Title
A Neural Network Theory of Frontal Lobe Function

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 8(0)

Author
Levine, Daniel S.

Publication Date
1986

Peer reviewed

A NEURAL NETWORK THEORY OF FRONTAL LOBE FUNCTION

Daniel S. Levine
Department of Mathematics
University of Texas at Arlington
Arlington, TX 76019

INTRODUCTION

The frontal cortex is six-layered only in primates and is the neocortical area best connected to the hypothalamus. For these reasons and many others, Fuster (1980, p. 144) stated:

> "The central notion...is that the prefrontal cortex plays a role in the temporal structuring of behavior. The prefrontal cortex is thought to be essential for the synthesis of cognitive and motor acts into purposive sequences."

This article attempts to integrate this qualitative notion with existing neural network theories of motivation and cognition.

Grossberg (1975) discusses the striving for balance between two subsystems in a network. The attentional system seeks stable response to fluctuating sensory cues by focusing attention on important subclasses of cues. The arousal system enables adaptation to unexpected events and new reinforcement contingencies. Frontal lesions often change the balance between attention and arousal.

REVIEW OF BEHAVIORAL RESULTS

Delay Tasks, Perseveration, and Novelty

In *delayed response* (Jacobsen, 1935), an animal first sees food placed under one of two or more identical covers. After one-half to two minutes in which all covers are hidden, the animal must choose which cover to lift for food. Intact chimpanzees, monkeys, dogs, and cats perform this task easily, but frontally lesioned primates perform it poorly.

The delayed response deficit does *not* reflect short-term memory loss. Konorski and Lawicka (1964) found that most delayed response errors of frontally lesioned dogs involved perseveration of the response made on the previous trial, indicating that memory of cues had not been abolished. Interfering tasks between trials weakened perseveration.

Frontal monkeys also perform poorly on *delayed alternation* (Stamm, 1964) and *delayed matching to sample* (Spaet and Harlow, 1943). In delayed alternation, food is placed, concealed from an animal, alternately under the left and right of an identical pair of containers, and each time the animal must look again for food after a delay. Frontal monkeys tend to repeatedly choose one container that was once rewarding, even in the face of errors. In delayed matching to sample, the animal is first presented with a "sample" or visual stimulus, then after a delay is presented with a configuration of stimuli that includes the original one. The animal is then rewarded for choosing the sample correctly.

These results suggest that perseveration is a general consequence of frontal lesions. Milner (1964) confirmed this idea by asking frontal-lobe patients to sort cards based on any one of three criteria (color, shape, or number shown on the card). The patients were not told which criterion was correct, but at each placement were told whether their choice was right or wrong. Frontal patients initially deduced the correct strategy. However, when the experimenter changed the criterion, the patients preserved their earlier strategy. In the same vein, frontal patients asked to draw, in succession, a cross, two circles, and a triangle often draw four crosses instead (Luria and Homskaya, 1964).

In spite of perseverative tendencies, frontally damaged animals show increased preference for novel stimuli over familiar ones, regardless of reward value. Pribram (1961) gradually increased the number of objects. When a new object was introduced the peanut was placed under it. Frontal animals showed less tendency than normals to perseverate their choice of the object under which the peanut had been previously placed.

EEG Data

Walter (1964) and Walter et al. (1964) recorded a negative-going potential shift, the contingent negative variation (CNV), in humans anticipating a motor response. The CNV originates in the frontal lobes and spreads thence to other areas of neocortex. A similar potential change, also of frontal origin, accompanies a rhesus monkey's anticipation (Donchin et al., 1971).

In normal subjects, verbal instructions to await a visual or tactile signal cause enhancement of potentials the signal later evoked in the corresponding sensory cortex (Luria, 1969). Frontal patients, however, lack this potential change.

Dorsal Versus Ventral Frontal Cortex

The dorsal part of the frontal cortex has reciprocal connections with secondary sensory cortices. The ventral (or orbital) part has reciprocal connections (some via the mediodorsal thalamus) with the hypothalamus and limbic system. Hence:

> "...lesion studies indicate that the cortex of the dorsal and lateral prefrontal surface is primarily involved in cognitive aspects of behavior. The rest of the prefrontal cortex, medial and ventral appears to be mostly involved in affective and motivational functions...." (Fuster, 1980, p. 74).

Nevertheless, dorsal and ventral regions are extensively interconnected. This article will view these two areas as parts of a system, one part primarily motivational and the other cognitive, but both related to goal-directed behavior.

THE MODEL: MOTIVATIONAL ASPECTS

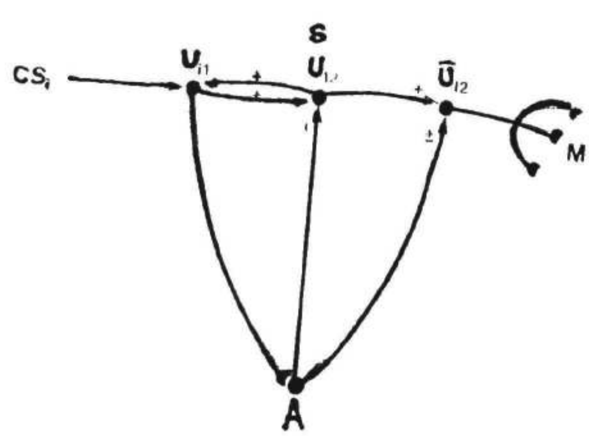
Both perseveration and enhanced novelty reaction can be seen as parts of deficit in drive-related incentive motivation. In other words, the frontal lobes serve to bias the organism's behavior toward actions that have current reward value, as opposed to actions that were once rewarding and have become motor habits, or actions that are exploratory in purpose. We shall now review how attention and arousal interact in some model neural networks.

Grossberg (1971) discussed the synchronization problem of classical conditioning, how the conditioned stimulus (CS) and an unconditioned stimulus (US) can become associated even when those stimuli are presented with different time lags on different trials. The solution of this problem involved "arousal" cell which include drive representations. Also, to permit secondary conditioning, it was found necessary to have two sensory representation stages for each stimulus.

Figure 1 reviews a general network, variants of which appear in Grossberg (1971, 1975, 1982) and Levine (1983). In Figure 1, the i th conditioned stimulus CS_i excites the cell population U_{i1} of its representation. Sensory representations are denoted generically by S. After receiving the CS_i input, U_{i1} sends signals to stage U_{i2} of the i th sensory representation and to all the arousal representations.

FIGURE 1

General Network for Classical or Operant Conditioning



Semicircles denote modifiable, arrows non-modifiable synapses. A is unconditionally activated by US, becomes activated by CS. Excitatory ("+") synapses from A to U_{i2} to U_{i1} lead to selective attention to stimuli conditioned to positive or negative arousal. Excitatory or inhibitory ("±") synapse from A to U_{i2} leads to enhancement of movement by stimuli conditioned to positive arousal and suppression of movement by stimuli conditioned to negative arousal. (Modified from Levine, 1983; reprinted with permission from Elsevier Science Publishing Company, Inc.).

The arousal representations (A in Figure 1) include, for example, A_h which subserve hunger and A_f which subserve fear. A given arousal population sends signals back to level U_2 of S only if it receives a large sensory input from level U_1 and a large drive input (such as hunger level or electric shock level). The synapse $U_{i1} \rightarrow A$ is always strong for an unconditioned stimulus, and is strengthened during learning for a conditioned stimulus.

Suppose that a hungry animal lifts a given cup, causing food to appear. Then A_h is excited and creates a positive $A_h \rightarrow S$ signal to all recently active sensory representations, such as those of the cup (S_c) and of proprioceptive feedback from the lifting response (S_{ir}). The U_{i2} stages of S_c and S_{ir}, having received U_{i1} and A_h inputs, can fire and send signals to M. The $A_h \rightarrow S_{ir}$ connection supplies positive incentive motivation for the motor act of lifting the cup.

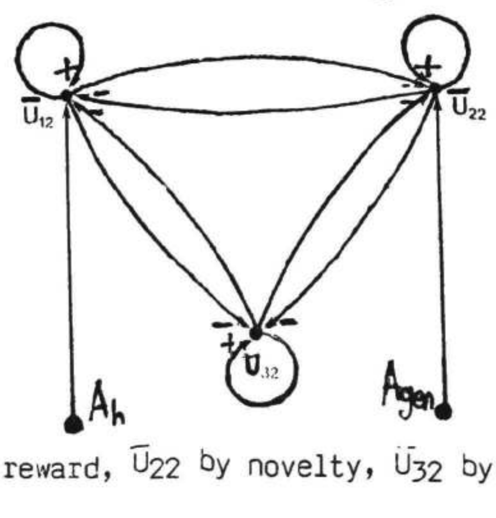
If lifting the cup leads to shock instead of food, then inhibitory $A_f \rightarrow S_{ir}$ connections supply negative incentive motivation for cup lifting, which can suppress $S_c \rightarrow M$ and $S_{ir} \rightarrow M$ firing. Negative incentive motivation can also come from frustration if expected reward is absent (Grossberg, 1975).

A CS conditioned to a given drive activates $A \rightarrow \bar{U}_{i2}$ positive incentive motivations, enabling signals from U_{i2} to M. Such signals influence $U_{i2} \rightarrow M$ synaptic habit strengths. Habit strengths are also influenced, less strongly, by repeated performance of a motor act even without current reward. Finally, habits can be influenced by the reward value of novelty (Berlyne, 1969). Response-contingent change (whether up or down) in light intensity in a rat's cage can reinforce bar pressing. Grossberg (1975) explained this effect using a nonspecific arousal locus (A_{gen}) that excites all the drive representations (A_j in Figure 1).

Figure 2 shows three stimulus representations, $\bar{U}_{12}$ which is excited by drive-specific incentive motivation because the stimulus is conditioned to that

FIGURE 2

Competition Between the Representations $\bar{U}_{i2}$ of Three Conditioned Stimuli



$\bar{U}_{12}$ is excited by reward, $\bar{U}_{22}$ by novelty, $\bar{U}_{32}$ by habit.

drive, $\bar{U}_{22}$ which is excited by nonspecific arousal because the stimulus is novel, and $\bar{U}_{32}$ which is initially active because the stimulus is one to which the animal has developed the habit of going. The representations of these objects are translated into target motor patterns via the $\bar{U}_{i2} \rightarrow M$ connection of Figure 1. Self-excitation and mutual inhibition between the $\bar{U}_{i2}$ creates competition for short-term storage.

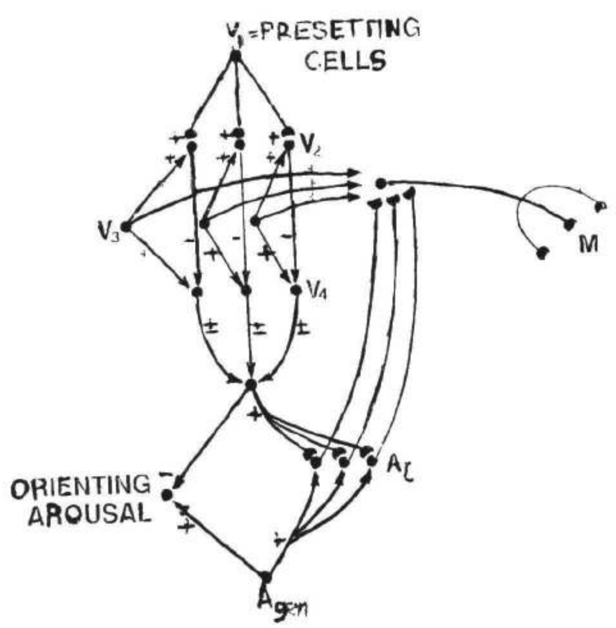
Figure 3 shows how unexpected events excite and expected events inhibit nonspecific arousal. The nonspecific arousal source is reminiscent of reticular areas, and the presetting cells, perhaps, of cerebellar areas.

My hypothesis that the frontal lobes are part of a major incentive motivational pathway is consistent with known anatomy. The A cells are reminiscent of drive-related areas of the hypothalamus. The frontal cortex is the only neocortical area known to have reciprocal monosynaptic connections with the lateral and preoptic hypothalamus (Nauta, 1971).

Thus a frontally lesioned animal can learn a response that leads to food reward, since some hunger-related incentive motivation still exists. Once that response has been established, however, even if reward ceases, perseveration occurs because incentive motivation for a competing response is weakened. Also, negative incentive motivation from frustration, which would normally occur when food is no longer found, is diminished.

FIGURE 3

A Network Where Expected but not Unexpected Patterns Inhibit Orienting Arousal



Activities of presetting cells represent a stored expected pattern. (Modified from Levine, 1983; reprinted with permission from Elsevier Science Publishing Company, Inc.)

If the new response involves a novel stimulus, perseveration is overcome by novelty. The approach to a novel object is stronger than the motor habit of approaching the familiar object. Also, the hunger-related incentive motivation exciting approach to the familiar object is weaker than in normal monkeys.

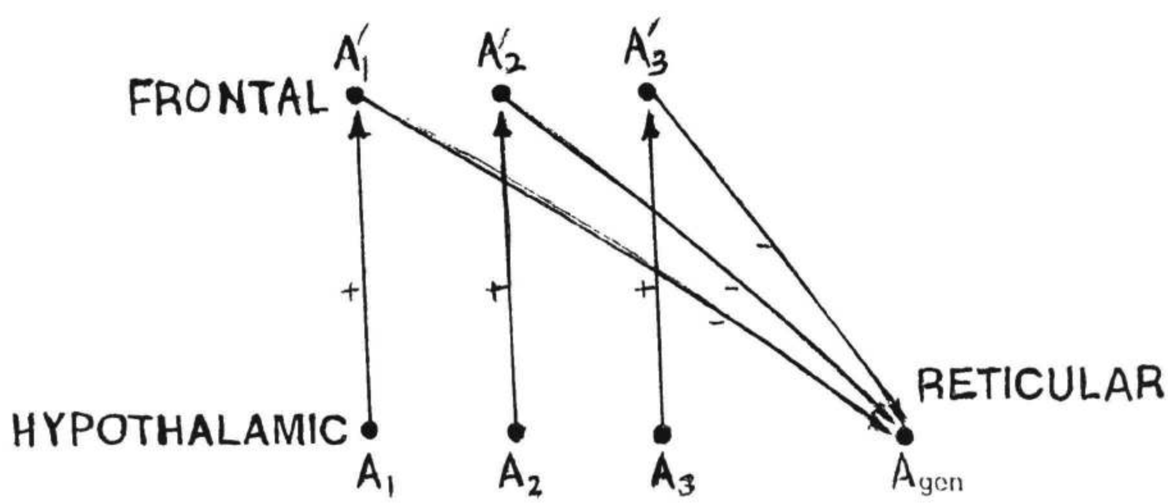
These results suggest a physiological prediction illustrated in Figure 4. Each "hypothalamic" A cell locus excites its own "frontal" representation A′, which inhibits the "reticular" nonspecific arousal locus. Frontally lesioned animals have trouble suppressing orienting reactions (Fuster, 1980, p. 61), which also supports this hypothesis.

THE MODEL: COGNITIVE ASPECTS

Weakening of incentive with frontal lesions is amplified by loss of neural preparation for expected sensory consequences of movement. Such preparation arises from the coding of sequences that include representations of stimuli, responses, and reinforcements in particular orders. It includes corollary discharge (Teuber, 1964), the compensation that the retina makes for eye movements.

FIGURE 4

Hypothesis for Frontal Participation in Incentive Motivation



The contingent negative variation accompanies expectation of one stimulus S_2 while another stimulus S_1 is present. This wave therefore depends on internal representations for S_1 and S_2 separately and for the temporal sequence S_1S_2. Further evidence that such sequences, and longer ones, are represented in frontal cortex is that frontally lesioned monkeys are easily distracted from sequential tasks (Grueninger and Pribram, 1969).

Grossberg (1978) discussed coding of "higher-order chunks" longer sequence representations in short-term memory models. He stated (p. 325) the following rule:

> "Self-Similar Coding Rule: Other things being equal, higher-order chunks have greater STM activity and longer duration than lower-order chunks."

This rule promotes goal-directed behavior, because longer stimulus sequences predict events better than shorter sequences.

The behavioral data suggest that the self-similar coding rule occurs particularly at the frontal cortex. Electrophysiological results from the dorsal (Fuster et al., 1982) and in the ventral frontal cortex (Rosenkilde et al., 1981). Both frontal areas in monkeys contain different types of cells whose activities change in correlation with each event in a delayed matching to sample sequence (sample/cue, choice stimuli, instrumental response, reinforcement). Some cue-sensitive cells respond to particular cue features such as color or location. Moreover, cells with similar properties may be organized into columns.

The dorsolateral frontal area known as the frontal eye field also shows variety in cell responses. Some cells in this area of monkeys discharge during but not before eye movement (Bizzi, 1968 and Bizzi and Schiller, 1970). Other frontal eye field cells fire before saccades, falling into three categories:

> "Visual activity occurred in response to visual stimuli whether or not the monkey made saccades. Movement activity preceded purposive saccades, even those made without visual targets. Anticipatory activity preceded even the cue to make a saccade if the monkey could reliably predict what saccade he had to make" (Bruce and Goldberg, 1985,p. 603).

The self-similar coding rule can best be understood by considering how short-term memory biases can develop in the selective coding of features. That issue was studied by Grossberg and Levine (1975). Their network was an on-center off-surround field, that is, each population excited itself and inhibited the others. The network (without biases) had been introduced by Grossberg (1973) to explain how noise can be suppressed and significant parts of a pattern contrast-enhanced. The activities x_i of the populations satisfied a system of differential equations of the form

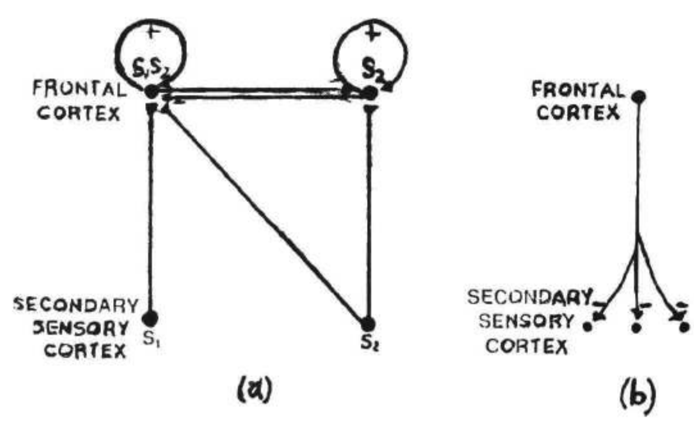
$$dx_i \,/\, dt = -Ax_i + (B_i - x_i)(f(x_i) + I_i) - x_i \sum_{k \neq i} f(x_k) \qquad (1)$$

where I_i are outside inputs, f is a monotone increasing function reflecting averaged neuronal input-output transformations, and B_i denotes maximum possible activity. If B_i is interpreted as number of cell sites, self-excitation of a population is proportional to number $B_i - x_i$ of "inactive sites," and inhibition of a population by the others is proportional to number x_i of "active sites."

Grossberg and Levine (1975) discussed how populations with larger B_i tend to suppress activities of populations with smaller B_i. Differences in B_i often arise from developmental and attentional biases. It is consistent that populations coding longer temporal sequences should have higher B_i values, perhaps reflecting more inputs from an earlier processing stage.

I conjecture that the dorsal frontal cortex contains on-center off-surround fields of populations coding chunks of all orders (see Figure 5a). Frontal afferents could also influence other on-center off-surround fields at the sensory cortices themselves (Figure 5b). The network of Grossberg (1973) has a <u>quenching threshold</u>, that is, an intensity below which stimuli are suppressed. Quenching threshold is lowered by tonic inhibition, leading to sharper decisions between stimuli. Hence I also predict that dorsal frontal cortex tonically inhibits secondary sensory cortex, thereby increasing the masking of irrelevant stimuli by relevant ones.

FIGURE 5



A) Mechanism for self-similar coding. Frontal representation for sequence S_1S_2 receives inputs from more populations than does representation for stimulus S_2 alone. Also, in frontal on-center off-surround field, larger B_i in equation (1) causes bias in favor of S_1S_2 (as represented by darker self-excitatory arrow). B) Tonic inhibition supplied to secondary sensory cortex by frontal cortex.

THE BEHAVIORAL PICTURE

The frontal cortex integrates sensory information from the neocortex with visceral information from the hypothalamus and limbic system. Hence frontal damage leads to "interoceptive agnosia" (Nauta, 1971,p. 182), including distractibility, lack of foresight, and inappropriate behavior. Frontal patients have been reported, for example, to urinate in public or tell off-color jokes at a funeral. Such behavior suggests disconnection between "reptilian" (instinctual), "old mammalian" (emotional), and "new mammalian" (rational) brains (MacLean, 1964).

Interfacing between the "three brains" seems to depend on orbito-dorsal connections within frontal cortex. Little is known about the structure of such connections. The flexibility of motivational responses and the known connections of dorsal frontal cortex with other neocortical areas and orbital frontal cortex with limbic areas suggest that orbito-dorsal connections should be nonspecific and modifiable in both directions. This should facilitate motivationally-based decisions between competing long-term plans.

REFERENCES

Berlyne, D. E. (1969). The reward-value of indifferent stimulation. In: Reinforcement and Behavior, J. T. Kapp, Editor, Academic Press, pp. .

Bizzi, E. (1968). Discharge of frontal eye field movements during saccadic and following eye movements in unanesthetized monkeys. Experimental Brain Research 6:69-80.

Bizzi, E. & Schiller, P. H. (1970). Single unit activity in the frontal eye fields of unanesthetized monkeys during eye and head movement. Experimental Brain Research 10:151-158.

Bruce, C. J. & Goldberg. M. E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. Journal of Neurophysiology 53:603-635.

Donchin, E., Otto, D., Gerbrandt, L. K. & Pribram, K. (1971). While a monkey waits: Electocortical events recorded during the fore-period of a reaction time study. Electroencephalography and Clinical Neurophysiology 31:115-127.

Fuster, J. M. (1980). The Prefrontal Cortex; Anatomy, Physiology, and Neuropsychology of the Frontal Lobe. Raven Press.

Fuster, J. M. (1985). The prefrontal cortex and temporal integration. In: Cerebral Cortex, E. G. Jones and A. Peters, Editors, Volume 4, Plenum Press, pp. 151-177,

Fuster, J. M., Bauer, R. H. & Jervey, J. P. (1982). Cellular discharge in the dorsolateral prefrontal cortex of the monkey during cognitive tasks. Experimental Neurology 77:679-694.

Goldberg, M. E. (1980). Cortical mechanisms in the visual instantiation of movement. Experimental Brain Research 41:A32-A33.

Grossberg, S. (1971). On the dynamics of operant conditioning. Journal of Theoretical Biology 33:225-255.

Grossberg, S. (1973). Contour enhancement, short-term memory and constancies in reverberating neural networks. Studies in Applied Mathematics 52:217-257.

Grossberg, S. (1975). A neural model of attention, reinforcement and discrimination learning. International Review of Neurobiology 18:263-327.

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. Progress in Theoretical Biology 5:233-374.

Grossberg, S. (1982). Processing of expected and unexpected events during conditioning and attention: A psychophysiological theory. Psychological Review 89:529-572.

Grossberg, S. & Levine, D. S. (1975). Some developmental and attentional biases in the contrast enhancement and short-term memory of recurrent neural networks. Journal of Theoretical Biology 3:341-380.

Grueninger, W. E. & Pribram, H. H. (1969). Effects of spatial and non-spatial distractors on performance latency of monkeys with frontal lesions. Journal of Comparative and Physiological Psychology 68:203-209.

Hirsch, H. V. B. & Spinelli, D. N. (1970). Visual experience modifies distribution of horizontally and vertically oriented receptive fields in cats. Science 168:869-871.

Jacobsen, C. F. (1935). Functions of the frontal association area in primates. Archives of Neurology and Psychiatry 33:558-569.

Konorski, J. & Lawicka, W. (1964). Analysis of errors of prefrontal animals on the delayed-response test. In: The Frontal Granular Cortex and Behavior, J. M. Warren and K. Akert, Editors, McGraw-Hill, pp. 271-294.

Levine, D. S. (1983). Neural population modeling and psychology: A review. Mathematical Biosciences 66:1-86.

Luria, A. R. (1969). The origin and cerebral organization of man's conscious action. Evening lecture at The XIX International Congress of Psychology, London. Moscow University Press.

Luria, A. R. & Homskaya, E. D. (1964). Disturbance in the regulative role of speech with frontal lobe lesions. In: The Frontal Granular Cortex and Behavior, J. M. Warren and K. Akert, Editors, McGraw-Hill, pp. 353-371.

MacLean, P. D. (1964). Man and his animal brains. Modern Medicine (February 2), pp. 95-106.

Milner, B. (1964). Some effects of frontal lobectomy in man. In: The Frontal Granular Cortex and Behavior, J. M. Warren and K. Akert, Editors, McGraw-Hill, pp. 313-334.

Nauta, W. J. H. (1971). The problem of the frontal lobe: A reinterpretation. Journal of Psychiatric Research 8:167-187.

Pribram, K. H. (1961). A further experimental analysis of the behavioral deficit that follows injury to the primate frontal cortex. Experimental Neurology 3:432-466.

Rosenkilde, C. E., Bauer, R. H. & Fuster, J. M. (1981). Single cell activity in ventral prefrontal cortex of behaving monkeys. Brain Research 209:375-394.

Spaet, T. & Harlow, H. F. (1943). Problem solution by monkeys following bilateral removal of the prefrontal areas. II. Delayed reaction problems involving use of the matching-to-sample method. Journal of Experimental Psychology 32:424-434.

Stamm, J. S. (1964). Retardation and facilitation in learning by stimulation of frontal cortex in monkeys. In: The Frontal Granular Cortex and Behavior, J. M. Warren and K. Akert, Editors, McGraw-Hill, pp. 102-125.

Teuber, H.-L. (1964). The riddle of frontal lobe function in man. In: The Frontal Granular Cortex and Behavior, J. M. Warren and K. Akert, Editors, McGraw-Hill, pp. 410-444.

Walter, W. G. (1964). Slow potential waves in the human brain associated with expectancy, attention and decision. Archiv fuer Psychiatrie und Nervenkrankheiten 206:309-322.

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C. & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensori-motor association and expectancy in the human brain. Nature 203:380-384.