A Physiological Mechanism for Hebb’s Postulate of Learning
(theoretical/synapse/membrane/neurobiology/polarity reversal)

GUNTHER S. STENT

Department of Molecular Biology, University of California, Berkeley, Calif. 94720

Communicated by Donald A. Glaser, January 22, 1973

ABSTRACT  Hebb’s postulate of learning envisages that activation or inactivation of extant synaptic contacts in plastic neural networks depends on the synchronous impulse activity of pre- and postsynaptic nerve cells. The physiological mechanism proposed here for this process posits that at synapses acting according to Hebb’s postulate, the receptors for the neurotransmitter are eliminated from the postsynaptic membrane by the transient reversals of the sign of membrane polarization that occur during action potential impulses in the postsynaptic cell. But, since the release of neurotransmitter drives the membrane potential of the synaptic zone towards a level about half-way between the negative-inside resting potential and the positive-inside action potential, it would follow that the membrane patches surrounding the receptors of a synapse whose activity has contributed to setting off the postsynaptic impulse would be spared the full extent of the noxious polarity reversal. This mechanism can account for a neurophysiologically documented example of the operation of Hebb’s postulate, namely the plasticity of the connections between fourth- and fifth-order neurons in the visual cortex of cats.

In 1949 D. O. Hebb (1) formulated his “neurophysiological postulate” of learning which states: “When an axon of cell A is near enough to excite a cell B and repeatedly and persistently takes part in firing it, some growth or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.” Recently D. Marr (2) has applied Hebb’s postulate to the neural circuitry of the cerebellum and shown that the idea of activation of extant, albeit ineffective, synaptic contacts by synchronous impulse activity of pre- and postsynaptic cells can explain, in theory at least, how that organ could learn to refine gross body movements commanded by the cerebral cortex. Although no direct neurophysiological evidence in support of the operation of Hebb’s postulate in the cerebellum has yet been adduced, the results of studies by T. N. Wiesel and D. H. Hubel (3, 4) on the plasticity of connections in the cerebral cortex lend strong support to the following complementary statement of that postulate: “When the presynaptic axon of cell A repeatedly and persistently fails to excite the postsynaptic cell B while cell B is firing under the influence of other presynaptic axons, metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is decreased.” The purpose of this paper is to propose a physiological mechanism that explains how activation on inactivation of extant synaptic contacts could depend on the synchronous impulse activity of pre- and postsynaptic cells.

Binocular cortical connections

The pair of eyes of carnivores and primates see very nearly the same visual field. The sensory inputs received from any given point in that visual field by the mosaic of primary photoreceptors of the two retinas is conducted separately via second-, third-, and fourth-order neurons to the cerebral cortex, where the two monocular pathways converge on a common fifth-order neuron, or binocular cortical cell. Hubel and Wiesel showed that in newborn kittens, many of the monocular neuronal connections that lead from retina to cortical binocular cells already exist before any visual experience (5). However, they found that during the first few months of a kitten’s postnatal development there occurs a critical period during which abnormal visual experience can cause a profound and life-long alteration of these connections (3). Thus, if the lid of one eye of the newborn kitten is sewn closed and allowed to open only after the first 3 months of life, then the fifth-order cortical neurons are no longer excited by visual inputs into the deprived eye, despite the fact that the primary photoreceptors of that eye and their connections to the fourth-order neurons are demonstrably intact. The congenitally binocular cortical cells have thus become monocular and henceforth are excited only by the normal eye. It is not, however, simple disuse that causes the congenital connections from fourth- to fifth-order neurons to wither, since many of these connections survive in a kitten whose both eyes were closed during the first three months of life. Hence, it appears that during the critical period the converging presynaptic fourth-order neurons from the two eyes compete for control over the fifth-order postsynaptic cell, so that a reduction in the frequency of use of one set of synapses permits the other set to take complete charge. Thus, under monocular closure the excitatory influence of the normal eye would be enhanced at the expense of the deprived eye. In binocularly deprived kittens, however, there would occur a severe reduction in the frequency of synaptic use of the connections from both eyes, and hence neither connection has to give way to the other.

Hubel and Wiesel were able to gain a most revealing further insight into the nature of the binocular competition for control of the postsynaptic neuron (4). For they found that in kittens rendered crosseyed by cutting one of the eye muscles soon after birth, most of the fifth-order cortical cells become monocular, some cells responding only to the left and others only to the right eye. This means that the conjoint survival of both sets of cortical synapses requires not merely their equally frequent, but their synchronous activity. Such synchronous activity occurs in the case of the normal kitten whose two eyes fix the scene so that a given point of the visual field projects on corresponding parts of the two retinas. But, since the crosseyed kitten is unable to fix both eyes on the same point in its visual field and corresponding areas of its two retinas thus receive projections from two completely different spatial points, the fifth-order neuron congenitally connected to those corresponding areas now receives asynchronous, though on the average equally frequent, activity from the two eyes. Hence, maintained asynchrony of presynaptic activity appears to allow
one or the other of the two sets of synapses to take complete charge of the postsynaptic cell.

A detailed study of the critical period showed that the vulnerability of cortical connections to abnormal visual experience begins with the first month of the kitten’s life. In the fifth or sixth week, a brief 3-day closure of one eye suffices to convert binocular into monocular fifth-order neurons and induce long-lasting blindness in the deprived eye. The critical period comes to an end during the fourth month of the kitten’s life. Temporary monocular eye closure or surgical production of cross-eyedness after that time no longer has any effect on the binocular character of the connections from fourth-order neurons to fifth-order cortical cells in the visual pathway (6).

A postsynaptic effect
In order to fathom the mechanism that may be responsible for this plasticity of the binocular connections, we may ask first how the synaptic contact made by cell A could sense whether or not its activity is asynchronous with that of other cells converging on the same postsynaptic cell B. This sensing might be done presynaptically, by reciprocal collateral connections between the converging axons. But, it would seem very unlikely that such reciprocal presynaptic collaterals exist in the manner required by this explanation among the hundreds of axons that actually converge on a single fifth-order cortical cell. For this very elaborate presynaptic network of reciprocal collaterals among converging fourth-order axons would not only add a great mass of neural tissue to the surround of each cortical cell, but would also introduce formidable complications to the already sufficiently complicated problem of how, during embryonic development, the specific connections between fourth-order and fifth-order cells arise in the first place. It appears much more likely, therefore, that the asynchrony is detected postsynaptically by a very simple rule: the activity of the synapse of cell A upon cell B is manifestly asynchronous with the activity of synapses of other cells converging on cell B if most of the impulses that arise in cell B occur while the synapse of cell A is inactive. This is, in fact, a complementary statement of the rule for postsynaptic detection of synchronous activity which Marr posited in his cerebellar theory of learning (2).

Second, we may inquire into the nature of the metabolic change that causes the synapse of cell A to decrease its efficiency in contributing to the firing of cell B upon an asynchronous activity pattern of the two cells. This change could occur either presynaptically or postsynaptically. A presynaptic change would result in a decrease in the amount of neurotransmitter released per impulse travelling in the axon of cell A, whereas a postsynaptic change would result in a decrease in the effect that the release by cell A of a given amount of transmitter exerts on the specific ion permeability of the membrane of cell B. The physiological mechanism to be proposed here assumes that the change is, in fact, of the postsynaptic kind, mainly for two reasons. First, since, as was already argued, the asynchrony of activity is itself likely to be detected postsynaptically, it is simpler to assign the detection system and metabolic change to the same cell. Second, whereas no long-term presynaptic metabolic changes of the kind needed to explain the connectional plasticity of the visual cortex have thus far come to light, a type of long-term postsynaptic metabolic change is known which, as will now be set forth, does provide a suitable model system.

The acetylcholine receptor
Acetylcholine is the neurotransmitter at the synapses formed by the axon terminals of motor neurons upon skeletal muscle fibers (7). The acetylcholine released upon arrival of an impulse at the presynaptic axon terminal combines with a specific protein, the acetylcholine receptor, present in the postsynaptic muscle-fiber membrane (8). This combination results in a depolarization of the postsynaptic membrane that sets off an impulse in the muscle fiber (9). In a normal muscle fiber the presence of the acetylcholine receptor protein is restricted to the immediate vicinity of the synapse, but upon denervation of the muscle by cutting its motor nerve, receptors appear within a few days along the entire length of the fiber membrane (10–13). The denervated muscle thus returns to the embryonic state in which prior to receipt of the first motor innervation, receptors are uniformly distributed over the whole membrane (14). A recent study by Lømo and Rosenthal (15) has shown that impulse activity of the postsynaptic fiber can re-restrict the presence of acetylcholine receptors in denervated adult muscle to the synaptic zone. In this study the motor nerve leading to a rat leg muscle was cut, so as to reduce greatly the frequency of impulses in the muscle fibers. A few days later, by which time newly formed acetylcholine receptors had already appeared along the whole fiber length, electrodes were attached to the muscle surface and a quasi-normal impulse frequency was restored to the muscle fibers by chronic direct electrical stimulation. It was found that such chronic stimulation for a few days causes the receptors to disappear again from the fiber membrane and to be reconfined to the narrow zone formerly occupied by the now-defunct synapse. Since the newly formed acetylcholine receptors of denervated rat muscle are metabolically stable in nonstimulated muscle for periods of days (13), it would follow that impulse activity in stimulated muscle causes an active elimination of the receptor from all parts of the membrane except the synaptic zone.

In order to explain the active elimination of receptor by impulse activity, the theory to be put forward here makes the following simple postulation, for which there is no present experimental support: the stability of the receptor protein in the cell membrane, in whose lipid phase, according to the "fluid mosaic" view of membrane structure the receptor "floats" (16), depends on the maintenance of an inward gradient of electrical potential. Such a gradient exists, of course, across the membranes of all nerve and muscle cells, thanks to the negative resting potential of a few dozen millivolts of the cell interior. This postulated dependence of the stability of the receptor on membrane polarization might be attributable to its being an electric dipole, thanks to which the protein molecule is suitably oriented in the field lines of the membrane dielectric. However, an action potential impulse, which brings the cell interior to a positive potential of a few dozen millivolts, generates an outward gradient of potential across the membrane. This sudden change in the directional sign of the field lines in the membrane dielectric would then dislodge the receptor molecules from their functional positions in the fluid membrane.

Protection of receptor in the synaptic zone
How, in the light of this postulate, is it possible that despite continuous impulse activity in a normal muscle, functional receptor protein survives in the membrane of the synaptic zone? Here it is necessary to introduce a second postulate,
namely, that there are two postsynaptic membrane states with respect to receptor stability, an immature, plastic and a mature, nonplastic state. In the mature, nonplastic state a modification of the membrane of the synaptic zone has taken place—for instance, insertion into the membrane of an additional protein species—that protects the receptor protein against the noxious effects of sign reversal of the potential field lines. This modification would be only one of several developmental “trophic” effects which the presynaptic axon terminal appears to exert on the postsynaptic membrane. For instance, after receiving its innervation the embryonic muscle-fiber membrane first thickens in the synaptic zone and then develops the deep infoldings that greatly increase the postsynaptic surface area facing the presynaptic axon terminal (17). Or, upon innervation of the embryonic muscle, the activity of the enzyme acetylcholinesterase is greatly increased in the fiber membrane of the synaptic zone (18, 19). Thus, from this point of view, in normal adult muscle fibers the membrane of the synaptic zone has reached the mature, nonplastic state, whereas the remainder of the adult membrane still remains in the immature plastic state characteristic of its embryonic past.

We now reach the crux of my theory: the postulation of the elimination of the receptor from the cell membrane by impulse-associated reversals of polarity sign allows also for the stability of the receptor protein in the synaptic zone of an immature, plastic membrane, if—and only if—in accord with Hebb’s postulate, the presynaptic axon terminal repeatedly and persistently takes part in firing the postsynaptic cell. As was found by Fatt and Katz (9) soon after Hebb put forward his postulate, the amplitude of an action potential set off in a muscle fiber upon stimulation of the motor nerve is lower by 10–20 mV when recorded by an electrode inserted into the fiber near the synaptic zone than the maximum amplitude recorded at a distance of a few millimeters from the synapse. However, if the action potential is set off by direct passage of current into the muscle fiber, then the action potential reaches its maximum amplitude in the synaptic zone. This finding was later explained by del Castillo and Katz (20) and Takeuchi and Takeuchi (21), who showed that acetylcholine released by the stimulated motor nerve terminal effects its depolarization of the postsynaptic membrane by greatly increasing the local selective permeability for both Na\(^{+}\) and K\(^{+}\). This increase of permeability drives the synaptic membrane toward a potential of about −10 to −20 mV, (about midway between the positive Na\(^{+}\) and the negative K\(^{+}\) equilibrium potentials) and, thus, to a potential still of the same sign as the resting potential. Hence, in the synaptic zone itself the transmitter prevents the resulting action potential from reaching the full amplitude attained elsewhere in the fiber.

It follows, therefore, that the patches of membrane surrounding the receptors in the zone of a synapse whose activity has set off (or has contributed to setting off) an impulse in the postsynaptic cell would not experience as extensive a reversal of polarity in the potential field lines during the action potential as the general membrane outside the synaptic zone*

*The difficulty of an exact theoretical treatment, as well as the lack of relevant data, make it hard to estimate the actual magnitude of this effect. The net resistance of an individual acetylcholine-activated ion channel, estimated to be about 10\(^{12}\) ohms (22), is the sum of the intrinsic resistance of the channel itself and of the convergence resistance in the volume conductors

**Crossed innervation**

Although the formation and disappearance of acetylcholine receptors under various conditions have now been studied extensively for many years, no case has so far been put in evidence where this phenomenon is actually responsible for a change in the efficacy of synaptic transmission. However, L. A. Marotte and R. F. Mark (29) have observed functional changes in synapses between a motor neuron and a muscle fiber that can be readily explained in these terms.

Their work concerned the antagonism between the synaptic contacts of both homonymous and heteronymous nerves on one of a pair of antagonistic eye muscles in carp and goldfish. In these experiments one of the oculomotor nerves was cut at its point of emergence from the brainstem far away from the eye. The second nerve was cut near the point of termination on its muscle and that muscle was removed from the eye. A month or so after this operation, the proximal stump of the second nerve had crosstovernerved the heteronymous muscle, to which it was much nearer than the short proximal stump of the first, homonymous nerve. Functional synapses arose in that crosstovernervation that resulted in inappropriate reflexive rotation of the eyeball. Within another month, however, the short proximal stump of the first nerve had regrown towards the eye and managed to reinnervate its homonymous muscle, reestablishing its own functional synapses upon the already incorrectly innervated muscle fibers. Soon after this reinnervation, the heteronymous nerve lost its power to cause contractions in the crosstovernervated muscle and appropriate reflexive eye rotation was reestablished.

The loss of influence of the heteronymous nerve upon reinnervation by the homonymous nerve does not seem to be simply attributable to a gross degeneration of the crosstovernervating axon terminals (such as that which develops upon cutting the motor nerve), since electronmicroscopic examination of the muscle fibers at this stage suggested that the synapses made by the heteronymous nerve still seem to be structurally intact. Instead, it would appear that because of some metabolic change invisible by electron microscopy, e.g., loss of receptors for the transmitter on the postsynaptic membrane, the synapses of the heteronymous nerve lose their capacity to transmit the presynaptic nerve impulses to the postsynaptic muscle fibers.

**Binocular cortical cells**

We are finally ready to propose a physiological mechanism for the operation of Hebb’s postulate, as exemplified by the plasticity of the connections between the monocular fourth-order and binocular fifth-order neurons in the visual cortex of kittens. The identity of the neurotransmitter responsible for impulse transduction at these cortical synapses has not been ascertained. Hence, for the purpose of this proposal it is envisaged, first of all, that this transmitter resembles acetyl-
choline in two essential regards (one itself hypothetical and the other factual): its active receptor molecules are susceptible to elimination from the postsynaptic membrane by polarity reversals induced by action potentials; and its effect on the ionic permeability of the postsynaptic cell drives the membrane potential to a level considerably more negative than the positive Na\(^+\) equilibrium potential to which the cell is driven in the self-regenerating action potential mechanism.

Second, it is envisaged that the binocular fifth-order neurons of the visual cortex of neonatal kittens are still in the embryonic, plastic state. In this state the postsynaptic neurotransmitter receptors retain their susceptibility to elimination by polarity reversals induced by action potentials while, however, still being replaceable because of the formation of new receptors by the protein-synthesizing machinery of the immature cell. The developmental maturation of the fifth-order neurons would then proceed in two stages. The first stage corresponds to the beginning of the critical period after the first month of life and marks the end of the process of receptor formation. The second stage corresponds to the end of the critical period during the fourth month of life and marks the attainment by the cell membrane of the nonplastic state, in which receptors have become stabilized against polarity reversal.

The cortical consequences of abnormal visual experience during the kitten's critical period can now be accounted for. Under **monocular closure**, a high impulse traffic reaches only those fourth-order axon terminals that report visual input from the open eye. Transmission of these impulses via congenitally active synapses sets off frequent action potentials in the binocular fifth-order neurons, and the associated polarity reversals eliminate receptor molecules from the postsynaptic membrane facing the terminals of the quiescent axons reporting from the closed eye. Hence, these synapses become permanently nonfunctional. The receptors facing the active axon terminals are, of course, protected, since here the temporal congruence of the pattern of presynaptic transmitter release and of postsynaptic action potentials reduces the amplitude of the noxious polarity reversals in the synaptic zone. Thus, the congenitally binocular cortical cells are transformed into monocellular ones. Under **binocular closure**, little impulse traffic reaches the cortical cells from either eye, and hence there occur only infrequent action potentials and their associated polarity reversals. Thus, the congenital receptors would not be eliminated from the postsynaptic membrane facing either class of quiescent axon terminals. Most of the synapses thus remain functional and the binocular character of the cortical cells is preserved despite total visual deprivation. Finally, under **cross-eyedness**, a high impulse traffic reaches the visual cortex from both eyes, and hence there occur frequent action potentials and polarity reversals in the cortical cells. But, since here the connections converging on the same postsynaptic cell from both eyes receive inputs from different parts of the visual field, their axon terminals carry an asynchronous pattern of impulse activity. Thus, while the axon terminals reporting from the right eye release transmitter and set off an action potential in the cortical cell, the resulting polarity reversal eliminates some receptor molecules from the postsynaptic membrane facing the momentarily quiescent axon terminals reporting from the left eye. And a postsynaptic action potential set off by the axon terminals reporting from the left eye similarly eliminates some receptors from the membrane facing the momentarily quiescent axon terminals reporting from the right eye. Inasmuch as the degree to which an axon terminal can influence postsynaptic events by release of transmitter depends on the density of receptor molecules present at that moment in its opposite postsynaptic membrane, it follows that under cross-eyedness the congenital binocularity of the cortical cell is inherently unstable. As soon as, due to the character of the visual input pattern, the axon terminals reporting from one eye happen to have been much more active for some period than those reporting from the other eye, the synaptic connections from the first eye will have become stronger than those from the second eye and will thus dominate the later postsynaptic impulse pattern, even though the average activity of the axon terminals reporting from both eyes may henceforth be equal. Because of this feature of positive feedback in the competition for receptor survival, the connections from the first eye will therefore become stronger and stronger relative to those from the second eye, until all receptors have been eliminated from the postsynaptic membrane facing the set of axon terminals from the second eye and the cell will have been rendered monocular.

The postulates of this theory readily account for the character of the critical period. Abnormal visual experience prior to its onset has no effect on the binocular character of the cortical cells, because receptor formation still proceeds and any receptors eliminated from the postsynaptic membrane by an asynchronous input activity pattern can still be replaced. After the onset of the 2–3-months-long critical period, transient monocular closure lasting for only a few days produces a long-lasting modification of the cortical cells, because receptor formation no longer proceeds and, once eliminated from the postsynaptic membrane, the receptors are irreplaceable. Abnormal visual experience after the end of the critical period has no effect on the binocular character of the cortical cells because the membrane has reached the mature, nonplastic state and the receptors are no longer susceptible to elimination by any asynchronous input activity pattern.

### Inhibitory synapses

Although the basic formulation of Hebb's postulate did not refer to changes in the efficacy of inhibitory synapses, which are likely to play no less important a role in learning than excitatory synapses, the following statement extends the postulate to the plasticity of inhibitory connections: "When the presynaptic axon of cell A repeatedly and persistently fails to inhibit the postsynaptic cell B while cell B is firing under the influence of other presynaptic axons, metabolic changes take place in one or both cells such that A's efficiency, as one of the cells inhibiting B, is decreased." This version of the postulate is, in fact, teleologically even more plausible than its original form since a cell would in general have little need for maintaining an inhibitory synaptic connection if the corresponding presynaptic axon were mainly inactive during the activity of other axons supplying excitatory inputs to the same cell. The physiological mechanism proposed here can readily account also for this version of the postulate. For, if an inhibitory synapse is consistently inactive during impulse activity of the postsynaptic cell, then the polarity reversals produced by the action potentials would eliminate the receptors from its postsynaptic membrane. Hence, the synapse would lose its efficacy. But, transmitter released from the postsynaptic axon terminal at an inhibitory synapse greatly in-
creases the local selective permeability for Cl− and/or K+, so that the synaptic membrane is driven towards the negative Cl− and K+ equilibrium potentials, or to a level even more negative than the resting potential of the cell (24). Hence, while transmitter happens to be present at an inhibitory synapse the amplitude of an action potential set off in the postsynaptic cell by excitatory inputs will be very greatly attenuated in the zone of that synapse. The receptors of an active inhibitory synapse would thus be spared the noxious polarity reversals and its efficacy would be maintained.

Tests of the theory
The central assumption of the theory put forward here is that the neurotransmitter receptors at synapses that behave according to Hebb’s postulate of learning are eliminated from the postsynaptic membrane by reversals in the sign of membrane potential. It should be possible to test this assumption, at least for the model system of skeletal muscle fibers. First, the assumption implies that the restriction of acetylcholine receptors to the synaptic zone of denervated muscle by chronic electrical stimulation is a direct consequence of the action potentials induced in the muscle fibers and not, as might also be possible, an indirect consequence, involving from a mechanical or chemical effect of the muscle contractions initiated by these action potentials. Thus, the assumption demands that restriction of the receptors to the synaptic zone by chronic stimulation occurs also under conditions where the contractile apparatus has been rendered inoperative. Second, the assumption predicts that chronic stimulation of a denervated muscle should not restrict acetylcholine receptors to the synaptic zone if pulses of acetylcholine are applied directly to extrasynaptic regions of the muscle-fiber membrane in synchrony with the chronic electrical stimuli. [Because of the desensitization of the receptors upon maintained exposure to their transmitter (25), this experiment could not, of course, be done by simply bathing the muscle in an acetylcholine solution during the chronic stimulation.] Two other key assumptions of the theory are, at least in principle, testable for the plastic binocular connections in the visual cortex of kittens. If, as assumed here, the asynchrony of convergent activity is detected postsynaptically, then pharmacological or electrophysiological prevention of the occurrence of action potentials in the fifth-order cortical neurons during a 3-day period of monocular closure in the critical period should avert the loss of the binocular character of the cells. And if, as assumed furthermore, the metabolic change that decreases the efficiency of influence from the deprived eye occurred postsynaptically, then direct localized application of transmitter to the membrane of the deactivated synaptic zones of the abnormal, monocular cortical neurons should be without effect on the postsynaptic membrane potential. As far as the visual cortex is concerned, both of these tests are presently within the realm of Gedankenexperimente, since formidable technical difficulties would be in the way of their realization. But it would seem that with nervous ensembles of a complexity less awesome than the mammalian cerebral cortex such tests might soon be feasible. In particular, the physiological basis of long-term plastic changes in the character and efficacy of synapses in comparatively simple invertebrate central nervous systems—for instance the changes recently discovered in the leech ventral nerve cord by Jansen and Nicholls (26)—ought to be susceptible to direct experimental exploration.

I am indebted to H. B. Barlow, W. B. Kristan, T. N. Wiesel, and R. Zucker for helpful comments and suggestions. This investigation was supported by Grants GM 17568 from the National Institute of General Medical Sciences, U.S. Public Health Service; and GB 31933X from the National Science Foundation.