The Book of Hebb

Terrence J. Sejnowski*
Howard Hughes Medical Institute
The Salk Institute
10010 North Torrey Pines Road
La Jolla, California 92037
Department of Biology
University of California, San Diego
La Jolla, California 92093

It has been 50 years since Donald Hebb published the Organization of Behavior: A Neuropsychological Theory in 1949. This book was written at a time when behaviorism was dominant in North American psychology. The approach that Hebb advocated, based on what was then known about the brain, was out of favor among psychologists who believed that only external sensory stimuli and motor responses ought to be included in any explanation of behavior. Most neuroscientists have heard about the “Hebb synapse,” but few know why he postulated this learning rule. This is a good time to take a closer look at this book and let Hebb speak for himself.

Hebb was on the faculty of the Psychology Department at McGill University. One of his research interests was the behavioral effects of brain lesions, and he had collaborated with Wilder Penfield, a colleague at McGill, but he was more broadly interested in the development of behavior and learning, which he saw as intimately related. In 1949, much of what we now take for granted about the organization of the nervous system and the properties of neurons was not yet discovered. Hodgkin and Huxley’s landmark series of papers on the ionic basis of the action potential would appear in 1952; the classic paper by Fatt and Katz on the quantal theory of synaptic transmission would appear in the same year. Not much was known about the localization of function in the cortex outside primary sensory and motor areas, and Lashley’s theory of equipotentiality of the cerebral cortex was still influential.

Most of what was then known about cortical neurons and circuits was based on static pictures of neurons stained with the Golgi technique. Even though the “neuron doctrine” went back to Cajal, conclusive evidence that the neuron was indeed a functional unit awaited the electron microscope in the 1950s and recordings from single cortical neurons in the 1960s. In the introduction to his book, Hebb states that his theory “is evidently a form of connectionism, one of the switchboard variety, though it does not deal in direct connections between the afferent and efferent pathways: not an ‘S–R’ psychology, if R means a muscular response. The connections serve rather to establish autonomous central activities, which then are the basis for further learning” (xix). One of the few figures in the book, reproduced in Figure 1, depicts the connections between area 17 (primary visual cortex) and area 18 (extrastriate visual cortex), and is remarkably modern in including feedback projections and long-range lateral connections within cortical areas as well as feedforward connections. He recognized that single synapses were generally too weak in cortex to cause a postsynaptic neuron to fire a spike and that patterns of converging synaptic inputs were required.

The Motivation for the Hebb Synapse

The central problem that concerned Hebb was the origin of what he believed was relatively autonomous activity in the cerebrum: “... we know practically nothing about what goes on between the arrival of the excitation at a sensory projection area and its later departure from the motor area of the cortex” (xvi). Hebb conjectured that cortical circuits admit self-sustaining activity that reverberated in what he called “cell assemblies.” This idea was inspired by evidence for recurrent connections between neighboring cells in the cortex. Although reverberatory activity lasting for up to half a second had been observed by Lorente de Nó, Hebb went further and suggested that such activity in one cortical circuit could through converging projections activate other areas of cortex and lead to a sequence of activations he called a “phase sequence.” Although these ideas remain highly speculative, they reflect recent issues such as spike timing and spike synchrony that today are at the forefront of theoretical research on the cortical neural code (Abbott and Sejnowski, 1999).

Hebb needed a way to sustain persistent reverberatory activity (a “trace”) in cortical circuits. He proposed that patterns of connections between neurons could sustain reverberatory activity if their strengths could be adjusted by an activity-dependent mechanism for synaptic plasticity that he called a “Neurophysiological Postulate”:

**Figure 1. Summary Diagram of Connectivity between Cells in Visual Cortex Taken from Figure 8 in Hebb (1949)**

Cells A and B in primary visual cortex (area 17) receive strong excitation from a visual stimulus in their receptive fields (as do other cells in the cross-hatched region). Cells C and D in extrastriate visual cortex (area 18) provide feedback connections to area 17. Cell E does not receive strong visual input in its receptive field but does receive feedback input from horizontal connections within the cortex. Hebb’s interpretation of this diagram anticipates the recently discovered modulation of the primary receptive field responses in area 17 from visual stimuli outside the classical receptive field.

*E-mail: terry@salk.edu.
When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes 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. (62)

This passage is the origin of the “Hebb learning rule” and made Hebb an adjective. Most good ideas have precursors, and earlier versions of Hebb’s can be found in books by Jerzy Konorski and even William James, but it was Hebb’s version that proved most influential. His words have been interpreted to mean that synaptic plasticity should be based on coincidence detection; that is, strengthening of the synapse should occur when the release of neurotransmitter from a presynaptic terminal coincides with the depolarization of the postsynaptic cell. Evidence for a coincidence detection mechanism has been found in the hippocampus, where long-term potentiation (LTP), discovered there in 1973 by Tim Bliss and Terje Lomo, was shown to be Hebbian (Kelso et al., 1986). LTP of synapses on hippocampal neurons can be elicited by pairing synaptic input with strong depolarizing current, when neither alone produces a long-lasting change, consistent with this interpretation. Furthermore, the induction of LTP at some synapses is controlled by the NMDA receptor, which requires both binding of glutamate and depolarization to allow entry of calcium into the cell. Insofar as the NMDA receptor is a coincidence detector, it might even be called a “Hebb molecule.” The only part that Hebb had apparently not gotten quite right was his statement about the firing of cell B, since LTP could still be induced after fast spiking was abolished by blocking active currents in the postsynaptic neuron, suggesting that cooperativity with other synaptic inputs might be needed to depolarize the dendrite sufficiently to open the NMDA receptor. Another issue is that increases in the strength of a synapse from random coincidences will end inexorably in saturation. Hebb suggested that unused synapses might decay, and a form of long-term depression (LTD) induced by low-frequency activity might provide such decay from spontaneous activity in the cortex. However, if synaptic strengths are to encode long-term memories, it is important to have a mechanism for LTD as specific as that for LTP.

Temporally Asymmetric Synaptic Plasticity
Monosynaptic connections between pairs of cells are best examined with dual intracellular recordings in cortical slices. In an experiment designed to test the importance of relative timing of the presynaptic release of neurotransmitter and the postsynaptic activity to LTP, Markram et al. (1997) paired stimulation of cell A either 10 ms before or after spike initiation in cell B. They found reliable LTP when the presynaptic stimulus preceded the postsynaptic spike, but, remarkably, there was LTD when the presynaptic stimulus immediately followed the postsynaptic spike. Similar results have been found for hippocampal neurons grown in culture (Bi and Poo, 1998; Debanne et al., 1998), between retinal axons and neurons in the optic tectum of frogs (Zhang et al., 1998), and in the electrical line organ of weakly electric fish—this is different from the others in that it is of opposite polarity (presynaptic release before the postsynaptic spike causes LTD) (Bell et al., 1997). Thus, this temporally asymmetric form of synaptic plasticity is widespread in cerebellar as well as cortical structures. In Figure 2, where the time delay between the synaptic stimulus and the postsynaptic spike was varied over a wide range, the window for plasticity is around ±20 ms and the transition between LTP and LTD occurs within a time difference of a few milliseconds.

This temporally asymmetric form of synaptic plasticity has many nice features. First, it solves the problem of balancing LTD and LTP in a particularly elegant way, since chance coincidences should occur about equally with positive and negative relative time delays. Second, when sequences of inputs are repeated in a network of neurons with recurrent excitatory connections, this form of synaptic plasticity will learn the sequence, and the pattern of activity in the network will tend to predict future input. This may occur in the visual cortex where simulations of cortical neurons can become directionally selective when exposed to moving visual stimuli (Rao and Sejnowski, 2000). Similar models have been proposed for neurons in other brain regions, although the temporal window for synaptic plasticity was taken to be 100 ms in the hippocampus (Blum and Abbott, 1996), where there is evidence that the locations of place cells shift to earlier locations in rats running repetitively through a maze (Mehta et al., 1997), and <1 ms in a model for learning auditory localization by the relative timing of spikes from two ears (Gerstner et al., 1996).

Is the temporally asymmetric learning algorithm Hebbian? The rapid transition between LTP and LTD at the moment of temporal coincidence does not conform to the traditional view of a Hebbian synapse. Notice that in Hebb’s formulation the synapse increases in strength “when an axon of cell A is near enough to excite cell B