

**U.S. copyright law  
(title 17 of U.S. code)  
governs the reproduction  
and redistribution of  
copyrighted material.**

**Downloading this  
document for the  
purpose of  
redistribution is  
forbidden.**

# **Neurocomputing**

## **Foundations of Research**

**Edited by James A. Anderson and Edward Rosenfeld**

© 1988 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

The MIT Press  
Cambridge, Massachusetts  
London, England

## 5 Introduction

(1950)  
K. S. Lashley

### In search of the engram

*Society of Experimental Biology Symposium, No. 4: Psychological Mechanisms in Animal Behavior*

Cambridge: Cambridge University Press, pp. 454–455, 468–473, 477–480

Karl Lashley was a Harvard neuropsychologist who studied the formation and storage of memory. This famous paper, which summarized years of work on the biology of memory, addresses directly the important question of distributed versus localized representation of information, which is central to modeling of the nervous system. We have included Lashley's summary, and a couple of small excerpts that bear on this issue.

As has often been pointed out, metaphors for the brain are usually based on the most complex device currently available: in the seventeenth century the brain was compared to a hydraulic system, and in the early twentieth century to a telephone switchboard. Now, of course, we compare the brain to a digital computer. One of the editors (JAA) remembers a number of talks, reports, and papers in psychology in the 1960s and 1970s that presented block diagrams of mental functions that looked for all the world like machine code for a PDP-8. Perhaps in the late 1980s we shall invert this process and make metaphors for computer design based on brain architecture!

The telephone switchboard analogy in its crudest form had a sensory input joined to a motor output by way of specific connections in the central exchange, i.e., the cerebral cortex. Memory was largely concerned with setting up the internal switching. That is, if  $x$  is seen, then do  $y$ , because the signal is properly routed. Real models of this type were, of course, more subtle, but they tended to predict that if you destroyed a chunk of brain containing the discrete paths between input and output, the memory would vanish along with the connection.

When Lashley tried to do this by lesioning various parts of rat brains, it did not work. It turned out that quite large areas of rat cortex could be removed and the animal was still capable of demonstrating learned behavior. The robustness of these results led Lashley to suggest at first that there was relatively little localization of function in cerebral cortex. More precise modern techniques have shown that this is untrue, and that there is a great deal of localization in cerebral cortex: for example, there are numerous topographic maps of visual space onto the surface of our cortex, maps of the body surface in areas of cortex associated with the somatosensory system, maps of frequency in the auditory system, and so on. In fact, fairly precise topographic localization of function is ubiquitous in cortex, and in other regions of the brain as well.

However, it would be equally in error to suppose that stored information is located only in a very small set of cells and their connections. For example, given the nature of the visual world, any complex image is liable to excite a large number of neurons in the visual system. Exactly how many is a matter for debate. Those believing in localization might estimate as few as a few hundred discharging neurons are adequate to represent complex perceptions. Those believing that information is more spread out

would estimate a few percent of the relevant parts of cortex are necessary for adequate representation. A fascinating pair of papers by Barlow (1971,1985)discusses this issue.

This debate is alive and well in the connectionist literature. Models have appeared assuming extreme element selectivity, where one active element corresponds to a complex,high level perception, and others have assumed widely distributed representations, where activity of a single element tells little about the item being represented. These different assumptions are important, because they tend to lead to different kinds of computations and have quite different statistical properties.

It should be pointed out that one natural way to realize semantic networks with connectionist hardware often requires highly selective elements, if one makes the natural assumption that single nodes in the network correspond to single active elements in a connectionist system. If one wants to use connectionist networks as computing devices—say, in artificial intelligence—there is absolutely no reason not to use highly selective,highly meaningful primitive elements (seeFeldman and Ballard, paper 29).

However, if one is trying to model brain or psychological function, the evidence seems to us to be compelling that information is distributed, in that many elements must be simultaneously active to represent information.

Lashley's summary states strongly his conclusions opposing localization and in favor of distribution. The modern picture is somewhat less clear. Some low level reflex pathways seem to be quite localized. There is also evidence that several kinds of memory, apparently associated with different brain structures, exist (Thompson, 1986).

However for the cerebral cortex, even though Lashley overstates some of his conclusions and modern evidence indicates a greater degree of localization than he thought existed, we think that there is compelling evidence that in the brain "... every instance of recall requires the activity of literally millions of neurons" and that "... the reservation of individual synapses for special associative reactions is impossible." Lashley conceives of brain operation as large scale patterns of activation involving a great many active neurons leading to other large patterns of activity. Many, though not all, connectionist models incorporate this assumption explicitly or implicitly.

## References

- H. Barlow (1971), "Single units and sensation: a neuron doctrine for perceptual psychology," *Perception* 1:371–394.
- H. Barlow (1985), "The role of single neurons in the psychology of perception," *Quarterly Journal of Experimental Psychology* (The Twelfth Bartlett Memorial Lecture) A37:121–145.
- R. F. Thompson (1986), "The neurobiology of learning and memory," *Science* 233:941–947.

(1950)  
K. S. Lashley

**In search of the engram**

*Society of Experimental Biology Symposium, No. 4: Psychological Mechanisms in  
Animal Behavior*

Cambridge: Cambridge University Press, pp. 454–455, 468–473, 477–480

**I. Introduction**

‘When the mind wills to recall something, this volition causes the little [pineal] gland, by inclining successively to different sides, to impel the animal spirits toward different parts of the brain, until they come upon that part where the traces are left of the thing which it wishes to remember; for these traces are nothing else than the circumstance that the pores of the brain through which the spirits have already taken their course on presentation of the object, have thereby acquired a greater facility than the rest to be opened again the same way by the spirits which come to them; so that these spirits coming upon the pores enter therein more readily than into the others.’

So wrote Descartes just three hundred years ago in perhaps the earliest attempt to explain memory in terms of the action of the brain. In the intervening centuries much has been learned concerning the nature of the impulses transmitted by nerves. Innumerable studies have defined conditions under which learning is facilitated or retarded, but, in spite of such progress, we seem little nearer to an understanding of the nature of the memory trace than was Descartes. His theory has in fact a remarkably modern sound. Substitute nerve impulse for animal spirits, synapse for pore and the result is the doctrine of learning as change in resistance of synapses. There is even a theory of scanning which is at least more definite as to the scanning agent and the source of the scanning beam than is its modern counterpart.

As interest developed in the functions of the brain the doctrine of the separate localization of mental functions gradually took form, even while the ventricles of the brain were still regarded as the active part. From Prochaska and Gall through the nineteenth century, students of clinical neurology sought the localization of specific memories. Flechsig defined the association areas as distinct from the sensory and motor. Aphasia, agnosia and apraxia were interpreted as the result of the loss of memory images either of objects or of kinaesthetic sensations of movements to be made. The theory that memory traces are stored in association

areas adjacent to the corresponding primary sensory areas seemed reasonable and was supported by some clinical evidence. The extreme position was that of Henschen, who speculated concerning the location of single ideas or memories in single cells. In spite of the fact that more critical analytic studies of clinical symptoms, such as those of Henry Head and of Kurt Goldstein, have shown that aphasia and agnosia are primarily defects in the organization of ideas rather than the result of amnesia, the conception of the localized storing of memories is still widely prevalent (Nielsen, 1936).

While clinical students were developing theories of localization, physiologists were analysing the reflex arc and extending the concept of the reflex to include all activity. Bechterew, Pavlov and the behaviourist school in America attempted to reduce all psychological activity to simple associations, or chains of conditioned reflexes. The path of these conditioned reflex circuits was described as from sense organ to cerebral sensory area, thence through associative areas to the motor cortex and by way of the pyramidal paths to the final motor cells of the medulla and cord. The discussions of this path were entirely theoretical, and no evidence on the actual course of the conditioned reflex arc was presented.

In experiments extending over the past 30 years I have been trying to trace conditioned reflex paths through the brain or to find the locus of specific memory traces. The results for different types of learning have been inconsistent and often mutually contradictory, in spite of confirmation by repeated tests. I shall summarize to-day a number of experimental findings. Perhaps they obscure rather than illuminate the nature of the engram, but they may serve at least to illustrate the complexity of the problem and to reveal the superficial nature of many of the physiological theories of memory that have been proposed.

**VI. The Engram within Sensory Areas (Equipotential Regions)**

The experiments reported indicate that performance of habits of the conditioned reflex type is dependent

upon the sensory areas and upon no other part of the cerebral cortex. What of localization within the sensory areas? Direct data upon this question are limited, but point to the conclusion that so long as some part of the sensory field remains intact and there is not a total loss of primary sensitivity, the habit mechanism can still function. Thus, in a series of experiments attempting to locate accurately the visual cortex of the rat, parts of the occipital lobes were destroyed in a variety of combinations. In these experiments it appeared that, so long as some part of the anterolateral surface of the striate cortex (the projection field of the temporal retina corresponding to the macula of primates) remained intact, there was no loss of habit. Any small part of the region was capable of maintaining the habits based on discrimination of intensities of light (Lashley, 1935b).

In a later experiment an attempt was made to determine the smallest amount of visual cortex which is capable of mediating habits based upon detail vision. The extent of visual cortex remaining after operation was determined by counting undegenerated cells in the lateral geniculate nucleus. Discrimination of visual figures could be learned when only one-sixtieth of the visual cortex remained (Lashley, 1939). No comparable data are available on postoperative retention, but from incidental observations in other experiments I am confident that retention would be possible with the same amount of tissue.

In an early study by Franz (1911) the lateral surfaces of the occipital lobes of the monkey were destroyed after the animals had been trained in pattern and colour discrimination. These operations involved the greater part of what is now known to be the projection field of the macula. There was no loss of the habits. I have destroyed the cortex of the retrocalcarine fissure (the perimacular field) without destroying visual memories. The results with monkeys thus support the more ample data for the rat; the visual memory traces survive any cortical lesion, provided some portion of the field of acute vision remains intact.

This lack of definite habit localization might really have been predicted from psychological data alone. Analysis of the effective stimuli in discriminative learning reveals that the association is independent of particular sensory nerve fibres. It is a response to a pattern of excitation which may vary widely in position on the sensory surface and consequently in cortical projection. The reactions involved in motor habits show the same sort of functional equivalence; a motor habit is not a predetermined set of muscular contractions but is a series of movements in relation to bodily posture and to the complex pattern of the environment. The

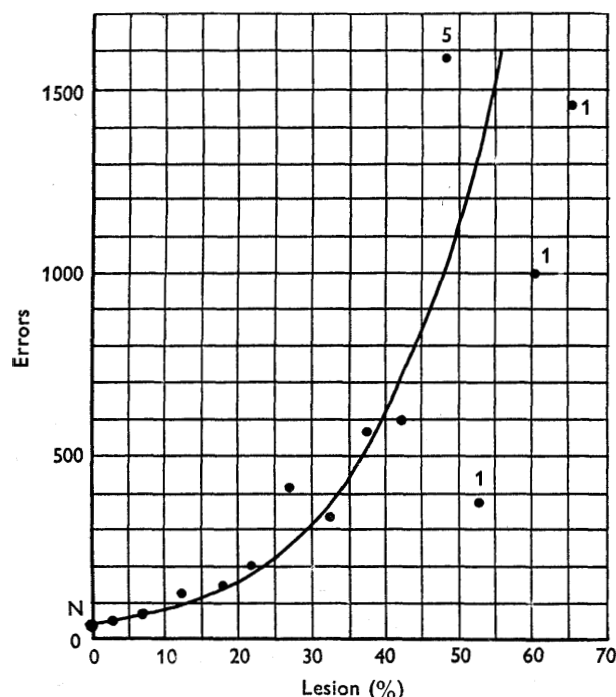
writing of one's name, for example, is not a stereotyped series of contractions of particular muscles but is a series of movements in relation to the body planes which can be performed with any motor organ and with any degree of amplitude.

I have not time here to report in detail the experiments which justify the conclusion that neither the afferent path nor the efferent is fixed by habit. The mass of evidence accumulated by gestalt psychologists shows conclusively that it is the pattern and not the localization of energy on the sense organ that determines its functional effect. Similar motor equivalence is demonstrated by a variety of less systematic evidence. The psychological studies, like the more limited direct experiments on the brain, point to the conclusion that the memory trace is located in all parts of the functional area; that various parts are equipotential for its maintenance and activation.

### VII. Facilitative Functions in Learning and Retention (Mass Action)

The experiments thus far reported have been concerned almost entirely with discriminative habits requiring only an association between a single sensory stimulus and a motor response. A very different picture develops in experiments with other types of learning. If rats are trained in the maze and then have portions of the cortex removed, they show more or less loss of the habit. If a small amount of cortex is destroyed, 5–10%, the loss may be scarcely detectable. If large amounts, say 50% or more, are destroyed, the habit is completely lost, and relearning may require many times as much practice as did initial learning. The amount of loss, measured in terms of the practice required for relearning, is, on the average, closely proportional to the amount of cortex destroyed. Text-fig. 8 shows the relation for one group of rats on a relatively difficult maze with eight *culls de sac*. There is some evidence that the more difficult the task, the greater the relative effect of the larger lesions (Lashley, 1929; Lashley & Wiley, 1933). Similar results have been obtained with latch-box learning and retention (Lashley, 1935a). So far as it is possible to analyse the data from more than 200 diverse operations, the amount of loss from a given extent of cortical destruction is about the same, no matter what part of the cerebral hemispheres is destroyed, provided that the destruction is roughly similar in both hemispheres.

The explanation of this quantitative relationship is difficult. In learning the maze the rat certainly employs a variety of sensory cues, visual, tactile, kinaesthetic, olfactory, possibly auditory. Brain injuries destroy vari-



**Figure 8** The relation of errors in maze learning to extent of cerebral damage in the rat. The extent of brain injury is expressed as the percentage of the surface area of the isocortex destroyed. Data from 60 normal and 127 brain-operated animals are averaged by class intervals of 5% destruction. The curve is the best fitting one of logarithmic form. For lesions above 45% the number of cases (indicated by numerals on the graph) is too small for reliability. (After Lashley & Wiley, 1933.)

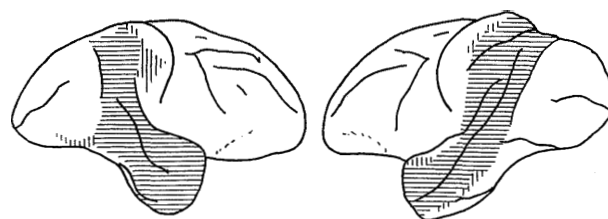
ous sensory fields and the larger the lesion the greater the reduction in available sense data. The production of different amounts of sensory deficit would thus appear to be the most reasonable explanation of the quantitative relation between habit loss and extent of lesion (Hunter, 1930; Finley, 1941). Sensory deficit certainly plays a role in it. In the experiment on effects of incisions through the cortex, which was described earlier, the severity of loss of the maze habit correlated highly with the interruption of sensory pathways, as determined from degeneration of the thalamus.

However, sensory loss will not account for all of the habit deterioration. There is evidence which shows that another more mysterious effect is involved. In the first place, destruction of a single sensory area of the cortex produces a far greater deficit in maze or latch-box performance than does loss of the corresponding sense modality. A comparison was made of the effects on retention of the latch-box habits of combined loss of vision, vibrissae touch, and the anaesthesia to touch and movement produced by sectioning the dorsal half of the spinal cord at the third cervical level. This latter

operation severs the columns of Gall and Burdoch, which convey tactile and kinaesthetic impulses, and also severs the pyramidal tracts which have a dorsal position in the rat. The combined peripheral sense privation and section of the pyramids produced less loss of the latch-box habits than did destruction of a single sensory area of the cortex (Lashley, 1935a). Secondly, when blind animals are trained in the maze, the removal of the primary visual cortex produces a severe loss of the habit with serious difficulty in re-learning, although the animals could have used no visual cues during the initial learning (Lashley, 1943).

A possible explanation of this curious effect was that the rat forms concepts of spatial relations in visual terms, as man seems to do, and that the space concepts are integrated in the visual cortex. The visual cortex might then function in the formation of spatial habits, even when the animal loses its sight. To test this Tsang (1934) reared rats blind from birth, trained them as adults in the maze, then destroyed the visual cortex. The resultant loss of the maze habit by these animals was as severe as in animals which had been reared with vision. The hypothesis concerning the formation of visual space concepts was not confirmed.

Our recent studies of the associative areas of the monkey are giving similar results to those gained with rats. Visual and tactile habits are not disturbed by the destruction singly, either of the occipital, parietal, or lateral temporal regions, so long as the primary sensory fields remain. However, combined destruction of these regions, as shown in Text-fig. 9, does produce a loss of the habits with retarded relearning. Higher level functions, such as the conditional reaction, delayed reaction, or solution of the multiple stick problem, show deterioration after extensive damage in any part of the cortex. The capacity for delayed reaction in monkeys, for example (to remember in which of two boxes food was placed), may be seriously reduced or abolished by removal either of the prefrontal lobes or of the occipital associative cortex or of the temporal lobes. That is, small lesions, embracing no more than a single associative area, do not produce loss of any



**Figure 9** Minimal lesion which produces disturbances in tactile or visual memory in the monkey.

habit; large lesions produce a deterioration which effects a variety of habits, irrespective of the sensori-motor elements involved.

Results such as these have led me to formulate a theory of mass action or mass facilitation. It is, essentially, that performance of any function depends upon two variables in nervous activity. The reaction mechanism, whether of instinctive or of learned activity, is a definite pattern of integrated neurons with a variable threshold of excitability. The availability of such patterns, the ease with which they can be activated, is dependent upon less specific facilitative effects. This facilitation can come from a variety of sources. Some instinctive behaviour seems to require hormonal activation, probably a direct chemical effect upon specific nervous elements. Emotional facilitation may produce a temporary activation. Continued activity of related mechanisms may facilitate the whole group of associated reactions; a sort of warming-up effect.

There are indications (Krechevsky, 1936), although little systematic evidence, that the severity of postoperative amnesia varies with the intensity of motivation. Rats trained in a discrimination without punishment with electric shock for errors may show loss of the habit after lesions which do not produce loss in animals which were trained with punishment. The greater effects of cortical lesions in monkeys than in rats may be in part a result of the greater difficulty in getting consistent motivation in the higher animals. In man an amnesia often seems to be a difficulty rather than impossibility of recall; recall may be possible but only with extreme effort and fatigue. I believe that the evidence strongly favours the view that amnesia from brain injury rarely, if ever, is due to the destruction of specific memory traces. Rather, the amnesias represent a lowered level of vigilance, a greater difficulty in activating the organized patterns of traces, or a disturbance of some broader system of organized functions.

In interpreting apparent loss of memory after cerebral damage, extreme caution is necessary. The poor performance in tasks may be due to the destruction of specific associative connexions, but is instead generally, I believe always, the result rather of interference with a higher level functional patterning. Some experiments of Dr Klüver's (personal communication) illustrate this point. Monkeys were trained in a variety of discriminative reactions calling for use of different sense modalities by a method that required them to pull in the stimulus objects by attached strings. Extensive lesions in different cortical areas all caused loss of these habits. The monkeys simply pulled the strings at random. They were retrained in the discrimination of

weights. When this was learned, the habits based on other sense modalities (reactions to intensities of light, for example) returned spontaneously. What had been disturbed by all the operations was the set or attitude to compare stimuli, not the specific memory of which one was correct.

This example, perhaps illustrates at a primitive level the characteristic of amnesias as seen clinically. Apparent loss of memory is secondary to a disorder in the structuring of concepts. Some physiological mode of organizing or integrating activity is affected rather than specific associative bonds.

## X. Summary

This series of experiments has yielded a good bit of information about what and where the memory trace is not. It has discovered nothing directly of the real nature of the engram. I sometimes feel, in reviewing the evidence on the localization of the memory trace, that the necessary conclusion is that learning just is not possible. It is difficult to conceive of a mechanism which can satisfy the conditions set for it. Nevertheless, in spite of such evidence against it, learning does sometimes occur. Although the negative data do not provide a clear picture of the nature of the engram, they do establish limits within which concepts of its nature must be confined, and thus indirectly define somewhat more clearly the nature of the nervous mechanisms which must be responsible for learning and retention. Some general conclusions are, I believe, justified by the evidence.

(1) It seems certain that the theory of well-defined conditioned reflex paths from sense organ via association areas to the motor cortex is false. The motor areas are not necessary for the retention of sensori-motor habits or even of skilled manipulative patterns.

(2) It is not possible to demonstrate the isolated localization of a memory trace anywhere within the nervous system. Limited regions may be essential for learning or retention of a particular activity, but within such regions the parts are functionally equivalent. The engram is represented throughout the region.

(3) The so-called associative areas are not storehouses for specific memories. They seem to be concerned with modes of organization and with general facilitation or maintenance of the level of vigilance. The defects which occur after their destruction are not amnesias but difficulties in the performance of tasks which involve abstraction and generalization, or conflict of purposes. It is not possible as yet to describe these defects in the present psychological terminology. Goldstein (1940) has expressed them in part as a shift from the abstract

to the concrete attitude, but this characterization is too vague and general to give a picture of the functional disturbance. For our present purpose the important point is that the defects are not fundamentally those of memory.

(4) The trace of any activity is not an isolated connexion between sensory and motor elements. It is tied in with the whole complex of spatial and temporal axes of nervous activity which forms a constant substratum of behaviour. Each association is oriented with respect to space and time. Only by long practice under varying conditions does it become generalized or dissociated from these specific Co-ordinates. The space and time Co-ordinates in orientation can, I believe, only be maintained by some sort of polarization of activity and by rhythmic discharges which pervade the entire brain, influencing the organization of activity everywhere. The position and direction of motion in the visual field, for example, continuously modifies the spinal postural adjustments, but, a fact which is more frequently overlooked, the postural adjustments also determine the orientation of the visual field, so that upright objects continue to appear upright, in spite of changes in the inclination of the head. This substratum of postural and tonic activity is constantly present and is integrated with the memory trace (Lashley, 1949).

I have mentioned briefly evidence that new associations are tied in spontaneously with a great mass of related associations. This conception is fundamental to the problems of attention and interest. There are no neurological data bearing directly upon these problems, but a good guess is that the phenomena which we designate as attention and interest are the result of partial, subthreshold activation of systems of related associations which have a mutual facilitative action. It seems impossible to account for many of the characters of organic amnesias except in such general terms as reduced vigilance or reduced facilitation.

(5) The equivalence of different regions of the cortex for retention of memories points to multiple representation. Somehow, equivalent traces are established throughout the functional area. Analysis of the sensory and motor aspects of habits shows that they are reducible only to relations among components which have no constant position with respect to structural elements. This means, I believe, that within a functional area the cells throughout the area acquire the capacity to react in certain definite patterns, which may have any distribution within the area. I have elsewhere proposed a possible mechanism to account for this multiple representation. Briefly, the characteristics of the nervous network are such that, when it is subject to any pattern of excitation, it may develop a pat-

tern of activity, reduplicated throughout an entire functional area by spread of excitations, much as the surface of a liquid develops an interference pattern of spreading waves when it is disturbed at several points (Lashley, 1942a). This means that, within a functional area, the neurons must be sensitized to react in certain combinations, perhaps in complex patterns of reverberatory circuits, reduplicated throughout the area.

(6) Consideration of the numerical relations of sensory and other cells in the brain makes it certain, I believe, that all of the cells of the brain must be in almost constant activity, either firing or actively inhibited. There is no great excess of cells which can be reserved as the seat of special memories. The complexity of the functions involved in reproductive memory implies that every instance of recall requires the activity of literally millions of neurons. The same neurons which retain the memory traces of one experience must also participate in countless other activities.

Recall involves the synergic action or some sort of resonance among a very large number of neurons. The learning process must consist of the attunement of the elements of a complex system in such a way that a particular combination or pattern of cells responds more readily than before the experience. The particular mechanism by which this is brought about remains unknown. From the numerical relations involved, I believe that even the reservation of individual synapses for special associative reactions is impossible. The alternative is, perhaps, that the dendrites and cell body may be locally modified in such a manner that the cell responds differentially, at least in the timing of its firing, according to the pattern of combination of axon feet through which excitation is received.