

## ARTS & HUMANITIES

# Contributions of Yale Neuroscience to Donald O. Hebb's *Organization of Behavior*

Bilal Haider

*Department of Neurobiology, Yale University School of Medicine,  
New Haven, Connecticut*

The neuropsychological concepts found in Donald Hebb's *The Organization of Behavior* have greatly influenced many aspects of neuroscience research over the last half century. Hebb's ideas arose from a rich tradition of research. An underappreciated contribution came from pioneering studies at Yale University. Here, we wish to reconsider these developments, placing particular emphasis on the roles of the neurophysiologists John Fulton, J.J. Dusser de Barenne, and Warren McCulloch and the psychologists Donald Marquis and Ernest Hilgard. These neuroscientists all contributed significantly to the intellectual climate that gave rise to Hebb's remarkable synthesis.

Donald Hebb's *The Organization of Behavior*, published in 1949, is one of the most influential monographs of neuroscience produced in the last century. In it, he formulated his "dual trace mechanism," whereby interconnected and coactive circuits permanently modify the efficacy of activated pathways. Drawing on data and concepts from the molecular to systems level, Hebb's synthesis has come to inspire investigators working on mechanisms of learning and memory at all levels of brain function. Although it seems to modern neuroscientists that Hebb's postulate burst forth fully formed, in fact, like most advances in science, it arose from a rich tradition of research. Hebb's nascent psychological ideas took shape over many

years and combined with emerging physiological concepts to inform the final thesis of his book.

An underappreciated factor contributing to Hebb's synthesis was the dynamic intellectual climate at Yale University. At Yale, pioneering investigations of brain function were occurring at the anatomical, physiological, and behavioral levels. These investigations involved several distinctive personalities, mostly now forgotten, that included American neurophysiologist John Farquhar Fulton, Spanish anatomist Rafael Lorente de Nó, neurophysiologists J.J. Dusser de Barenne from Holland and Warren McCulloch from the United States, and psychologists Donald Marquis and Ernest Hilgard from the United States. We would

---

To whom all correspondence should be addressed: Bilal Haider, Department of Neurobiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510; Tele: 203-737-2926; Fax: 203-785-5263; E-mail: bilal.haider@yale.edu.

†Abbreviations: LTP, long-term potentiation.

like to note briefly how the investigations of these individuals — on anatomical, physiological, and behavioral levels, respectively — help enrich the background to Hebb's work.

Nineteen thirty-eight was a landmark year for Yale neurophysiology. In that year, John Farquhar Fulton, professor of neurophysiology and chairman of the Department of Physiology, in addition to founding the *Journal of Neurophysiology*, published the first comprehensive textbook on the physiology of the nervous system. For the chapter on the cerebral cortex, Fulton turned to the eminent neuroanatomist Rafael Lorente de Nó. De Nó was the last, and perhaps best known, student of Santiago Ramon y Cajal, the Spanish neuroanatomist widely regarded as the founder of modern neuroscience. Lorente de Nó published several studies of the cellular architecture of the cerebral cortex in the 1920s and 1930s. These initial studies were bolstered with a considerable amount of new findings and presented together with theoretical insights for Fulton's book in "Chapter XV, Cerebral Cortex: Architecture, Intracortical Connections, Motor Projections" [5].

Lorente de Nó did not take this chapter merely as an exercise in summarizing his anatomical studies of cortical neurons, but rather to advance new and insightful functional interpretations. In a letter to Fulton concerning the chapter, he writes, "One of the reasons why writing it has been so laborious is that I have verified in my collection of brain sections the truthfulness of every statement in the text and of every line in the drawings" [5]. Most of the chapter is concerned with meticulous characterization of the laminar and cellular organization of the cortex, detailed presentation of comparative anatomy across cortical regions, the termination zone of specific and non-specific thalamic afferents, the axonal and dendritic distributions of pyramidal cells, and the distribution of interneurons across layers. The reader is urged to review the clarity of description and depth of thought in his chapter. Building on the work of his mentor, de Nó provided a basic diagram of cellular connec-

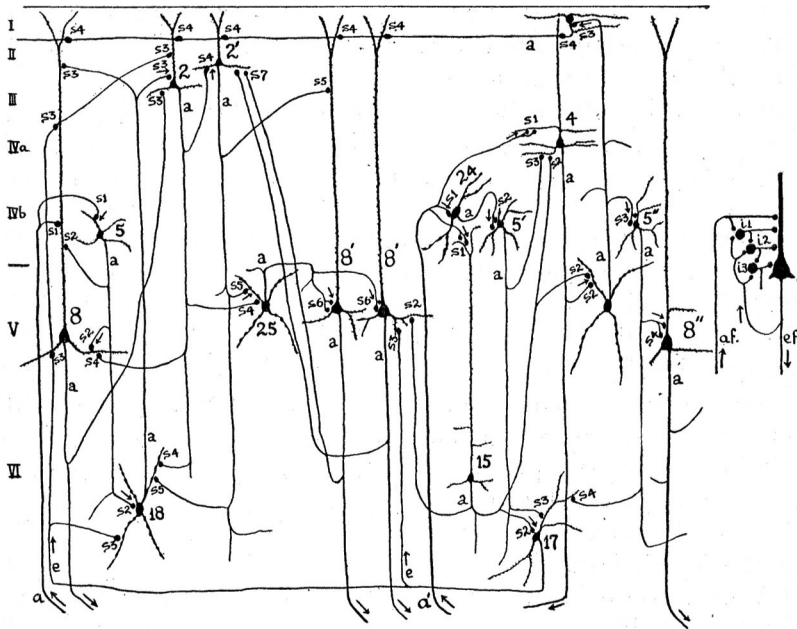
tivity in the cortex, which was to stimulate all further studies of this structure. He did not fail to elaborate on the implications of this cortical scaffold:

Since the impulses conducted by a fibre necessarily passes [sic] into its collaterals, and branches of the descending axons are distributed in the same territories as the cortical afferents, there can be no doubt that the effect of the impulses entering the cortex depends largely upon the impulses at that moment circulating through the descending axons as a result of existing cortical activity.

The importance of this proposition cannot be overstated; Lorente de Nó distinguishes the ongoing activity in the cortex as a process distinct from the arrival of the sensory inputs. De Nó makes a clean break with the theory that signaling through cortical neurons is a strictly hierarchical and serial progression of information. This idea of ongoing, or reverberating, activity would become central to Hebb's thinking when he eventually breaks from the strictly Pavlovian tradition of stimulus-response associations.

In the concluding sections of his chapter in Fulton's book, Lorente de Nó tries to form a comprehensive view of this ongoing activity traveling through connected neurons in the cortex. The chapter contains the essential six-layered "basic circuit" found in all areas of cortex that he examined (Figure 1). His inferences are insightful, effectively formulating an outline of the modern concept of cortical circuit function. He goes on to integrate these essential histological observations with his knowledge of physiological mechanisms and theorizes on the information flow of cortical circuits. He states:

... it is evident that each a impulse causes the cortical cells to be bombarded by a succession of impulses, thus creating in them a constant state of facilitation, and eventually stimulating them to discharge their axons.



**Figure 1. The essential six-layered cortical circuit diagram from Lorente de Nó, in Fulton (1938).** The detail at right illustrates the possibility for 'reverberatory' activity in the cortex.

Some paths are passable, because conditions for summation, such as instantaneous convergence of impulses, are given. Other paths are impassable, but may become passable later when cortical activity creates impulses capable of summing with afferent impulses.

In these propositions, Lorente de Nó essentially describes the dynamic, rapid, and flexible operation of groups of cells, which Hebb would call "cell assemblies," in a modern physiological framework. The outstanding omission is his failure to link explicitly the activity of these closed circuits with psychological theory or cognitive phenomena, as Hebb would later do. However, compared to Hebb's schematic interactions occurring among groups of amorphous cells (see Figure 10 in [15]), Lorente de Nó's model for cortical circuit function presented in Fulton's volume was a much more specific, though less cognitively relevant, fore-runner.

While Lorente de Nó was describing the local connectivity of cortical circuits, others were more concerned with the distant

interconnectivity of cortical regions. Among the earliest and most active contributors to the *Journal of Neurophysiology* was Fulton's Yale colleague and pioneering neurophysiologist J.G. Dusser de Barenne. Fulton's scholarly activities and research interests played a key role in persuading Dusser de Barenne to come to Yale. Fulton had been a doctoral student in Charles Sherrington's laboratory in the spring of 1924, while Dusser de Barenne was there visiting and performing experiments concerning localization of function in the cortex [6]. Fulton completed his doctoral work at Oxford in 1925, and he arrived at Yale as professor of physiology in 1929, whereupon he established a pioneering laboratory dedicated to primate neurophysiology. Dusser de Barenne was already a well-known primate electrophysiologist, and Dean Winternitz of the Yale School of Medicine specifically wanted such expertise in New Haven. Winternitz visited Dusser de Barenne in Utrecht in the spring of 1929 to persuade him to come to Yale [6]. Shortly thereafter, Dusser de Barenne attended the Ninth International Conference on Psychology in New Haven September 1-7, 1929. There, he had the op-

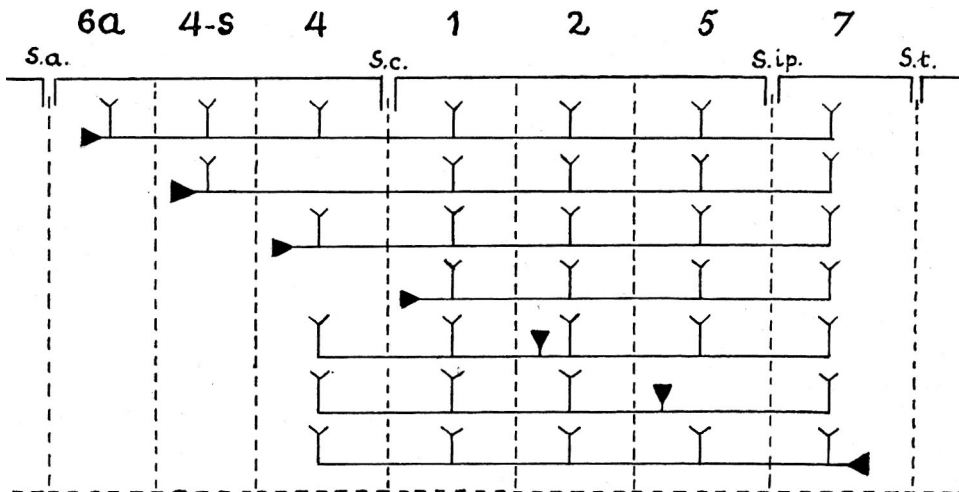
portunity to hear lectures by Pavlov and the American Psychological Association's Presidential Address from Karl Lashley, Hebb's future mentor [1], while sharing conversation and accommodations with other scholars in the Harkness dormitories [7]. This conference illustrated the urgent need for mechanistic investigations into the physiological basis for cognitive function, particularly in higher mammals. Shortly after this meeting, Dusser de Barenne decided to return to Yale to establish his laboratory. In September 1930, he arrived from Holland as the newly appointed Sterling Professor, and while his new laboratory in the Sterling Hall of Medicine was undergoing renovations, he set up at the Brady Laboratories with a variety of fruitful collaborators [6]. Fulton was instrumental in establishing this unique center for neurophysiology, and it was in this vibrant climate Dusser de Barenne continued his studies upon the connectivity between adjacent regions of the cortex, both in monkeys and apes.

Dusser de Barenne had established an *in vivo* method for inferring connectivity in anesthetized animals by a combination of localized application of strychnine to the surface of the cortex while simultaneously recording the extent and distribution of the ensuing "strychnine spikes" with multiple electrodes placed across the cortical surface. One of the first investigators to work with Dusser de Barenne in his new laboratory was the young neurologist Warren S. McCulloch. McCulloch arrived at Yale in 1934, and the two completed a series of seminal investigations over the next six years. Although Hebb himself cites a visual mechanisms chapter co-authored by McCulloch, Garol, and Von Bonin [8], he especially mentions the influential studies of Dusser de Barenne in his introductory chapter, and many of the ideas which McCulloch describes in his visual mechanisms chapter actually were first stated in his 1938 paper written with his mentor Dusser de Barenne in the new "Laboratory of Neurophysiology at the Yale University School of Medicine" [9]. In their series of experiments on the sensory cortex of macaques, McCulloch and

Dusser de Barenne firmly established that the connectivity of the cortex was simultaneously convergent and divergent but in no way "equipotential" in every direction. Additionally, they were among the earliest investigators to observe that the functional divisions of the sensory cortex were localized in a somatotopic and orderly manner. These observations were sharply opposed to the non-specific "mass action" ideas of cortical function espoused by Hebb's advisor Lashley [1], and also were in opposition to the dominant Pavlovian ideas at the time concerning irradiating waves and interference patterns traveling equally in all directions across the surface of the cortex.

A great deal of Hebb's initial (and lasting) appeal was his willingness to break from the prevailing psychological ideas of his advisor and others and instead conceptualize cognitive phenomena in terms of specific neural cells. The work done by McCulloch and Dusser de Barenne was critical in establishing the possibility of distantly, yet discretely, connected regions of cortex that would be essential to Hebb's conception of cell assemblies spanning across the cortex, yet containing discrete units of perception. McCulloch and Dusser de Barenne strongly suggest that the connectivity of the cortex is not random, but specifically convergent and divergent and unique for different cortical subdivisions. Their summary of feedforward and feedback connections within the sensory cortex of the macaque is summarized in Figure 2 [9]. In the final discussion of their experiments, they state:

Whatever the explanation of the strychnine-spikes may be, the fact that an area a "fires" an area b, whereas b does not "fire" a, must be an expression of directed functional relations, and, therefore, directed anatomical relations, between these two areas. Assuming that strychnine produces the spikes by acting on the perikarya of the cortex, the finding mentioned above must be interpreted to mean that nerve cell bodies in area a send their axons



**Figure 2.** A diagram of the directed functional relations between somatosensory cortical areas as revealed by electrophysiological recordings by Dusser de Barenne and McCulloch (1938). They demonstrate specific feedforward and feedback connections in the cortex.

into area b, whereas b has no cell bodies the axons of which extend into a.

Dusser de Barenne suffered a coronary seizure and died in 1940, just a few days after his 55th birthday. Following de Barenne's untimely death, McCulloch would continue these studies of functional relations between cortical areas. Such studies clearly informed Hebb's ideas of specific groups of neurons interconnected in distinct feedforward and feedback configurations (see Figure 7 in [15]), which would allow specific activation of unique cortical circuits to mediate and store unique information. The importance of these ideas of cortical connectivity also would influence the young McCulloch, and he subsequently developed the further abstract possibilities of interconnected neurons. In his seminal paper with Walter Pitts on the computational possibilities in "assemblies" of neurons [10], they formally demonstrate that different combinations of neurons could carry out different logical operations. McCulloch and Dusser de Barenne's early work convinced Hebb that coincident activation of interconnected circuits, when coupled with Lorente de Nó's ideas of reverberation, would lead to patterns of activity that could outlast the sensory stimulus and establish "associations"

among the relevant active neural sub-populations.

The major shortcoming of these early cortical connectivity diagrams described by Lorente de Nó, McCulloch, and Dusser de Barenne is that none of these investigators utilized their findings to propose mechanisms by which animals and humans could retain and manipulate information gained through experience. In this regard, Hebb's conception of activity in the cortex was more inclusive than any of the investigators reviewed thus far. However, two contemporaries of Hebb, Donald G. Marquis and his graduate student Ernest R. Hilgard in the Yale Department of Psychology, integrated both Lorente de Nó's ideas of reverberating activity in the cortex and Dusser de Barenne and McCulloch's ideas of connectivity with ideas about learning and memory in their aptly titled *Conditioning and Learning*, published in 1940 [11].

Hilgard had completed his doctoral research at Yale with Marquis and moved to Stanford in 1934. They spent the next six years collaborating and communicating by mail, exchanging completed chapters and revisions for their monograph. The majority of their book concerns a characterization of classical and instrumental conditioning at the behavioral level, painstakingly catalogu-



ing the most recent data on thresholds for conditioning and extinction, reaction times, and learning theory. Their last chapter (written by Marquis) is devoted to an analysis of the neurophysiological mechanisms possibly underlying conditioning and, in Hilgard's retrospective view, was "one of the better chapters" [12]. In the closing chapter, they first carefully examine the theory of cortical function proposed by Pavlov, which, as previously mentioned, involved waves of excitation and inhibition traveling over the cortex, with the exact temporal relationship depending upon the elicitory stimuli. In this manner, the same stimuli could serve conditioning or extinction, depending upon the exact time course of excitation vs. inhibition. Marquis's colleagues in the Laboratory of Neurophysiology had thrown this "irradiating wave" hypothesis into question, since their results clearly indicated this diffuse spread of activity was highly unlikely, as it seemed to be dependent upon the underlying connectivity of cortical regions and not on arcane surface wave phenomena and geometrical relationships.

Accordingly, Hilgard and Marquis suggest that conditioning may instead occur at the level of discrete groups of specifically interconnected cells. They discuss several ideas, but particularly emphasize two mechanisms: one in which the physico-chemical properties of nerve cells change upon co-activation, and another in which conditioning is instead mediated by continuous activity in closed neural chains. They are forthright about the lack of knowledge regarding the mechanisms of the first hypothesis, but form a remarkable concept based upon the current evidence. This is shown in their Figure 38 [11], and they state:

Excitation of a neuron is believed to be a process of breaking down (depolarizing) the membrane at the point of the synaptic connection. Among the changes suggested as necessary for learning are changes in one or the other of these properties: a concentration of some substance within the cell, a reorientation of molecules of the membrane, changes in the

colloidal dispersion within the cell, changes in the surface tension, changes in the interior-exterior potential gradient. There is no dependable knowledge relating these processes to learning.

The primary problem of neurophysiology of learning, however, is to specify the conditions of this change, whatever it may be ... Translated into the terms of Figure 38, his [Pavlov] conception would assume that simultaneous excitation by fiber a by two axons somehow produces a long-lasting decrease in threshold ... The essential basis of all such theories is the simultaneous activation of a neuron from two sources. Recent experimental work (Lorente de Nó, 1938b) has demonstrated that spatial summation of at least two simultaneous impulses is necessary to excite a neuron under any conditions. No explanation has yet been proposed why such excitation produces a permanent threshold change in some cases and not in others.

Hilgard and Marquis state here an essential tenet of Hebb's postulate: They suggest that co-activation of two inputs are necessary to produce a lasting change in the physiological characteristic of the post-synaptic neuron, which likely will have an effect on its signaling ability. They also integrated their ideas into the framework of classical conditioning, something Hebb himself (perhaps purposefully) neglects. In this manner, they formulate the simplest neural connection capable of explaining an observed behavior, i.e., the classically conditioned response, as indicated in the labels of Figure 38. Their diagram, coupled with their insights about plasticity and behavior, predates not only Hebb, but also tantalizingly suggests the basis for later work on classical conditioning and synaptic modifications in *Aplysia*, where similar simple circuits have been shown to mediate conditioning [13].

They next turn to the possibility of mediating conditioning without synaptic modifications, instead relying on summation of impulses through activity in closed chains of

neurons. In their Figure 39 [11], they propose the following:

The internuncial neurons are arranged in closed chains, in which a collateral excites a circle composed of several neurons. In the latter case the chain of neurons may maintain its activity indefinitely in the absence of peripheral afferent impulses. This arrangement suggests another possible mechanism of learning which would not necessarily involve any permanent alteration in the physico-chemical properties of the neurons. Closed chains, set into activity by the training procedure and continuing in the absence of any external excitation, would summate with otherwise inadequate afferent impulses to produce the conditioned response.

This concept takes the work on circuits done by Lorente de Nó and integrates it into a behavioral framework. The resulting hypothesis is remarkably close to modern ideas about certain cognitive phenomena, such as the physiological basis for working memory: Continued circulation of impulses in cortical circuits represents activation by past sensory experience, and this ongoing activity can then appropriately guide behavioral output [14]. In one short, concluding chapter, Hilgard and Marquis have synthesized the most current ideas about anatomy and physiology and framed them in the context of classical conditioning. They have formed key insights into the possible mechanisms for learning and memory — and all of this a full nine years before Hebb’s monumental work.

In his book, Hebb fully acknowledges the influence of all of the investigators discussed above [15]. He cites the influence and work of Dusser De Barrenne and McCulloch regarding functional connectivity in the cortex, and through their experiments he can envision the formation and permanence of unique cell assemblies. He also specifically mentions the influence of Hilgard and Marquis. In his “Chapter 5: Growth of the Assembly,” he reproduces a figure based

upon the work of Lorente de Nó. The ideas he proposes are a more sophisticated synthesis compared with those of Hilgard and Marquis, but they do not differ markedly in their propositions. Hebb essentially combines the physico-chemical modification hypothesis stated earlier by Hilgard and Marquis with reverberatory activity as proposed by Lorente de Nó to formulate his “dual trace” mechanism. In this manner, not only are assemblies of cells formed rapidly, but, importantly, their permanence and reactivation is made more likely because of activity-dependent changes in the entire assembly of participating neurons. Furthermore, he says of his mechanism of physiological modification that “some growth process or metabolic change takes place in one or both cells.” He further adds that this change occurs at the “synaptic knobs,” increasing the strength of the synapse. This is the novelty and essence of his postulate, which bears his name today.

Hebb envisions his cell assemblies taking hold in association cortex, and it is in these groups of cells, far downstream of sensory activation, that ongoing patterns of activity are of primary importance. The landmark paper describing medial temporal lobe resection and consequent memory impairment in patient H.M. would be published by Hebb’s Montreal Neurological Institute colleagues James Scoville and Brenda Milner some eight years after the publication of *The Organization of Behavior* [16]. Hebb himself makes no mention of the hippocampus, or its possibilities for associative phenomena, although such investigations of “Hebbian” mechanisms in this structure generate great interest today. This interest in “Hebbian” mechanisms of learning and memory in the hippocampal formation greatly increased following the discovery of long-term potentiation (LTP $\dagger$ ) and its proposed role as a cellular mechanism for activity-dependent modulation of synaptic efficacy [17]. This cellular correlate of Hebb’s postulate would be taken up by computational neuroscientists in the 1980s who could invoke this mechanism to modify synaptic weights, and thereby store information in network models [18].

Hebb's ideas have been influential in many areas of neuroscience and extend far beyond his attempt "to bridge the gap between neurophysiology and psychology" [15]. The significance of Hebb's *Organization of Behavior* lies in his postulate concerning synaptic modification, but he develops his ideas of the cell assembly much more fully than his proposed mechanism for cellular changes accompanying neural co-activation. Accordingly, there has been a recent shift in the appraisal of Hebb's work to reflect his important ideas about the cell assemblies in the cortex [1,2,3,4]. However, as illustrated here, ideas are rarely unitary events in the course of science and inevitably contain common themes and collaborative influences spanning many years. Hebb's legacy stems from his remarkable integration of the current psychological and physiological thinking into a cohesive yet sufficiently abstract framework.

Our review indicates that several of the ideas for which he is credited actually took shape in the intellectual climate at Yale in the 1930s and '40s and clearly were stated by these early investigators of cortical function. Fulton's first textbook of nervous system physiology facilitated widespread exposure to Lorente de Nó's views on cortical circuits and his creation of the first scholarly journal and a laboratory dedicated specifically to neurophysiology gave Dusser de Barenne and McCulloch a platform upon which they established the functional connectivity between regions of cortex. Following Hilgard and Marquis' initial attempt at a "physiological psychology," Hebb, drawing on the influences of all these researchers, transformed their ideas into a powerful framework describing how the cerebral cortex is capable of rapidly associating, dissociating, and storing circuit activity.

**Acknowledgment:** The author thanks Dr. Gordon Shepherd for his comments on the manuscript and for his enthusiasm and support of this project in all of its phases.

## REFERENCES

- Orbach J. The neuropsychological theories of Lashley and Hebb. Lanham: University Press of America; 1998.
- Brown RE, Milner PM. The legacy of Donald O. Hebb: more than the Hebb Synapse. *Nat Rev Neurosci.* 2002;4:1013-18.
- Cooper SJ. Donald O. Hebb's learning rule and synapse: a history and commentary. *Neurosci Biobehav Rev.* 2005;28(8):851-74.
- Harris KD. Neural signatures of cell assembly organization. *Nat Rev Neurosci.* 2005;6:399-407.
- Fulton JF. *Physiology of the Nervous System.* New York: Oxford University Press; 1938.
- Fulton JF. Obituary of J.G. Dusser de Barenne. *J Neurophysiol.* 1940;4:283-92.
- E.G.B. Ninth International Congress of Psychology. *Am Jour Psy.* 1929;41(4):684-6.
- Von Bonin G, Garol HW, McCulloch WS. The functional organization of the occipital lobe. In: *Visual Mechanisms.* Kluver H, editor. Biological Symposia. Volume 7. Lancaster, Pennsylvania: Jacques Cattell Press; 1942. pp. 165-192.
- Dusser de Barenne JG, McCulloch WS. Functional organization in the sensory cortex of the monkey. *J Neurophysiol.* 1938;1:69-85.
- McCulloch WS, Pitts W. Logical calculus of the ideas immanent in nervous activity. *Bull Math Biol.* 1943;52(1-2):99-115.
- Hilgard ER, Marquis DG. *Conditioning and Learning.* New York: D. Appleton-Century; 1940.
- Hilgard ER. ISI Current Contents Citation Classics. 18 April 1998;16:14.
- Bailey CH, Kandel ER, Si K. The persistence of long-term memory: a molecular approach to self-sustaining changes in learning-induced synaptic growth. *Neuron.* 1994;44(1):49-57.
- Goldman-Rakic PS. Cellular basis of working memory. *Neuron.* 1995;14(3):477-85.
- Hebb DO. *The Organization of Behavior.* New York: Wiley; 1949.
- Scoville WB, Milner B. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry.* 1957;20:11-21.
- Bliss T, Lømo T. Long lasting potentiation of synaptic transmission in the dentate area of the anesthetized rabbit following stimulation of the perforant path. *J Physiol.* 1973;232(991):331-56.
- Seung HS. Half a century of Hebb. *Nat Neurosci.* 2000;3:1166.