemble rather than by the exclusive behavior of any specifiable neuron in the ensemble. The same ensemble can represent many different items, each with a different coherent pattern of deviation from randomness or from its baseline pattern. The theory is configurational in that new responses are based upon the establishment of new temporal patterns of ensemble activity, rather than upon the elaboration of new pathways or connections. Learning increases the probability that particular temporal patterns of orderly activity will occur in coupled ensembles of neurons. By this process, the representational system acquires the capability of releasing the specified common mode of activity as a whole if some significant portion of the system enters the appropriate mode.

It should be emphasized that this is not a "field" theory, nor does it deny the highly organized structure of the brain. The firing pattern of neural ensembles undoubtedly depends upon connections between neurons. Configurations of activity in representational systems are presumed to become established by modification of interneuronal relationships, perhaps by changes located at the synapse although other alternatives are conceivable. However, in this laboratory we doubt that memory is based upon the establishment of new connections rather than upon the modification of existing relationships. We consider representation of information by statistical features of temporal patterns of ensemble behavior more likely than by the localized activity of specific cells.

Experimental Support for Statistical Configuration Theory

Studies of neural activity in animals acquiring and performing CR's are particularly relevant for evaluating the statistical configuration theory.

Widespread changes in evoked responses occur during learning. When a meaningless intermittent stimulus, serving no cue values, is presented to an alert animal performing an uncued instrumental behavior, evoked responses with different waveshapes appear in a few regions of the brain. If that intermittent stimulus is then used as a "tracer" conditioned stimulus to cue the same instrumental behavior, it comes to evoke "labeled" responses at the frequency of the stimulus in many brain regions (*26, 27*). Since these widespread responses are not observed in the performing animal before the behavior is brought under stimulus control, they cannot be attributed to "pseudoconditioning." These changes can be seen in Fig. 1.

Similarity or "common modes" in different regions. The anatomically extensive labeled responses to the CS after training seem to be grossly similar in a number of different brain areas. These similarities have been noted by many experimenters when both subjective and

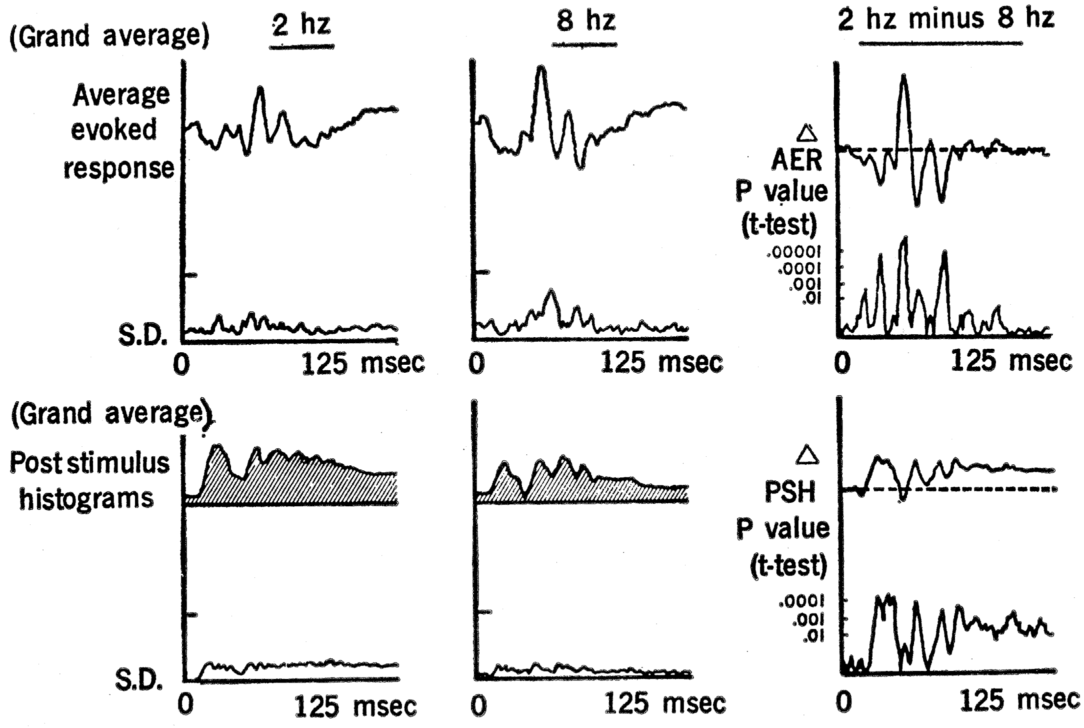


Fig. 4. (Top left) The top curve shows the grand average of the AER's elicited by the 2-hz CS across all electrode positions in the mapped region, while the lower curve shows the standard deviation (S.D.) of the group of AER's. (Bottom left) The top curve shows the grand average of the PSH's elicited by the 2-hz CS across the same electrode positions and the lower curve shows the S.D. (Top center) The grand average of the AER's elicited by the 8-hz CS and the corresponding S.D. (Bottom center) The grand average PSH elicited by the 8-hz CS, and its S.D. (Top right) The top curve shows the difference waveshape resulting from the subtraction of the grand average AER elicited by the 8-hz CS from the grand average AER elicited by the 2-hz CS. The lower curve shows the *P* value, as computed by the *t*-test, for each point of the difference wave. (Bottom right) The top curve shows the difference waveshape resulting from the subtraction of the grand average PSH elicited by the 8-hz CS from the grand average PSH elicited by the 2-hz CS. The lower curve shows the *P* value for each point of the difference. [Data from E. R. John and P. P. Morgades (*32*)]

quantitative criteria have been used (*26, 27*). The evolution of similarity in the modes of activity elicited in different neural ensembles by a CS is illustrated in Fig. 2. In many regions, the waveshape of the response which emerges during conditioning resembles a tilted "W," with two negative (downward-going) peaks. In addition to the findings in this laboratory, it has been reported in numerous studies that a new late component appears in the evoked response after conditioning (*28*). As conditioning progresses, the waveshape becomes more complex and detailed in some structures, and the later negativity becomes more pronounced. The similarity of electroencephalograph (EEG) rhythms or evoked response waveshapes shown in Figs. 1 and 2 is interpreted as a reflection of the common mode of activity established by the learning experience in the representational system, sampled by electrodes in different brain regions. The ease of detecting such labeled responses in so many different regions, both sensory-specific and nonspecific, suggested that very large numbers of neurons must be involved.

Orderly firing patterns in extensive neural ensembles. Since many workers have found high intercorrelations between the waveshape of the average evoked response, or slow wave activity, and the discharge of neural units within the same domain (*29*), it was expected that in trained animals the average time course of discharge of some neural ensembles after a CS would reflect the time course of fluctuations in the average evoked responses. In order to examine the nature of the neural activity in different brain regions when a CS elicited similar evoked potential waveshapes, and to rule out the possibility that such similarity might reflect distant events being picked up by the electrodes because of volume conduction rather than local neural events arising in the immediate vicinity of the electrode tips, a "multidrive" instrument consisting of two to five independently movable microelectrodes was developed (*30*). This instrument, which is permanently implanted, can record simultaneously the evoked potentials and the discharges of the group of neurons close to each of the microelectrode tips as they are moved slowly, over a period of many months, through the brain of an unrestrained animal. A cat carrying such an instrument is shown in Fig. 3A.

Studies with such instruments (*31, 32*) revealed that: (i) positive and negative peaks in evoked potentials correspond to maxima and minima in the curve describing probability of firing in the neural population as a function of time after the stimulus (poststimulus histogram, or PSH), confirming the findings of Fox and O'Brien and Verzeano *et al.* (*33*); (ii) characteristic evoked potential and PSH waveshapes were elicited by a given CS across a very large anatomical domain. Gradients of evoked responses and neural discharge throughout the mapped domain are illustrated for a trained cat in Fig. 3B. The flatness of these gradients indicates that both evoked responses and neural discharges with the same average temporal pattern are widely distributed throughout the region.

Dependence of firing pattern on stimulus rather than electrode position. Variations between the response patterns caused by two differentiated stimuli at any place in the mapped domain were greater than variations within the set of responses caused across the domain by either stimulus. These results are shown in Fig. 4 and indicate that huge numbers of neurons distributed across extensive anatomical regions shared similar statistical features in their temporal patterns of response to a CS. Characteristic firing patterns to a particular CS were basically the same from place to place, while in any place the two differentiated stimuli elicited significantly different firing patterns.

Statistical invariance and individual variability. In those studies (*31, 32*) single cells were found to display extremely variable responses to a given

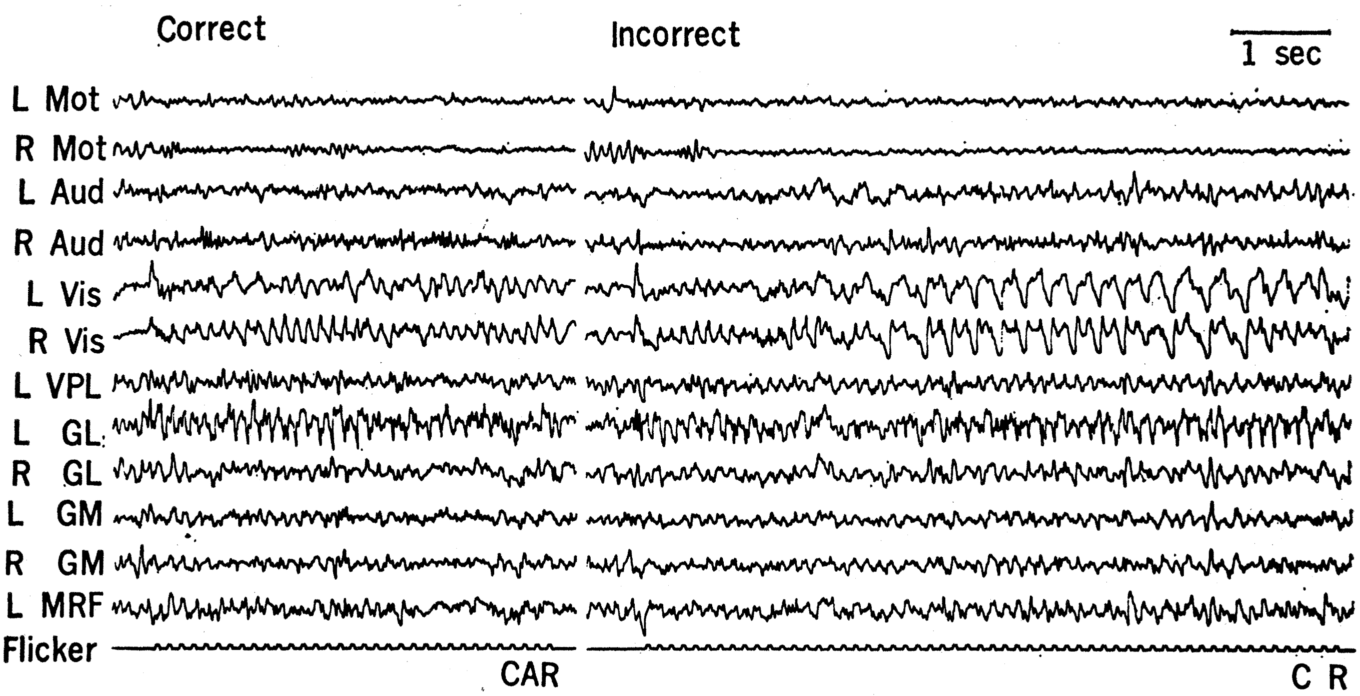


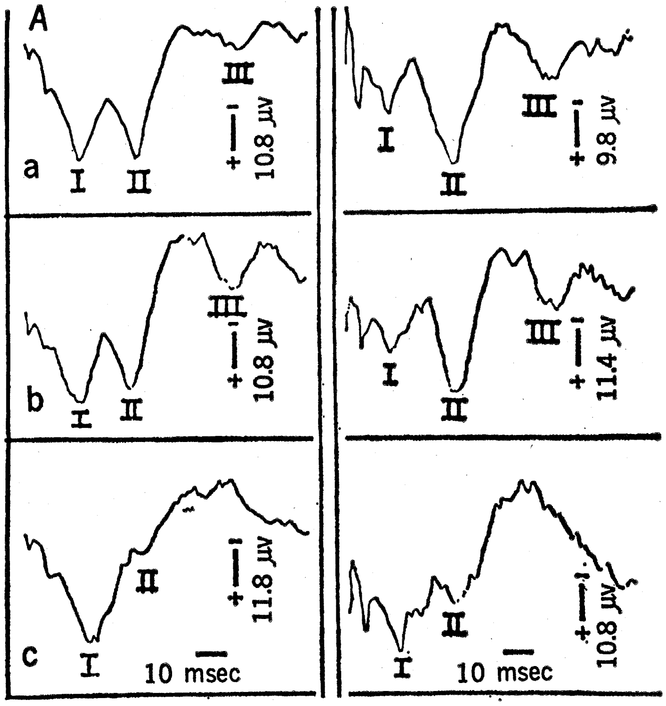
Fig. 5. Comparison of electrophysiological activity elicited by the same visual signal interpreted in two different ways. Records on the left were obtained as the differentially trained cat responded correctly to a negative CS (7.7-hz flicker) by performing a conditioned avoidance response. Records on the right were obtained on the next trial, when the cat erroneously responded to the same flicker frequency by performing the conditioned approach response appropriate to the positive stimulus (3.1-hz flicker). (*Mot,* motor cortex; *Aud,* auditory cortex; *Vis,* visual cortex; *VPL,* ventroposterolateral nucleus; *GL,* lateral geniculate nucleus; *GM,* medial geniculate nucleus; *MRF,* mesencephalic reticular formation.) All records bipolar.

CS, while ensembles of cells displayed essentially invariant patterns. This suggested that the short-term average of an ensemble might represent a mean pattern to which the long-term activity of single elements of the ensemble would converge.

In current studies in these laboratories directed at this question, techniques for discriminating spike heights have been used to fractionate a multineuronal ensemble into small subgroups or single cells. The PSH's elicited by a specified stimulus from the different subgroups within the ensemble are essentially identical with respect to firing pattern, which varies from stimulus to stimulus. Short-term comparison of the activity within separate channels of the spike-height discriminator indicates that the various elements in the ensemble are firing asynchronously, with great variability. Thus, heterogeneously active elements converge to a homogeneously invariant average pattern. Preliminary results indicate that the changes caused by conditioning are distributed across the subgroups, causing closely similar changes in the average activity of each set of elements (*34*).

These results support the contention that the response to a single stimulus presentation, averaged across many elements of an ensemble, would possess the same stimulus-specific invariance observed in the response of a single element averaged across many presentations of the stimulus.

Our studies show little evidence of cells being selective to one but not the other stimulus. In general, cells or ensembles responsive to one CS also responded to the other. Differences between stimuli were reflected in temporal firing patterns. These results show that temporal differences in statistically orderly firing patterns are caused in extensive neural ensembles by different stimuli, and could serve to represent sufficient information to discriminate between the stimuli in many regions. The data do not seem to provide support for the proposal that neurons with different connectivity mediate sensory



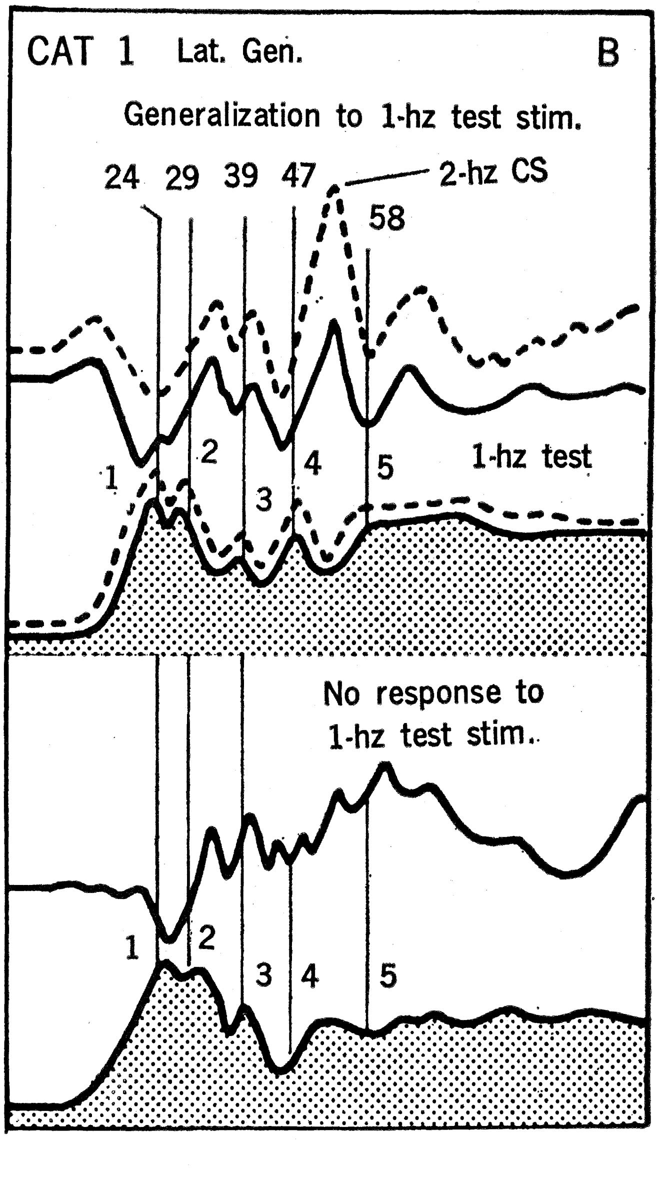


Fig. 6. (A) Computations of average responses obtained from the lateral geniculate nucleus and nucleus reticularis of the cat under various conditions during the same experimental session. First row of averages (a) are based upon 100, second (b) and third (c) rows are based upon 42 repetitions of the same stimulus applied during a number of behavioral trials. Analysis epoch was 90 msec. (a) Average responses evoked in structures by the 10-hz CS (flicker) actually used in training, during repeated correct behavioral performances. (b) Average responses evoked by a novel 7.7-hz CS, during repeated generalization behavior. Test trials with the 7.7-hz stimulus were interspersed among trials with the actual 10-hz CS, and were never reinforced. (c) Average responses evoked by the 7.7-hz flicker on presentations when no generalization behavior was elicited. The waveshape elicited by the actual CS is similar to the response evoked by the novel stimulus during generalization behavior. Notice the absence of the second positive component in the evoked potential when generalization behavior failed to occur. [Data from D. S. Ruchkin and E. R. John (*40*)] (B) (Top) Records of AER's and PHS's obtained during 18 trials that resulted in CR to the 2-hz CS (dotted curves) and during 32 trials that resulted in behavioral generalization in response to a 1-hz flicker used as a test stimulus (solid curves). The test stimuli were randomly interspersed between presentations of 2-hz (dotted curves) and 8-hz flickers in a long experimental session. (Bottom) Records of AER's and PSH's obtained during 17 trials that resulted in failure to elicit generalization behavior in response to the test stimulus. Note change in late components. Analysis epoch, 100 msec. [Data from E. R. John and P. P Morgades (*32*)]

discriminations in a logically exclusive way, by firing which is restricted to one or another pathway reserved for the representation of particular events (*6*).

Using only evoked potentials, we have shown that different visual forms equated for energy also cause characteristically different response patterns (*35*). These findings have been confirmed in several other laboratories (*36*). Therefore, the above results may well be related to general properties of the representation of information in the central nervous system rather than constituting a special case relevant only to flicker stimuli.

Functional significance of temporal patterns of activity. The data so far discussed demonstrate that widespread and marked changes in the electrophysiological responses to CS's occur during learning and show that, after learning, a variety of anatomical regions display similar evoked responses to the CS, correlated with strikingly invariant temporal patterns of firing in extensively distributed ensembles of neurons. These findings, however intriguing they might seem, are not sufficient by themselves to justify the conclusion that this diffuse network of characteristically responding cells constitutes a representational system which functions to store information about the CS. In spite of control data ruling out the gross effects of pseudo-conditioning, the observed changes might be caused by changes of attention or other unspecific factors altering the sensory response to the CS. Numerous experiments have been conducted to explore such alternate interpretations. The resulting evidence strongly supports the interpretation that the phenomena under discussion are closely related to the storage and retrieval of information about the stimulus.

Storage and Retrieval of Information

Similarity of responses across structures diminishes in behavioral errors. During correct discrimination behavior, similar EEG rhythms, similar evoked potential wave-shapes, and similar ensemble firing patterns appear in some (but not all) brain regions. However, during trials resulting in erroneous performance, certain regions in the "representational system" fail to display the characteristic response appropriate to the actual stimulus and produce an inappropriate response, corresponding to the characteristic effect of the signal for the behavior which is erroneously performed (*37*). A clear example of this phenomenon is shown in the EEG tracings of Fig. 5.

When a trained animal performs a CR to a novel stimulus, indicating behavioral generalization, some brain regions display EEG rhythms corresponding to the frequency of the CS actually used in training rather than to the frequency of the novel stimulus which elicits the behavior. The best demonstration that these effects reflect the release of a stored electrical pattern representing the CS, rather than any unspecific influence, has been provided by Majkowski (*38*), who showed that they appear only on the trained but not on the untrained side of cats in which the two cerebral hemispheres have been separated by surgical procedures. That the frequency of these rhythms released during generalization specifically corresponds to the absent CS has been shown in a differential design, by means of spectral analysis (*39*).

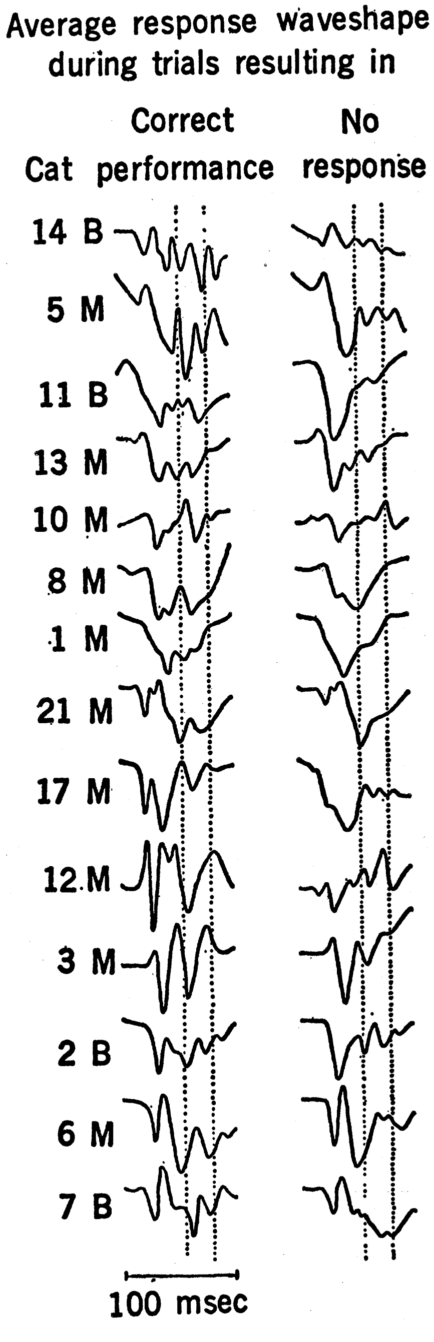


Fig. 7. Similar features of waveshape components in the average evoked responses from the visual cortex of 14 cats. Each row of data comes from a different cat. The waveshapes in the left column were computed from cats performing a correct CR during a group of trials with the same CS. The waveshapes on the right show the changes observed when the same CS failed to elicit behavioral response in the same session (NR). The vertical dotted lines bracket a negative (downward-going) component with latency between 55 and 80 msec, present during CR and smaller or absent during NR. Since the signal frequency of the CR varied widely in these animals, the inter-stimulus analysis epochs are necessarily different. All averages are based upon 200 repetitions of the CS. M, monopolar; B, bipolar.

Results analogous to those just described have been reported for the shape of the evoked potential (*40*) and the ensemble firing pattern (*31, 32*) in behavioral generalization and are illustrated in Fig. 6, A and B.

Effects of brain stimulation. The functional relevance of these widespread temporal patterns is further suggested by brain stimulation experiments. Electrical stimulation of the mesencephalic reticular formation (MRF) at either of the two different frequencies used for previous peripheral discrimination training results in almost immediate performance of the appropriate differentiated behavior (*37, 41*) and rapid achievement of criterion in training. Subsequent stimulation of MRF at either cue frequency simultaneously with presentation of the peripheral CS (flicker or click) at the other frequency results in performance of the behavior appropriate to the electrical stimulus (*41*).

It is obviously impossible to activate stimulus-specific pathways or connections by such gross electrical stimulation of neuronal masses. The relative ease of transfer from peripheral to MRF stimulation and the ability of MRF input to control behavior differentially even though contradicted by a simultaneous sensory CS suggest that the critical aspect of the representation of stored information was the temporal coherence of activity patterns in neuronal ensembles which were not stimulus-specific, rather than the activity of a cell or set of cells in some CS-specific pathway.

Release of the "readout" component. As stated earlier, new late components appear in the responses evoked by the CS during learning. As Fig. 6 shows, these late components appear if behavioral generalization does occur in response to a novel stimulus, but they fail to appear if the novel stimulus does not elicit behavioral generalization. We have referred (*26*) to these late components as endogenous, released rather than determined by the afferent stimulus, while earlier components were termed exogenous, reflecting afferent input information about the stimulus. By subtracting waveshapes obtained when a stimulus failed to elicit behavioral generalization ("no go"—exogenous input alone) from waveshapes obtained when generalization occurred ("go"—exogenous input plus endogenous readout), a picture of the difference waveshapes representing the "readout process" was constructed (*26*).

The difference waveshape was similar in a wide variety of structures, including the sensory cortex and the MRF, and seemed to propagate centrifugally to the lateral geniculate body, where it appeared latest. Thus, the evidence from these initial studies of behavioral generalization showed that an extensively distributed system could fire autonomously in response to a novel stimulus, releasing patterns of response resembling those usually caused by the CS. The released activity was widespread and seemed to involve very large numbers of neurons.

The consistency of the readout process released in the visual cortex by a visual CS is shown in Fig. 7, for 14 differentially trained cats. A marked difference in the surface-negative (downward) wave with latency between 55 and 80 milliseconds is apparent when these "go" and "no go" outcomes are compared.

Specificity of readout processes. Differences between "go" and "no go" might reflect unspecific factors such as attention, arousal, or level of motivation. In order to conclude that these released patterns reflected activation of specific representational systems, it was necessary to show different patterns when the same signal elicited two different "go" outcomes. For this purpose, we utilized a technique which we call differential behavioral generalization (*42*).

In this method, animals are trained to discriminate between two different flicker frequencies, V_1 and V_2, by performing two different behavioral responses, CR_1 and CR_2. After substantial overtraining, a third test stimulus, V_3, midway between the two differential stimuli is occasionally presented. The animal sometimes treats V_3 as equivalent to V_1, performing CR_1, and sometimes reacts by performing CR_2 as if V_2 had been presented. The electrophysiological responses evoked by V_3 are analyzed separately, according to the behavioral outcome.

A wide range of frequencies between 1 and 12 hz have been used for V_1 and V_2 in the 16 cats for which differential generalization data have been obtained, and basically similar results have been observed in every case. Typical data from the lateral geniculate body (bipolar recording) of one cat are illustrated in Fig. 8. Two markedly different waveshapes were elicited by the same physical stimulus, V_3, depending upon the particular behavior which was subsequently performed. These results show unequivocally that what we will refer to as the "readout waveshape" is a mode of the evoked response which is independent of the novel afferent stimulus. This readout waveshape or mode is released from memory and is a close facsimile of the effects of the familiar CS which is the usual specific cue for the generalized behavior.

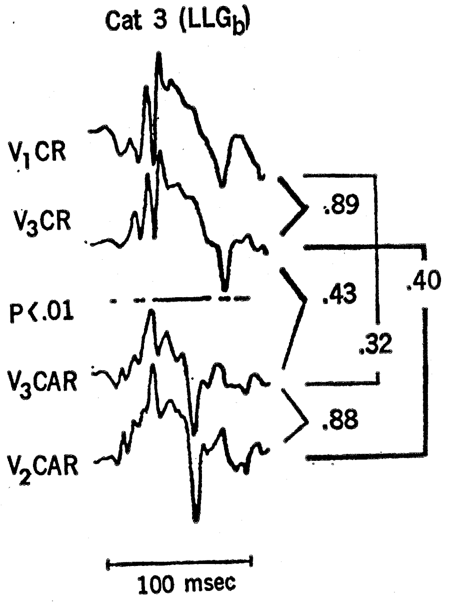


Fig. 8. Waveshapes of average responses recorded from the lateral geniculate body (bipolar) under various stimulus-response contingencies: V_1CR, during trials resulting in correct performance of an approach response (*CR*) to a 3.1-hz flicker CS; V_2CAR, during trials in which a conditioned avoidance response (*CAR*) was correctly performed in response to a 7.7-hz CS; V_3CR, during generalization trials in which a neutral 5-hz test stimulus elicited *CR* behavior; V_3CAR, during generalization trials in which the same 5-hz test stimulus elicited *CAR* behavior. The interrupted line between V_3CR and V_3CAR indicates time intervals during which V_3CR and V_3CAR were significantly different at better than the $P = .01$ level. The numbers at the right indicate the correlation coefficients between the corresponding bracketed waveshapes.

Further control experiments were directed at other possible unspecific origins of the observed differences between readout waveshapes (*43*): (i) Animals were taught to respond to one auditory stimulus (A_1) with behavior CR_1 and to a second auditory stimulus (A_2) with behavior CR_2. These differential auditory CS's consisted of different steady tones (600 as opposed to 1200 hz) or clicks at different repetition rates. Subsequently, almost identical average responses were elicited by the meaningless flicker, V_3, whether it was yoked to A_1 during CR_1 performance or to A_2 during CR_2 performance. After transfer of training so that the flicker acquired cue value, V_3CR_1 responses became distinctly different from V_3CR_2 responses. This finding rules out the possibility that the potentials evoked by a sensory stimulus alter when a particular behavioral response bias or set becomes established. The finding also shows that the differences in readout waveshapes obtained during trials that result in two different behavioral outcomes do not merely reflect the different movements, body positions, or orientations assumed by the animal. Differences were observed only when the stimulus genre had real cue value. (ii) Animals were trained to perform an approach behavior CR_1 in response to V_1 on either the left or right side of the apparatus, and to perform a different behavior CR_2 to avoid shock in response to V_2 on either the left or right side of the apparatus. It was found that the evoked response to V_1 was the same no matter on which side of the apparatus the animal performed. Similarly, evoked responses to V_2 were the same whether the avoidance response was performed on the left or right side. However, responses evoked by the same CS during a correct performance were radically different from those evoked when the performance was in error, although both responses were performed on the same side of the apparatus. These results show that the differences in readout waveshapes are not attributable to the fact that different sets of movements are being carried out but seem to depend upon the significance of the behavior, that is, the specific meaning of the signal. In groups of other animals, response-specific readout waveshapes have been obtained by means of auditory stimuli or by using direct electrical stimulation of brain structures as the CS. The results also show that readout processes are not restricted to visual signals, nor can differences in readout processes be attributed to changes in position, orientation, or direction of gaze. They do not reflect pupillary dilation, since visually elicited readout waveshapes are not markedly altered by homatropine. They do not merely reflect feedback from performance of an instrumental movement, since readout waveshapes obtained in response to cues of one sensory modality usually do not appear in the same brain regions when the same CR's are performed for cues in other sensory modalities. It is interesting that in a few animals, markedly similar waveshapes have been elicited in certain structures by cues of different modalities, with latency differences appropriate to the different afferent transmission times. (iii) In one group of animals, CR_1 and CR_2 were two different approach responses, each one appropriate to a specific CS. In a second group, CR_1 and CR_2 were two different avoidance behaviors. In a third group, CR_1 was an approach behavior and CR_2 was an avoidance behavior. There were two tasks, lever pressing and hurdle jumping, and observational as well as instrumental learning situations were used. In all instances, if the same physical stimulus had sometimes elicited performance of behavior appropriate to one cue but sometimes performance of a behavior appropriate to a different cue, it was possible to demonstrate differences in the waveshape of the evoked potential. Thus, these differences do not depend upon differences in motivation for the two differential behaviors nor are they unique to a particular motivation. Furthermore, they are not restricted to a particular kind of learning or behavioral task.

Temporal sequence of readout waveshapes. It is particularly informative to study the temporal sequence of waveshapes of potentials evoked during trials with different behavioral outcomes, as illustrated in Fig. 9. These data provide an example of the way in which readout potentials emerge during differential generalization. In many instances, evoked potentials shift from one readout waveshape, or mode, to

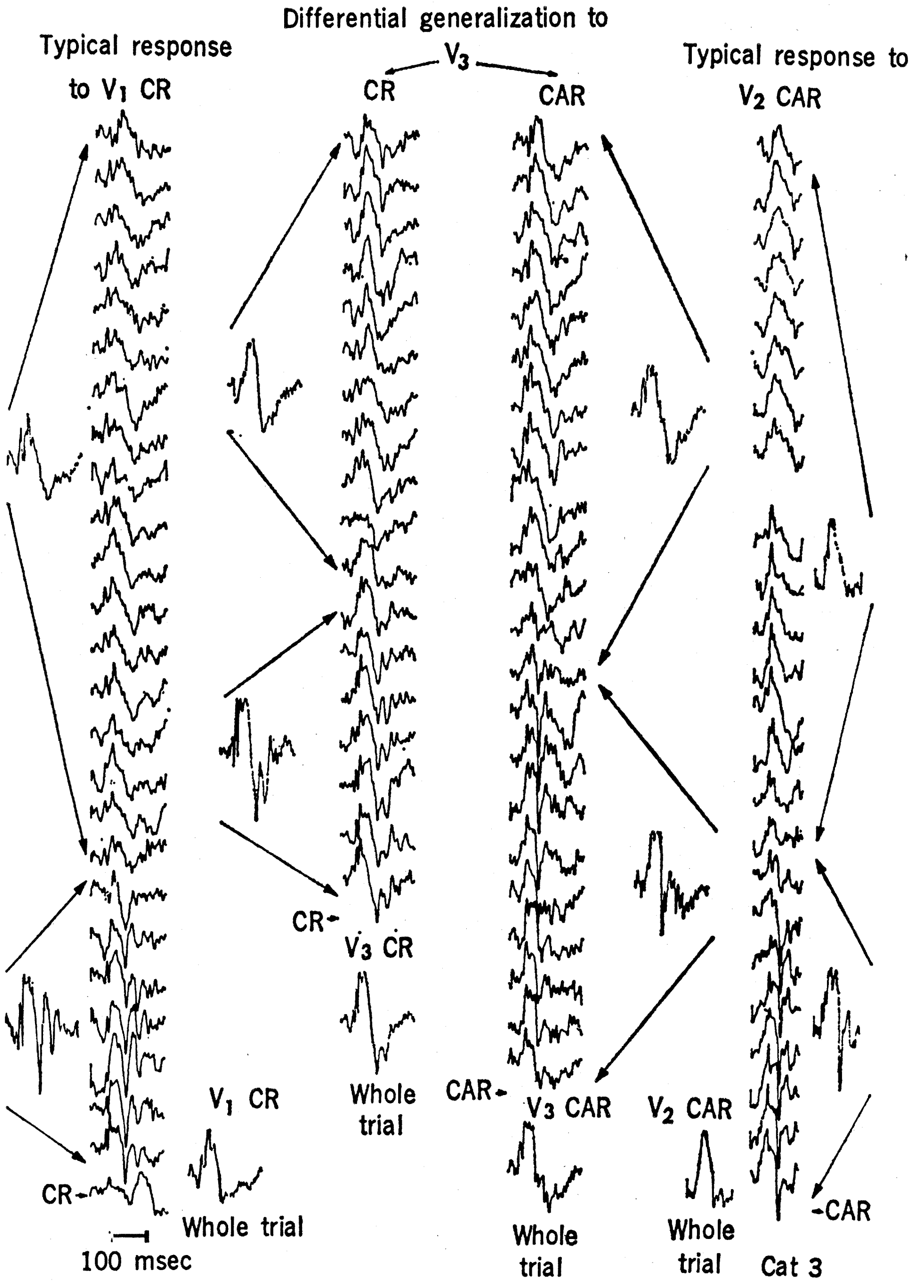


Fig. 9. Each column contains the series of evoked potentials which occurred in a behavioral trial, arranged from CS onset (top of column) to behavioral performance (bottom of column). Sequential (bipolar) potentials evoked from the lateral geniculate during a V_1CR trial are displayed in the first column, during a V_3CR trial in the second column, a V_3CAR trial in the third column, and a V_2CAR trial in the fourth column. Each line represents the sweep of the oscilloscope triggered by occurrence of a single flicker CS, at the respective repetition rates. The waveshapes shown beside or below each column represent average evoked responses computed over the whole behavioral trial or across the subset of evoked potentials indicated by the brackets. In the V_1CR trial, a radical change in the waveshape of the evoked potential (EP) can be seen on the 19th EP, indicated by the upper arrow of the lower bracket. The sharp, new, single, downward deflection assumes a double shape on the 20th EP; it is clearly evident on the 21st EP and remains visible until the end of the trial. In the trial shown in the fourth column, the first eight EP's have approximately similar shape. A sharp, short upward deflection caused by a late component appears in the 9th to 16th EP. On the 17th EP, at the upper arrow of the lower bracket, a downward deflection appears, which becomes quite sharp on the 18th EP and remains as a single sharp downward deflection until the end of the trial. Clear differences are evident in the EP's of the second and third columns. On the 12th EP, marked by the upper arrow of the lower bracket, the double downward deflection characteristic of the last potentials in the V_1CR trial appears. Its precursors can be seen in the 4th and 8th EP's. The remaining EP's of the V_3CR trial display this double downward deflection. The first eleven EP's of the V_3CAR contain a few single or double downward deflections at about the latency of the downward deflection seen in the 19th to 25th EP's of the V_1CR trial. However, in the 12th EP, a sharp, short downward deflection appears, and the remainder of the EP's in this trial closely resemble the final EP's of the V_2CAR trial.

another and back again at such a high rate that it is difficult to give credence that they correspond to some overt behavior as yet unperceived by the experimenter. As shown in Fig. 9, differences between V_3CR_1 and V_3CR_2 were often obscured by averaging the results for a whole trial. The differences could be accentuated by the experimenter selecting typical readout waveshapes from the later portions of behavioral trials (*42*). It was suggested that such selection by the experimenter made it possible for subjective bias to invalidate the apparent significance of these findings (*44*). Accordingly, a computer sorting program was devised to perform this analysis in an objective way, yielding homogeneous subgroups of evoked potentials with minimum variance from the heterogeneous population of waveshapes which characterize behavioral situations (*45*). By means of this program, the waveshapes objectively selected by the computer closely approximated the waveshapes subjectively selected on the basis of experimenter judgment (*46*). The waveshape formed by the evoked potentials seems to be an excellent predictor of the subsequent behavioral response. A film showing oscilloscopic traces of potentials evoked during differential generalization has been prepared. Numerous audiences have quickly learned to predict the results of trials by studying the waveshapes (*47*).

Nevertheless, it still had to be demonstrated objectively that readout waveshapes thus selected in any animal are generally distributed throughout a large number of behavioral trials in which the same outcome occurs, and that different readout waveshapes are significantly predictive of different behaviors.

It has been possible to show the extremely reproducible content of waveshapes or modes of evoked potentials identified by computer selection in a large number of behavioral trials with the same outcome, to establish the correlation of a specific readout waveshape or mode with a particular behavior, and to find such correlations in a large number of animals.

Figure 10 shows the distribution of three different modes of evoked response in six groups of different types of behavioral trials from the same animal. The cat from which these 59 trials were recorded was trained to press the left bar on a work panel to avoid electric shock from a floor grid within 15 seconds after the onset of a 4.0-hz flicker (V_1) and to avoid shock by pressing

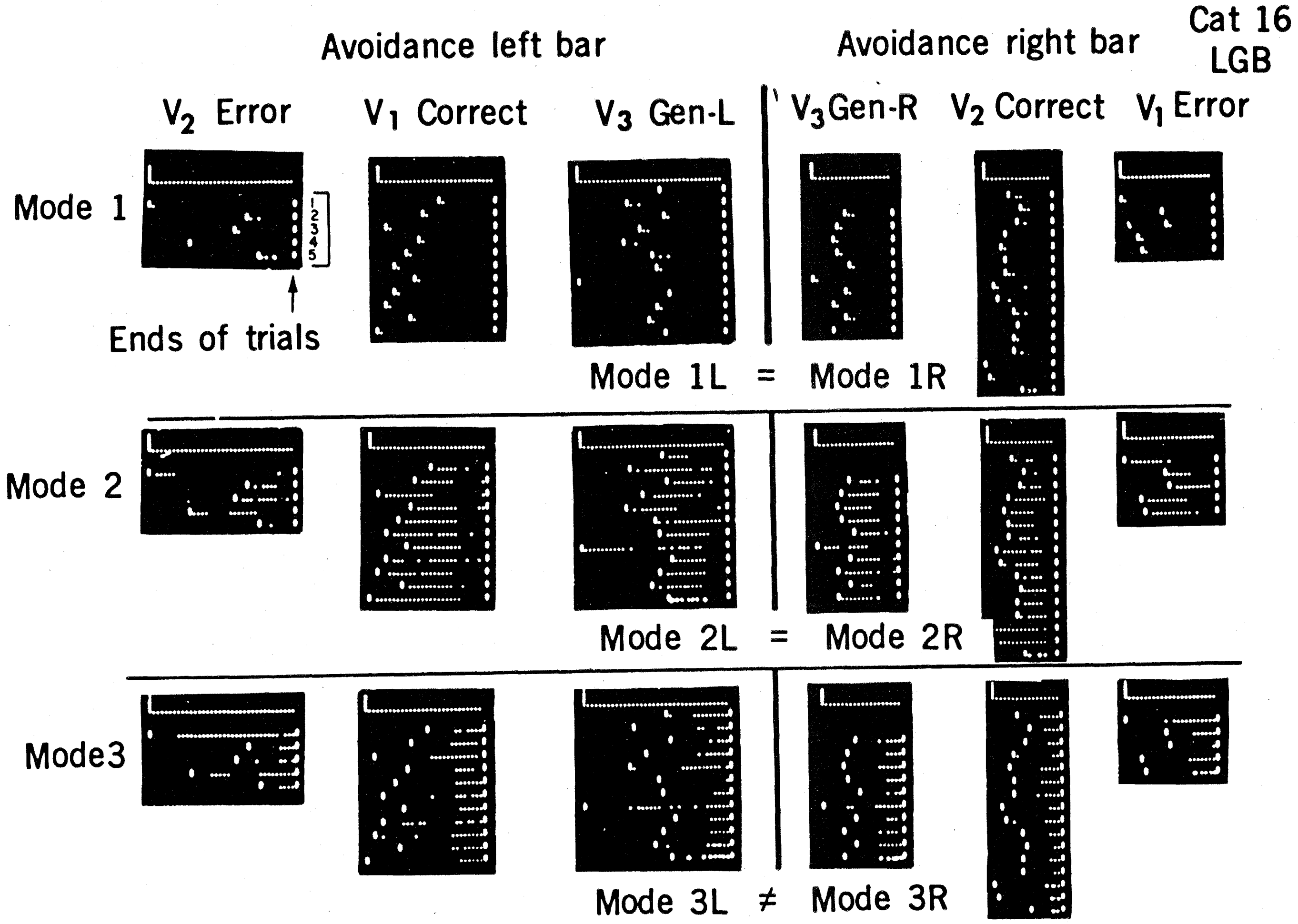
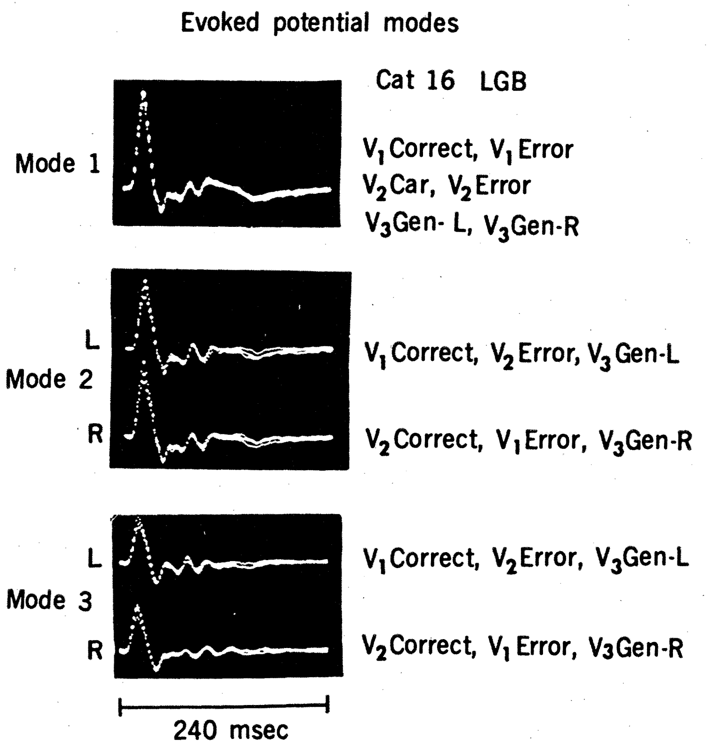


Fig. 10. Distribution of modes of evoked potentials. Each column represents a set of trials in which the same behavioral response was performed to a particular stimulus, as indicated in the heading. The V_1 (4.0-hz flicker) was the CS for pressing the left bar; V_2 (2.0-hz flicker) was the CS for pressing the right bar, and V_3 (3.0-hz flicker) was a test stimulus used to elicit differential generalization. The 59 trials included five errors in response to V_2, 11 correct responses to V_1, and 12 generalization-L (Gen-L) responses to V_3, all resulting in left bar responses; there were also 10 generalization-R (Gen-R) responses to V_3, 16 correct responses to V_2, and five errors in response to V_1, all resulting in right bar responses. Both left and right bar responses were to avoid electric shock from a floor grid. The bars were mounted side by side on a work panel. The top row of white dots in each black rectangle represents the occurrence of successive CS (flicker) in the behavioral trials. Each row of dots corresponds to a separate trial, beginning at the left vertical bar and ending at the right bar. The occurrence of a dot indicates that the potential evoked for that CS was classified as an example of the indicated mode of evoked response. Three modes were identified within all six types of trials. and are represented by the three horizontal arrays of data which are analyzed in Fig. 11.

Table 1. Distribution of CR_1 and CR_2 modes of evoked potentials within sets of trials resulting in either CR_1 or CR_2 behavior in response to a single flicker signal (usually V_3). Number of trials and evoked potentials in each set is indicated for each animal, as is the structure from which the data were recorded. The exact probability of obtaining the observed distributions from a randomly distributed population was calculated and is shown in the last column. LG, lateral geniculate body; VIS, visual cortex; MRF, mesencephalic reticular formation; M, monopolar derivation; B, bipolar derivation; + + represents approach-approach discrimination; − − represents avoidance-avoidance discrimination; + − represents approach-avoidance discrimination; 0 + represents no response as opposed to generalization of approach discrimination.

Cat	Structure	Number of trials with outcome		Number of evoked potentials from trials with outcome		Evoked potentials (percent)				$P <$
						Classified as CR_1 mode from trials with outcome		Classified as CR_2 mode from trials with outcome		
		CR_1	CR_2	CR_1	CR_2	CR_1	CR_2	CR_1	CR_2	
2	LG_B + −	7	10	95	159	0.48	0.16	0.22	0.54	10^{-8}
3	VIS_B + −	9	10	118	137	0.37	0.17	0.27	0.45	10^{-4}
4	LG_M 0 +	5	5	131	125	0.87	0.10	0.13	0.90	10^{-37}
5	LG_B + −	9	4	174	81	0.58	0.21	0.03	0.41	10^{-14}
7	LG_B + −	7	8	97	120	0.55	0.30	0.45	0.69	10^{-3}
8	VIS_B + −	3	5	56	104	0.45	0.14	0.27	0.62	10^{-5}
9	LG_B + +	5	5	128	128	0.58	0.34	0.27	0.50	10^{-4}
10	LG_B + −	7	7	128	128	0.47	0.14	0.00	0.41	10^{-20}
11	VIS_M + −	4	3	154	102	0.37	0.15	0.23	0.36	10^{-3}
12	MRF_M + −	9	9	101	149	0.77	0.50	0.23	0.50	10^{-4}
13	LG_M + −	5	5	104	150	0.57	0.04	0.10	0.31	10^{-16}
14	LG_M + −	3	4	79	177	0.73	0.24	0.27	0.76	10^{-12}
16	LG_B − −	11	11	63	55	0.70	0.31	0.13	0.53	10^{-5}
21	LG_B − −	6	6	128	128	0.79	0.46	0.18	0.48	10^{-7}



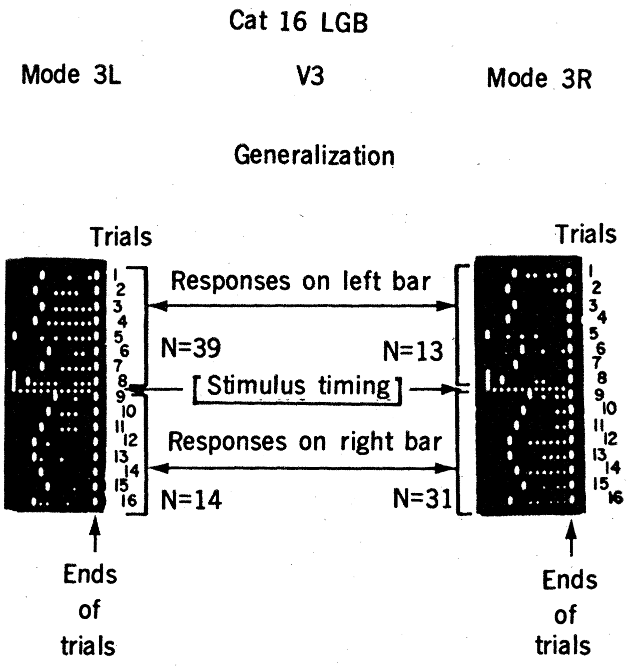


Fig. 11 (left). Mode 1 consisted of the initial evoked potential in 46 of the 59 trials, plus a small number of additional evoked potentials all of which occurred early in the trials. The mode 1 potentials from the six types of trials are superimposed in the top row of data. They were essentially identical. Mode 2 consisted of the evoked potentials which dominated the early portions of all types of trials. The second row of data, mode 2 left, shows the superimposed waveshapes of this mode for the three sets of trials which resulted in the cat pressing the left bar, while the third row, mode 2 right, illustrates the corresponding data for responses in which the right bar was pressed. Mode 3 consisted of the evoked potentials which dominated the final portions of all types of trials. The fourth row of data, mode 3 left, shows the superimposed waveshapes of this mode for the different sets of responses in which the left bar was pressed, while the corresponding data for responses in which the right bar was pressed are illustrated in the bottom row. The three different mode 3 left waveshapes are closely similar, independent of whether they were elicited by V_1, V_2, or V_3. Similarly, the three mode 3 right waveshapes are essentially identical. However, mode 3 left waveshapes are markedly different from mode 3 right. All data were recorded from the lateral geniculate nucleus, bipolar derivation. Fig. 12 (right). Correlation between modes of evoked response and behavior. Each rectangle contains data for eight V_3 trials which resulted in left bar behavioral generalization (trials 1–8) and eight V_3 trials which resulted in right bar behavioral generalization (trials 9–16). The rectangle on the left shows the incidence of mode 3 left waveshapes in 16 generalization trials, while the right rectangle shows the incidence of mode 3 right waveshapes in the same trials. The trials were randomly selected from those shown in Fig. 1, and were truncated so that only the late portion dominated by mode 3 waveshapes was subjected to this further analysis. These results showed a significant deviation from randomness in the probability that a particular waveshape mode would occur in a trial with a particular behavioral outcome ($P = .00001$).

the right bar after a 2.0-hz flicker (V_2). After overtraining, a number of presentations of a 3.0-hz (neutral) flicker (V_3). were randomly interspersed among a sequence of V_1 and V_2 trials. The set of trials analyzed here includes left and right bar responses to each of these three signals. All trials began with an evoked potential of mode 1, progressed to a period dominated by potentials of mode 2, and concluded with a period in which mode 3 potentials were most probable.

Mode 1 waveshapes from the six types of trials were essentially identical, indicating that the state of the nervous system of this animal was approximately the same at the onset of trials of each behavioral type. The mode 2 left waveshapes from trials resulting in the left bar being pressed in response to the three different stimuli were closely similar to each other, and also closely resembled mode 2 right waveshapes from trials in which the right bar was pressed in response to the three kinds of stimuli. Mode 3 left waveshapes were closely similar independent of the stimulus, as were mode 3 right waveshapes (see Fig. 11). However, mode 3 left and mode 3 right waveshapes were significantly different from each other. Although the differences between mode 3 left and mode 3 right seem slight, they can readily be identified by the computer. Differences between readout waveshapes vary from cat to cat. In some animals they are extremely striking and obvious, while in others they may be as slight as those illustrated in Fig. 11. However, relatively late components are usually involved.

The use of the mode 3 left and mode 3 right waveshapes as classification criteria for sorting showed clearly that the probability of these two kinds of waveshapes varied with the behavioral outcome of the trial. Mode 3 left waveshapes occurred much more frequently in V_3-left trials while mode 3 right waveshapes occurred more frequently in V_3-right trials. This distribution deviated significantly from random, as shown in Fig. 12. Table 1 shows comparable results for 14 cats. For each animal, a mixed set of trials was constituted from a large sample of trials which resulted in two different behavioral outcomes to the same physical stimulus. The evoked potentials recorded during the whole set of trials were then subjected to sorting analysis. Each mixed population contained two different readout modes, one correlated with each of the differential behaviors. The differential distribution of the two readout modes in the two kinds of behavioral trials was highly significant. The temporal distribution and actual waveshapes of these various readout processes will be presented in a forthcoming paper (*43*).

These results show that readout processes can be objectively defined and automatically identified, that they appear consistently in an extensive sample of behavioral trials from a given animal, that they can be found in most, if not all, animals in an experimental population, and that they are strongly correlated with a particular differential behavior.

The evidence which has been presented indicates that when a specific memory is retrieved, a temporal pattern of electrical activity peculiar to that memory is released in numerous regions of the brain. To that released set of waveshapes corresponds the average firing pattern of ensembles of neurons diffusely distributed throughout these widespread anatomical domains. Individual neurons within these ensembles display different momentary discharge patterns but the individual average firing patterns converge to the ensemble mean. This suggests that during retrieval of a particular memory, a unique and invariant temporal pattern of coherence occurs in the neural discharges averaged across a spatially distributed and diffuse ensemble of neurons, in which the variable activity of any individual neuron is significant primarily insofar as it contributes to the statistics of the population.

Remembering

If this description of the process of retrieval of a memory is accurate, then it is also a description of the subjective experience of remembering. We selected the problem of memory retrieval as of critical importance because it afforded an objective way to define the momentary content of consciousness in an animal, if one were willing to relax the conventional anthropocentric restriction of subjective experience to human beings. A substantial body of data from human beings not only supports the contention that these released patterns of electrical activity actually correspond to the activation of specific memories, but establishes unequivocally that there is a subjective correlate to the appearance of these released potentials. When an event expected by a man does not occur, a cerebral potential appears at a latency similar to that of potentials usually evoked by the expected stimulus. Evoked potentials elicited in man by absent but expected events have been reported by numerous workers (*48*), including, most recently, Weinberg *et al.* (*49*). We earlier reported similar findings in the cat (*50*). These cerebral events, termed "emitted potentials," have been interpreted by Weinberg *et al.* to reflect the generation of processes corresponding to the memory of past or imaginary stimuli. Working in our laboratory, Herrington and Schneidau (*51*) demonstrated that in some subjects the shape of the waveshape released when a particular geometric form was imagined closely resembled the waveshape evoked by actual presentation of that visual stimulus.

Thus, there is a close correspondence between the data for animals reported earlier in this article and findings with human subjects. This correspondence permits us to conclude that there is a subjective experience, "remembering" or "thinking," which is the concomitant of the release of the electrical waveshape representing a specific memory. Presumably, fundamentally similar physiological processes mediate the release of these waveshapes in man and in the cat. It is intriguing to speculate, in view of these similarities, whether subjective experience in man should be attributed to the temporal patterns of coherence in the enormous populations of neurons which make up our brains. This speculation suggests that statistical or cooperative processes involving large numbers of neural elements generate physical consequences and concomitants not to be predicted from the characteristics of the individual elements. The continuity and stability of content of human consciousness may arise from invariant features of the patterned fluctuations in coherence in statistical ensembles.

Conclusions

It has been demonstrated that part of the electrical activity which follows the presentation of a stimulus to a trained cat during differential generalization is independent of the afferent stimulus. This released activity is considered to reflect activation of a specific memory, and has been referred to as the "readout component."

Readout components have been observed in evoked potentials recorded from a wide variety of electrode placements, appearing in about two-thirds of the derivations studied in 25 trained cats. The wide anatomical distribution of these events and the great amplitude of the potentials displayed in certain structures indicate that readout activity involves large ensembles of neurons in different brain regions in cooperative processes, and that a rather high coherence of neural discharge is probably achieved in these extensive neuronal populations. The implications of these findings are corroborated by studies with permanently implanted moving microelectrodes. Although single cells within a region display highly variable responses to CS's, ensembles of neurons display essentially invariant temporal patterns. Similar firing patterns were observed in neural ensembles at many different positions along the electrode traverse. Differences between firing patterns elicited by two differentiated stimuli at a fixed position were greater than the differences within the set of patterns elicited by one CS across the mapped domain. The release of readout processes was observed in ensemble firing patterns as well as in evoked potentials.

Readout components have been observed in response to visual, auditory, and direct electrical conditioned stimuli. They have been observed in a variety of different tasks performed for different motivations. They are not elicited by meaningless stimuli yoked to conditioned cues, but appear when those stimuli become meaningful after transfer of training. Control studies indicate that they are unrelated to changes in attention, arousal, body position or orientation, direction of gaze, pupillary dilation, set or response bias. It seems reasonable to interpret this phenomenon as reflecting the activation and release of a specific memory.

The experimental results reported here strongly support the proposition that the common mode of activity of large numbers of neurons in anatomically extensive systems represents information about a learned experience. Different brain regions come to share a common mode during learning. When an experienced organism receives a novel and meaningless stimulus and generalization occurs, this new afferent input in a familiar context activates the representational system in such a way as to cause release of a common mode of activity like that stored during the learning experience. This released activity is not attributable to any of the numerous unspecific factors thus far investigated, and seems to reflect the memory of that experience. The relationship between the temporal distribution of release of a specific waveshape and subsequent occurrence of the appropriate behavior was found to be highly nonrandom.

These observations are compatible with the salient features of the statistical configuration theory which has been proposed. Moreover, they are difficult to reconcile with switchboard and place theories of learning, which thus far lack comparable experimental evidence.

Analogous data from experiments with human beings indicate that phase-locked potentials are released at the time that absent stimuli are expected to occur. Furthermore, the waveshapes released when particular visual stimuli are imagined resemble the waveshapes of potentials evoked by actual presentation of the imagined stimuli. These findings suggest that the subjective experience of remembering is correlated with the release of readout processes, and provide some basis for the speculation that coherent temporal patterns in the average activity of anatomically extensive neural ensembles may constitute the neurophysiological basis of subjective experience.

References and Notes

1. E. Tanzi, *Rev. Sper. Freniat.* **19**, 419 (1893).

→ K. S. Lashley, *Psychol. Rev.* **31**, 369 (1924); *Science* **73**, 425 (1931); *Symp. Soc. Exp. Biol.* **4**, 454 (1950).

3. D. O. Hebb, *Organization of Behavior: A Neuropsychological Theory* (Wiley, New York, 1949); J. Konorski, *Conditioned Reflexes and Neuron Organization* (Cambridge Univ. Press, Cambridge, 1948).
4. N. E. Miller, in *Symposium on Motivation*, M. R. Jones, Ed. (Univ. of Nebraska Press, Lincoln, 1963), p. 96.
5. E. A. Asratyan, in *Brain Mechanisms and Learning*, J. F. Delafresnaye, A. Fessard, R. W. Gerard, J. Konorski, Eds. (Blackwell, Oxford, 1961), p. 95; S. H. Barondes, in *Molecular Approaches to Biology and Memory*, W. L. Byrne, Ed. (Academic Press, New York, 1970), p. 27; J. A. Deutsch, *Annu. Rev. Psychol.* **20**, 85 (1969); J. C. Eccles, in *Brain and Conscious Experience*, J. C. Eccles, Ed. (Springer-Verlag, New York, 1966); A. Fessard and T. Szabo, in *Brain Mechanisms and Learning*, J. F. Delafresnaye, A. Fessard, R. W. Gerard, J. Konorski, Eds. (Blackwell, Oxford, 1961), p. 353; H. Gastaut, in *Neurological Basis of Behavior*, G. E. W. Wolstenholme and C. M. O'Connor, Eds. (Churchill, London, 1958); S. F. Grossman, in *Molecular Mechanisms in Memory and Learning*, G. Ungar, Ed. (Plenum, New York, 1970), p. 249; W. C. Halstead and W. B. Rucker, in *Molecular Approaches to Biology and Memory*, W. L. Byrne, Ed. (Academic Press, New York, 1970), p. 1; D. O. Hebb in *Brain Mechanisms and Learning*, J. F. Delafresnaye, A. Fessard, R. W. Gerard, J. Konorski, Eds. (Blackwell, Oxford, 1961), p. 37; H. H. Jasper and B. Doane, in *Progress in Physiological Psychology*, E. Stellar and J. M. Sprague, Eds. (Academic Press, New York, 1968), p. 79; J. Konorski, *Integrative Activity of the Brain* (Univ. of Chicago Press, Chicago, 1967); E. Kandel and W. A. Spencer, *Physiol. Rev.* **48**, 65 (1968); A. K. McIntyre, *Proc. Int. Physiol. Congr.* 19th Montreal (1953), p. 107; N. E. Miller, in *The Neurosciences: A Study Program*, G. C. Quarton, T. C. Melnechuk, F. O. Schmitt, Eds. (Rockefeller Univ. Press, New York, 1967), p. 6 → P. M. Milner, *Psychol. Rev.* **64**, 242 (1957); F. Rosenblatt, in *Molecular Mechanisms in Memory and Learning*, G. Ungar, Ed. (Plenum, New York, 1970), p. 103; R. W. Russel, *Brain, Memory, Learning* (Clarendon, Oxford, 1959); R. W. Sperry, in *Biological and Biochemical Bases of Behavior*, H. Harlow and C. N. Woolsey, Eds. (Univ. of Wisconsin Press, Madison, 1958), p. 401; W. H. Thorpe, *Learning and Instinct in Animals* (Methuen, London, 1956); J. F. Tonnies, *Arch. Psychiat. Nervenkr.* **182**, 478 (19 → J. Z. Young, *Proc. Roy. Soc. Ser. B* **139**, 18 (1951); H. von Foerster, in *The Future of the Brain Sciences*, S. Bogoch, Ed. (Plenum, New York, 1969), p. 19; J. Z. Young, *A Model of the Brain* (Clarendon, Oxford, 1964).
6. G. Ungar, in *Molecular Mechanisms in Memory and Learning*, G. Ungar, Ed. (Plenum, New York, 1970), p. 149.
7. R. Galambos, in *The Neurosciences: A Study Program*, G. C. Quarton, T. C. Melnechuk, F. O. Schmitt, Eds. (Rockefeller Univ. Press, New York, 1967), p. 637.
8. A. L. Bandura, *Symposium on Motivation*, M. R. Jones, Ed. (Univ. of Nebraska Press, Lincoln, 1963), p. 211; P. Chester, *Science* **166**, 901 (1969); J. A. Corson, *Psychon. Sci.* **7**, 197 (1967); C. L. Darby and A. J. Riopelle, *J. Comp. Physiol. Psychol.* **52**, 94 (1959); R. Gilbert and M. Beaton, *Psychon. Sci.* **8**, 1 (1967); M. J. Herbert and C. M. Harsh, *J. Comp. Physiol. Psychol.* **37**, 81 (1944); E. R. John, P. Chesler, F. Bartlett, I. Victor, *Science* **159**, 1489 (1968); E. L. Thorndike, *Psychol. Rev. Monogr.* **2**, 8 (1898); W. H. Thorpe, *Ibis* **93**, 1, 252 (1951); R. M. Yerkes, *J. Soc. Psychol.* **5**, 271 (1934).
9. H. Klüver, *Biol. Symp.* **7**, 263 (1942).
10. J. Rosen, D. Stein, N. Butters, *Science* **173**, 355 (19 → D. Stein, J. Rosen, J. Graziadei, D. Mishkin, J. Brink, *ibid.* **166**, 528 (1969).
11. R. W. Doty, in *The Visual System: Neurophysiology and Psychophysics* (Springer, Berlin, 1961), p. 228.
12. R. Galambos, T. Norton, G. P. Frommer, *Exp. Neurol.* **18**, 8 (1967).
13. S. S. Winans, *Science* **158**, 944 (1967); K. L. Chow, *Brain Res.* **9**, 363 (1968).
14. D. H. Hubel and T. N. Wiesel, *J. Physiol. London* **160**, 106 (1962); V. B. Mountcastle, *J. Neurophysiol.* **20**, 408 (1957); T. P. S. Powell and V. B. Mountcastle, *Johns Hopkins Med. J.* **105**, 108 (1959).
15. B. D. Burns and G. K. Smith, *J. Physiol. London* **164**, 238 (1962).
16. G. L. Gerstein and N. Kiang, *Exp. Neurol.* **10**, 1 (1964).
17. G. L. Gerstein, *Int. Biophys. Congr. 3rd* Cambridge, Mass. (1969).
18. J. H. O'Brien and S. S. Fox, *J. Neurophysiol.* **22**, 3 (1969).
19. F. Morrell, in *The Neurosciences: A Study Program*, G. C. Quarton, T. C. Melnechuk, F. O. Schmitt, Eds. (Rockefeller Univ. Press, New York, 1969), p. 452.
20. D. N. Spinelli and M. Weingarten, *Exp. Neurol.* **15**, 347 (1966); D. N. Spinelli, K. H. Pribram, M. Weingarten, *Fed. Proc.* **25**, 2173 (1966); M. Weingarten and D. N. Spinelli, *Exp. Neurol.* **15**, 363 (1966); K. L. Chow, D. F. Lindsley, M. Gollender, *J. Neurophysiol.* **31**, 729 (1968).
21. D. R. Humphrey, E. M. Schmidt, W. D. Thompson, *Science* **170**, 758 (1970).
22. E. R. Kandel and W. A. Spencer, *Physiol. Rev.* **48**, 65 (1968).
23. G. Adam, W. R. Adey, R. W. Porter, *Nature* **209**, 920 (1966); N. A. Buchwald and C. D. Hull, *Fed. Proc.* **25**, 105 (1966); Y. Hori and N. Yoshii, *Psychol. Rep.* **16**, 241 (1965); J. Bureš and O. Burešova, *Proc. Int. Congr. Physiol. Sci. 23rd* Tokyo (1965), p. 359; N. Yoshii and H. Ogura, *Med. J. Osaka Univ.* **11**, 1 (1960).
24. J. H. O'Brien and S. S. Fox, *J. Neurophysiol.* **23**, 3 (1969).
25. F. Morrell, J. Engel, W. Bouris, *Electroencephalogr. Clin. Neurophysiol.* **23**, 89 (1967).
26. E. R. John, *Mechanisms of Memory* (Academic Press, New York, 1967).
27. V. N. Dumenko, in *Contemporary Problems of Electrophysiology of the Central Nervous System*, I. N. Knipst, Ed. (Academy of Science, Moscow, 1967), p. 104; R Galambos and G. C. Sheatz, *Amer. J. Physiol.* **203**, 173

(1962); E. V. Glivenko, T. A. Korol'kova, G. D. Kuznetsova, *Fiz. Zh. SSSR Im. I. M. Sechenova* **48**, 1026 (1962); E. R. John and K. F. Killam, *J. Pharmacol. Exp. Ther.* **125**, 252 (1959); E. R. John, D. S. Ruchkin, J. Villegas, *Science* **141**, 429 (1963); I. N. Knipst, in *Contemporary Problems of Electrophysiology of the Central Nervous System*, I. N. Knipst, Ed. (Academy of Science, Moscow, 1967), p. 127; T. B. Shvets, *ibid.*, p. 160; M. N. Livanov, *Proc. Int. Cong. Physiol. Sci.* Leyden (Excerpta Medica Foundation, Amsterdam, 1962), vol. 1, p. 899; N. Yoshii, P. Pruvot, H. Gastaut, *Electroencephalogr. Clin. Neurophysiol.* **9**, 595 (1957).

28. E. A. Asratyan, *Proc. Int. Congr. Physiol. Sci.* (1968), vol. 7, p. 20; H. Begleiter and P. Platz, *Science* **166**, 769 (1969); K. F. Killam and A. J. Hance, *Proc. Int. Congr. Physiol. Sci.* Tokyo (1965), vol. 4, p. 1125; A. L. Leiman, thesis, Univ. of Rochester, New York (1962); G. T. Sakhuilina and G. K. Merzhanova, *Electroencephalogr. Clin. Neurophysiol.* **20**, 50 (1966); N. Yoshii, *Proc. 22nd Int. Congr. Physiol. Sci. 22nd* Leyden (1962), vol. 2, p. 1988.
29. O. D. Creutzfeldt, S. Watanabe, H. D. Lux, *Electroencephalogr. Clin. Neurophysiol.* **20**, 1 (1966a); J. Calvet, M. C. Calvet, J. Scherrer, *ibid.* **17**, 109 (1964); G. H. Fromm and H. W. Bond, *ibid.* **22**, 159 (1967); J. D. Frost, Jr., *ibid.* **23**, 179 (1968a); G. L. Gerstein, *ibid.* **20**, 68 (1961); J. D. Green, D. S. Maxwell, W. J. Schindler, C. Stumpf, *ibid.* **23**, 403 (1960); H. Jasper and C. Stefanis, *ibid.* **18**, 541 (1965); Y. Lass, *ibid.* **25**, 503 (1968); C. Stefanis, *ibid.* **15**, 1054 (1963); V. Fujita and T. Sato, *J. Neurophysiol.* **27**, 1011 (1964); D. P. Purpura and R. J. Shofer, *ibid.* **27**, 117 (1964); M. R. Klee, K. Offenloch, J. Tigges, *Science* **147**, 519 (1965); A. D. Robertson, *Nature* **209**, 757 (1965); J. Thomas, P. Groves, M. Verzeano, *Experientia Basel* **24**, 360 (1968); N. N. Vasilevs, *Byull. Eksp. Biol. Med.* **59**, 597 (1965).
30. E. R. John and P. P Morgades, *Electroencephalogr. Clin. Neurophysiol.* **27**, 205 (1969a); E. R. John and D. Walker, unpublished.
31. ———, *Exp. Neurol.* **23**, 412 (1969b).
32. ———, *Commun. Behav. Biol.* **3**, 181 (1969).
33. S. S. Fox and J. H. O'Brien, *Science* **147**, 888 (1965); M. Verzeano, R. C. Dill, E. Vallecalle, P. Groves, J. Thomas, *Experientia Basel* **24**, 696 (1968b).
34. F. Morrell, E. R. John, S. Nord, T. Hoeppner, E. Grastyan, unpublished data, 1971.
35. E. R. John, R. N. Herrington, S. Sutton, *Science* **155**, 1439 (1967).
36. M. Clynes, M. Kohn, J. Gradijan, *IEEE (Inst. Elec. Electron. Eng.) Int. Conv. Dig.* (1967), p. 218; M. Buchsbaum and P. Fedio, *Electroencephalogr. Clin. Neurophysiol.* **26**, 266 (1969); K. H. Pribram, D. N. Spinelli, M. C. Kambeck, *Science* **157**, 94 (196[illegible] C. Fields, *ibid.* **165**, 1377 (1969).
37. E. R. John, F. Bartlett, M. Shimokochi, in preparation.
38. J. Majkowski, *Electroencephalogr. Clin. Neurophysiol.* **23**, 521 (1967).
39. R. W. Thatcher, thesis, Univ. of Waterloo, Ontario, Canada (1970).
40. D. S. Ruchkin and E. R. John, *Science* **153**, 209 (1966).
41. D. Kleinman and E. R. John, unpublished data.
42. E. R. John, M. Shimokochi, F. Bartlett, *Science* **164**, 1519 (1969).
43. E. R. John, F. Bartlett, M. Shimokochi, in preparation.
44. M. Schwartz, *Science* **169**, 303 (1970).
45. D. S. Ruchkin, *Commun. Behav. Biol.* **5**, 383 (1971).
46. F. Bartlett and E. R. John, *Science* **169**, 304 (1970).
47. E. R. John, F. Bartlett, M. Shimokochi, *Neural Readout from Memory During Generalization*, 16 mm sound and color film (1969). This 40-minute teaching film may be borrowed from this laboratory.
48. J. S. Barlow, L. Morrell, F. Morrell, "Mechanisms of Orienting Reactions in Man," *Proceedings of an International Colloquium* (Czechoslovakia Academy of Sciences, Prague, 1967); M. Haider, P. Spong, D. B. Lindsley, *Science* **145**, 180 (1964); R. Klinke, H. Fruhstorfer, P. Finkenzeller, *Electroencephalogr. Clin. Neurophysiol.* **26**, 216 (1968); L. A. Riggs and P. Whittle, *Vision Res.* **7**, 441 (1967); V. A. Rusinov, in *Central Nervous System and Behavior*, M. A. B. Brazier, Ed. (Josiah Macy Jr. Foundation, New York, 1959), pp. 249–312; S. Sutton, P. Tueting, J. Zubin, E. R. John, *Science* **155**, 1436 (1967).
49. H. Weinberg, W. G. Walter, H. J. Crow, *Electroencephalogr. Clin. Neurophysiol.* **29**, 1 (1970).
50. E. R. John, in *Information Storage and Neural Control*, W. S. Fields and W. Abbot, Eds. (Thomas, Springfield, Ill., 1963), pp. 243–282.
51. R. N. Herrington and P. Schneidau, *Experientia* **24**, 1136 (1968).
52. This work was supported by NIH grant NS09924, NSF grant GB-27559 and the Health Research Council of New York grant I-375. I thank Drs. E. Grastyan and F. Morrell for discussions of these ideas and Drs. D. S. Ruchkin, P. Easton, and F. Bartlett for their programming and technical assistance.

How the Chinese Scientist Survives

Leo A. Orleans

Just as every poor harvest in China precipitates predictions of mass starvation, the periodic intensification of attacks against scientists—particularly foreign-trained scientists—inevitably brings forth prophesies that science in China has suffered a mortal blow from which it is not likely to recover. This is not really surprising, for the attacks against these "reactionary bourgeois authorities and specialists" are indeed unrestrained. The latest assaults on scientists during the Cultural Revolution (1966 to 1969)—including the disappearance of all scientific journals, prescribed sojourns on the commune or in a factory, and the requirement that scientists "lay bare their mistakes and shortcomings"—were especially severe and, for Westerners, difficult to accept. Yet, despite the excesses of the Cultural Revolution, China has once again bounced back, and Chinese science and technology seem to be none the worse for it. To understand how this is possible, it is necessary to look at the developments during the past half-dozen years not from a Western perspective, but from the vantage point of the Chinese scientist.

One law of nature that is particularly applicable to the scientist in China is natural selection. The human species is quite adept at adapting to, even prospering in, what might seem to be a hostile environment. This practical trait of the Chinese people is well known—they are masters at enduring adversity, whether caused by nature or by man. After almost 20 years of recurring thought reforms, surely the ideologically weak species have been wiped out by now, or, to use their favorite expression, "plucked out." Those who survived thought reforms prior to the Cultural Revolution surely managed to endure the most recent onslaught as well. If security and professional activity require reciting the thoughts of Mao or confessing the crime of one's heritage, then the pain is certainly compensated by the ensuing salvation.

The Experience of Kuo Mo-jo

Kuo Mo-jo—who is, among other things, vice-chairman of the Standing Committee of the National People's Congress, chairman of the Chinese Academy of Sciences, and China's top intellectual—is not the average scientist; but, because thousands of other scientists suffered an experience very similar to his during the Cultural Revolution, Kuo Mo-jo's well-publicized experience is worth relating. In early May 1966, Kuo made a public confession that was reported around the world. In his criticism of his past actions and thoughts he said (*1*):

> In the past scores of years, a pen was always in my hand, writing and translating works amounting to many millions of words. However, in the light of present-day standards, what I have written, strictly speaking, should all be burned. It has no value—none whatsoever.

The author is China research specialist in the reference department of the Library of Congress, Washington, D.C. 20540. This article is adapted from a larger study which appeared in *People's Republic of China: An Economic Assessment*, Joint Economic Committee of the U.S. Congress (Government Printing Office, Washington, D.C., 1972).