A Brief History of the Hebbian Learning Rule

PETER MILNER
McGill University

Abstract
Hebb included a version of his neural postulate of learning in his MA thesis (1932). It seems to be a translation of Pavlovian conditioning into neural terms. The version that appears in his book, *The Organization of Behavior* (1949), speaks of synapses rather than routes, but the idea of simultaneous firing of afferent and efferent elements is common to both versions. The postulate was adopted by groups interested in programming computers to learn and think. It was also one of the hypotheses tested by neurophysiologists in their search for the synaptic mechanism of learning. So far the neurophysiologists have had more success than the logicians; the NMDA synapse appears to explain the main features of Hebb's postulate.

In his monograph, *The Organization of Behavior,* Hebb (1949) presented a theory concerning the way stimuli might be represented in the brain. It was the first such attempt to be widely accepted by psychologists and it has had a strong influence on subsequent theories. Commonsense decrees that brain representations must be learned, at least in the case of representations of the shapes of human artefacts such as the letters of the alphabet, tools, and buildings. The radical empiricism of early 20th century psychology led Hebb to go the whole hog and base his model on the assumption that the neurons of the newborn visual cortex are randomly interconnected. In order to proceed, Hebb then had to specify the conditions under which visual input might organize these connections.

In 1932, when Hebb was a part-time graduate student at McGill, he submitted an MA thesis entitled Conditioned and Unconditioned Reflexes and Inhibition. The gist of the thesis is that spinal reflexes are the result of prenatal Pavlovian conditioning; in his more mature years Hebb referred to it as nonsense (Hebb, 1980). The McGill library had, in the meantime, lost its copy of the thesis, so Hebb's evaluation could not readily be verified. Recently, however, Professor Richard Brown of Dalhousie University, in connection with a presentation he made to the Society of Neuroscience, tracked down another copy of the thesis. The part of it dealing with the ontogeny of spinal reflexes may well be nonsense, as Hebb maintained, but Brown (2001, 2002) made the interesting discovery that the thesis includes an analysis of the neural learning mechanism underlying Pavlovian conditioning that foreshadows the one he later presented to explain the learning of visual representations. Based on the observations of Pavlov and others that a stimulus occurring repeatedly at about the same time as a response, acquires a connection to that response, Hebb concluded that:

> *An excited neuron tends to decrease its discharge to inactive neurons, and to increase this discharge to any active neuron, and therefore to form a route to it, whether there are intervening neurons between the two or not. With repetition this tendency is prepotent in the formation of neural routes.* (Hebb, 1932, p. 8)

An accompanying diagram makes it clear that this postulate was indeed a transposition of Pavlov's conditioning paradigm to the neural level.

When, some dozen or so years later, Hebb needed to specify the neural learning mechanism responsible for the acquisition of shape representations in the brain, it seems that he consciously or unconsciously returned to the formulation in his thesis. The neurological postulate was presented in *The Organization of Behavior* as follows:

> Let us assume then that the persistence or repetition of a reverberatory activity (or "trace") tends to induce lasting cellular changes that add to its stability. The assumption can be precisely stated as follows: *When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes a part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.* (Hebb, 1949, p. 62)

The idea that a connection between two neurons takes place only if both neurons are firing at about the same time is common to both of Hebb's formulations. This idea was not entirely new. From the earliest times philosophers and psychologists have recognized that learning often involves the establishment of an association between two or more concurrent events. Hartley (1746/1959) made the ingenious suggestion that stimuli produce neural vibrations, and that ideas consist of smaller after-vibrations (vibratiuncles). When two events occur in close succession, Hartley speculated that the frequencies of their vibrations draw together so that if one occurs it excites vibratiuncles of the other.

With the improved knowledge about the nervous system that came in the latter part of the 19th century, the psychologist Alexander Bain made a more modern attempt to explain associative learning in neural terms. In his book, *Mind and Body* (1874), written when Pavlov was still a medical student and even before the neuron doctrine had been enunciated (Waldeyer, 1891), Bain wrote:

> We know what are the conditions of making an acquirement, or of fixing two or more things together in the memory. The separate impressions must be made together, or flow in close succession; and they must be held together for a certain length of time, either on one occasion, or on repeated occasions. Now to each impression, each sensation or thought, there corresponds physically a group or series of nerve-currents; when two impressions concur, or closely succeed one another, the nerve-currents find some bridge or place of continuity, better or worse, according to the abundance of nerve-matter available for the transition. In the cells or corpuscles where the currents meet and join, there is, in consequence of the meeting, a *strengthened connexion* or *diminished obstruction* – a preference track for that line over other lines where no continuity has been established. (p. 117)

Of course, nobody spoke of the Bain synapse or Bain rules, probably because in 1874 the synapse was unknown, therefore unnamed, and nobody was performing computer simulations of brain processes.

HEBB'S POSTULATE

When it appeared in 1949, *The Organization of Behavior* was seen by some as the foundation upon which to build a working model of the brain. For them, Hebb's postulate, which they usually call the Hebb Rule, was a means to an end, an important feature of a system in which most of the functional connections had to be acquired by experience. Modelers have not hesitated to tinker with the rule to suit their own purposes. Neurophysiologists interested in synaptic plasticity constitute another group that refers frequently to Hebb's postulate, usually calling it the Hebb Synapse. It is not an integral component of their research, and their recognition of its relevance to their work came only after a delay of many years. Thus the postulate has been adopted by two very different disciplines, the brain modelers using it as a tool and the neurophysiologists as a hypothesis to be tested.

My assigned task for this introductory essay was to "present what Hebb actually said." Unfortunately, after the postulate was released to the world, Hebb actually said almost nothing more about it, or either of its subsequent lines of development. Most of what I shall present concerns the treatment his idea has received at the hands of others.

The Hebb Rule

In 1950, general purpose electronic computers were about to invade businesses and the universities, inspiring a few enthusiasts to dream of the day when all our heavy thinking could be unloaded onto them. One such enthusiast was Nat Rochester of IBM research lab who started a project called "Conceptor" to mechanize the cell assembly, Hebb's design for an element of thought.

If Hebb had any enthusiasm for the project he kept it under strict control. The notes for his book show that he struggled long and hard to work out his theory about the growth of cell assemblies, and he probably knew better than anyone the extent to which it was held together by wishful thinking. After its publication Hebb had the heavy responsibility of building a department as well as research projects of his own that he found more interesting than computer modeling.

The most obvious flaw in Hebb's design was that as connection increased in accordance with Hebb's postulate, activity would eventually spread and increase uncontrollably throughout the cortex. When Rochester came to Hebb for advice, I was in the process of writing a Mark II version of the cell assembly (Milner, 1957) in which this problem was overcome by the introduction of an inhibitory system, so Hebb sent Rochester to see me and, as far as I remember, took no further interest in the project. Several models were tested but from the point of view of this account, the main point is that the paper eventually published (Rochester, Holland, Haibt, & Duda, 1956) contained the following paragraph:

> Now another property of the neurons will be described. When neuron A participated in firing neuron B, the synapse that enabled A to stimulate B was increased in magnitude unless it already reached the limit of 938, in which case it remained constant. This characteristic was our version of Hebb's basic neurophysiological postulate (which was thereupon cited). (p. S82)

The paper stops short of referring to the postulate as Hebb's rule, but it does state that, in order to work, "an additional rule must be introduced" to prevent connections from growing without limit. Although the authors tried to be optimistic, the tests did not succeed in generating cell assemblies that would work as Hebb intended, as is indicated in this summary:

> It will be shown that Hebb's scheme is unlikely to work with neurons of the type described so far. It will also be shown that by suitably improving the neurons and by making the network more complex, cell assemblies can be made to form spontaneously. It will further be shown that these cell assemblies are not entirely satisfactory but that there is a plausible course for further investigation. (p. S83)

IBM abandoned the Conceptor project shortly afterwards, but one of the authors, Holland, continued the pursuit for a number of years.

Another visitor at about the same time was Frank Rosenblatt who was at the time a bright and very ambitious graduate student at Cornell University. He had read *The Organization of Behavior,* and like me, thought he could improve on it. Unlike me, he was a skilled mathematician and logician so most of his improvements were intended to make the theory more amenable to simulation and testing by computer. He presented a logical argument to show that cell assemblies could not be generated in a purely homogeneous mass of neurons and introduced a "minimally constrained" system that he called a perceptron (Rosenblatt, 1958). Some neurons were designated as association units, others as response units. Response units were facilitated by some of the association units and inhibited other association units.

Rosenblatt criticized Hebb's learning postulate on the somewhat precarious grounds that he could not imagine a learning process that would only work at successful synapses. Learning in Rosenblatt's perceptron affected the whole neuron and depended only on the number of times it fired. This suggests to me that although representations of stimuli may be generated, the perceptron would not be able to associate them with each other

The perceptron was ahead of its time. It incorporated some valuable features such as the influence of response activity on associations, and it introduced many of the elements of Parallel Distributed Processing (Rumelhart & McClelland, 1986). Unfortunately, it proved no more successful than the Conceptor and the work stopped when Rosenblatt died at an early age.

The Hebb Synapse

I do not know when the expression "Hebb synapse" first appeared in print. It certainly was not widely used in 1965 when Kandel and Tauc (1965a, b) investigated theories of synaptic change in a giant cell of the abdominal ganglion of *Aplysia*. They make no mention of Hebb, though one type of modification they observed appeared to be in accordance with Hebb's neurological postulate. Later experiments, cited by Kupfermann and Pinsker (1969, 1970) cast doubt on this finding, however. Hebb was cited by Wurtz, Castellucci, and Nusrala (1967) as an originator of the hypothesis they tested that the discharge of a neuron, paired with a subthreshold input to one of its synapses, strengthens that synapse. Their experiments, also using cells of the *Aplysia* abdominal ganglion, did not support the hypothesis.

In 1969, Marr presented a theory of cerebellar learning in which he speculated that the climbing fibres either release a substance, or fire postsynaptic Purkinje cells, in order to strengthen recently active synapses. He attributes the simultaneous pre and postsynaptic mechanism to Hebb (1949). It seems that 20 years after Hebb's book (and nearly 40 years after his MA thesis), physiologists were becoming aware of the postulate.

In 1973, Stent wrote a paper with the title "A physiological mechanism for Hebb's postulate of learning." He uses the finding of Wiesel and Hubel (1963), that kittens reared with one eye occluded do not have binocular cells in the visual cortex, to support the converse of Hebb's postulate:

> When the presynaptic axon of cell A repeatedly and persistently fails to excite the post synaptic 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 firing B is decreased.

The physiological mechanism he put forward to explain this was that discharge of a cell was damaging to the immature synaptic terminals of kittens unless they were protected from the large potential gradient by having discharged immediately before.

Rauschecker and Singer (1981) also used evidence from the effects of visual experience on the connections of the kitten visual cortex to support Hebb's postulate. The title of their paper, "The effects of early visual experience on the cat's visual cortex and their possible explanation by Hebb synapses," is the earliest appearance of "Hebb synapse" in print that I have found.

Carew, Hawkins, Abrams, and Kandel (1984) made a specific test of Hebb's postulate, stating that, ". . . it has had a major impact on thinking about the neural mechanisms of learning (see, for example, Beinenstock et al., 1982; Heidemann, and Changeaux, 1982; Marr, 1969; Rauschecker and Singer, 1981; Stent, 1973; Woody et al., 1978)."

They conclude that the mechanism postulated by Hebb is neither necessary nor sufficient to produce the associative change in synaptic strength that underlies conditioning in *Aplysia*. They still believed that the presynaptic mechanism proposed earlier by Kandel and Tauc (1964) provided a better explanation.

During the 1960s, experiments involving electrical stimulation of the brain became common and several people noticed that seizures sometimes developed after a few weeks. Goddard made a thorough study of this phenomenon, which he called "kindling" (Goddard, McIntyre, & Leech, 1969). Bliss and Lømo (1970) reported that short bursts of tetanic stimulation of the perforant path in the rabbit produced a long-lasting increase in the size of evoked potentials in the dentate gyrus. They called it long-term potentiation (LTP). As the increase in synaptic efficacy was in the same path as the tetanic stimulation, it was uncertain whether the effect was in accordance with Hebb's postulate or not. Wigström, Gustafsson, Huang, and Abraham (1984) showed that single input pulses to synapses on a hippocampal cell did not become more effective if the cell did not fire, but if the cell were depolarized during the pulse by current injected via a microelectrode, the synapse underwent LTP. Current injected 400 ms after a pulse did not change the effectiveness of the synapse The authors concluded that hippocampal LTP occurred in accordance with Hebb's postulate.

It was eventually discovered that one type of LTP takes place at n-methyl-d-aspartate (NMDA) receptors. Their channels are normally blocked by magnesium atoms, held there by the internal negative charge of the cell. When the cell discharges, or is otherwise depolarized, the magnesium is released, allowing a calcium channel to be opened by afferent impulses. The inflowing calcium is responsible for the long-term changes in synaptic effectiveness.

Hebb died in 1985 before this mechanism underlying his postulate was revealed, but if he had lived I do not think he would have been greatly impressed. Hebb was always a maverick; in his autobiography (1980) he wrote that he might have been a physicist now if his father had not pressed, and "For years and years I would read no novel recommended by others." When physiological psychology was beyond the pale Hebb was a physiological psychologist; when it became a bandwagon he jumped off.

Direct correspondence to Peter Milner, Department of Psychology, McGill University, 1205 Dr. Penfield Avenue, Montreal, Quebec, Canada H3A 1B1 (E-mail: peter.milner@mcgill.ca).

Résumé

Hebb a inclus une version de son postulat neural de l'apprentissage dans sa thèse de maîtrise (1932). Il semble s'agir d'une traduction du conditionnement pavlovien en termes neuraux. La version publiée dans son livre *The Organization of Behavior* (1949) renvoie à des synapses plutôt qu'à des routes, mais l'idée d'une décharge simultanée d'éléments afférents et efférents est commune aux deux versions. Le postulat a été adopté par les groupes intéressés à programmer les ordinateurs à apprendre et à penser. C'était aussi l'une des hypothèses testées par les neuropsychologues dans leur recherche du mécanisme synaptique d'apprentissage. Jusqu'à maintenant les neuropsychologues ont eu plus de succès que les logiciens; la synapse du récepteur NMDA semble expliquer les principales particularités du postulat de Hebb.

References

Bain, A. (1874). *Mind and body.* New York: Appleton.

Bienenstock, E. L., Cooper, L. N., & Munro, P. W. (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience, 2*, 32-48

Bliss, T. V. P., & Lømo, T. (1970). Plasticity in a monosynaptic cortical pathway. *Journal of Physiology, 207*, 61P.

Brown, R. E. (2001). D. O. Hebb and the origins of *The Organization of Behavior. Society of Neuroscience Abstracts, 22-23*, (Abstract).

Brown, R. E. (2002). D. O. Hebb's "lost" MA Thesis: The first draft of the Hebb synapse? *Society of Neuroscience, Poster #21.3.*

Carew, T. J., Hawkins, R. D., Abrams, T. W., & Kandel, E. R. (1984). A test of Hebb's postulate at identified synapses which mediate classical conditioning in *Aplysia. Journal of Neuroscience, 4*, 1217-1224.

Goddard, G. V., McIntyre, D. C., & Leech, C. K. (1969). A permanent change in brain function resulting from daily electrical stimulation. *Experimental Neurology, 25*, 295-330.

Hartley, D. (1959). *Various conjectures on the perception, motion, and generation of ideas (1746).* Los Angeles, CA: The Augustan Reprint Society.

Hebb, D. O. (1932). *Conditioned and unconditioned reflexes and inhibition.* Unpublished MA thesis, McGill University, Montreal, QC.

Hebb, D. O. (1949). *The organization of behavior.* New York: Wiley.

Hebb, D. O. (1980). D. O. Hebb. In G. Lindzey (Ed.), A history of psychology in autobiography (Vol. 7) (pp.273-303). San Francisco, CA: Freeman.

Heidmann, T., & Changeaux, J. P. (1982). Un modèle moléculaire de régulation d'efficacité au niveau postsynaptique d'une synapse chimique. *Comptes Rendus des Séances de l'Academie des Sciences- Série Iii, Sciences de la Vie, 295*, 665-670.

Kandel, E. R., & Tauc, L. (1964). Mechanism of prolonged heterosynaptic facilitation. *Nature*, *202*, 145-147.

Kandel, E. R., & Tauc, L. (1965a). Heterosynaptic facilitation in neurones of the abdominal ganglion of *Aplysia depilans*. *Journal of Physiology*, *181*, 1-27.

Kandel, E. R., & Tauc, L. (1965b). Mechanism of heterosynaptic facilitation in the giant cell of the abdominal ganglion of *Aplysia depilans*. *Journal of Physiology*, *181*, 28-47.

Kupfermann, I., & Pinsker, H. (1970). Cellular models of learning and cellular mechanisms of plasticity in *Aplysia*. In K. H. Pribram & D. E. Broadbent (Eds.), *Biology of memory* (pp.163-174). New York: Academic Press.

Marr, D. (1969). A theory of the cerebellar cortex. *Journal of Physiology*, *202*, 437-470.

Milner, P. M. (1957). The cell assembly: Mk. II. *Psychological Review*, *64*, 242-252.

Rauschecker, J. P., & Singer, W. (1981). The effects of early visual experience on the cat's visual cortex and their possible explanation by Hebb synapses. *Journal of Physiology*, *310*, 215-239.

Rochester, N., Holland, J. H., Haibt, L. H., & Duda, W. L. (1956). Tests on a cell assembly theory of the action of the brain, using a large digital computer. *IRE Transactions on Information Theory*, *IT-2*, S80-S93.

Rosenblatt, F. (1958). *The perceptron: A theory of statistical separability in cognitive systems*. Buffalo, NY: Cornell Aeronautical Laboratory.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing. Volume 1: Foundations*. Cambridge, MA: MIT Press.

Stent, G. S. (1973). A physiological mechanism for Hebb's postulate of learning. *Proceedings of the National Academy of Sciences, U.S.A.*, *70,* 997-1001.

Waldeyer, W. (1891). Über einige neuere Forschungen im Gebiete der Anatomie des Centralnervensystems. *Deutsche medizinischer Wochenschrift*, *17*, 1352-1356.

Wiesel, T. N., & Hubel, D. H. (1963). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology*, *26*, 1003-1017.

Wigström, H., Gustafsson, B., Huang, Y.-Y., & Abraham, W. C. (1986). Hippocampal long-term potentiation is induced by pairing single afferent volleys with intracellularly injected depolarizing current pulses. *Acta Physiologica Scandinavica*, *126*, 317-319.

Woody, C. D., Swartz, B. E., & Gruen, E. (1978). Effects of acetylcholine and cyclic GMP on input resistance of cortical neurons in awake cats. *Brain Research, 158*, 373-395.

Wurtz, R. H., Castellucci, V. F., & Nusrala, J. M. (1967). Synaptic plasticity: The effect of action potential in the postsynaptic neuron. *Experimental Neurology*, *18*, 350-368.