

A Model of Attention and Memory Based on the Principle of the Dominant and the Comparator Function of the Hippocampus

V. I. Kryukov (Igumen Feofan)

UDC 004.032.26(06)

Translated from Zhurnal Vysshei Nervnoi Deyatel'nosti, Vol. 54, No. 1, pp. 11–31, January–February, 2004. Original article submitted October 2, 2002, accepted October 28, 2002.

The six major questions of attention are described in terms of the dominant, as defined by Ukhtomskii. The dominant was in turn simulated as a systems manifestation of phase transitions in the brain. The theoretical and experimental bases for the existence of metastable states in the brain are reviewed, these states having lifetimes of 1 sec and more. This approach simultaneously provides solutions for all the major questions of attention and the “central controller.” A neurobiological model of attention and memory is proposed, based on the systems properties of Ukhtomskii’s dominant and the comparator function of the hippocampus as described by Vinogradova. New published data are presented to support the existence of an information processing system in the brain in which the hippocampus plays the central role.

KEY WORDS: attention, the Ukhtomskii dominant, long-term memory, the oscillator model of neural networks, phase transitions in the brain, comparator function of the hippocampus, septo-hippocampal system, phase-frequency control system.

1. MAJOR PROBLEMS IN SIMULATING ATTENTION

Over the last few decades, the hippocampus has attracted particular attention as a structure which may harbour the brain’s greatest secret. There is no significant cognitive function of the brain which has not at one time or another been assigned to the hippocampus. We believe that the closest solution to this mystery was proposed by Vinogradova in her comparator hypothesis [2, 67, 68]. Our attempts to understand this hypothesis in terms of a systems model of attention based on the septo-hippocampal formation [4, 69] came up against a number of insoluble problems in simulating attention. The aim of the present work was to consider all these problems from the single point of view of the principle of the dominant as defined by Ukhtomskii and, after uniformly solving them, to return to the secret of the hippocampus. We will start with the question of what attention is and what the major problems in simulating it are.

More than a century ago, William James, one of the founders of cognitive psychology, wrote: “Everyone knows what attention is. It is the possession by the mind, in a clear and distinct form, of one of several simultaneously possible objects or a sequence of thoughts. Its essence is focusing, concentration of consciousness. It implies distraction from several subjects in order to deal more effectively with those remaining.”

We were unable to find any more contemporary and more precise definition of attention. We present only a list of functions usually ascribed to attention and provide a brief description of a number of problems with simulating attention. By attention we understand a psychophysiological process closely associated with memory processes and fulfilling the following functions: a) ignoring of several objects in order to allow more effective analysis of other objects; b) formation of short-term and long-term memory; c) performance of intermodal integration; d) control of overall excitation (arousal); e) inhibition of unnecessary or inessential ongoing activity to provide an easier or more effective response to new or vitally important stimuli.

The main problem in simulating attention is as follows: if the input of the simulation consists of a mixture of signals from different sources, it must be able to recognize

Svyato-Danilov Monastery, Moscow;
e-mail: kryukov@mail.cnt.ru.

and identify any component of the mixture. The other functions of attention listed above are also difficult to model, but are strongly dependent on this main problem.

We will illustrate the importance of the problem using an example. Studies of human vision have provided evidence that the brain processes visual information in different ways depending on the ongoing task. This processing may be parallel (pre-attention) or sequential (focused attention). In the former case, processing is rapid and independent of the number of objects in the visual field, and the global properties of the image are equivalent to its local properties. In the latter case, analysis is relatively slow, though the whole (the recognized signal) is “greater” than the sum of its parts, because the low-significance background is discarded and the accent is on analysis of the “figure,” which is the most important part of the image. The suggestion is that the operative and automatic combination of these two types of analysis also explains the whole mystery of human preeminence in recognizing images as compared even with the most powerful computers.

In what way is attention linked to perception, memory, and other cognitive functions? This question has always been central to both psychology and neuroscience. However, progress has been slow because of poor understanding of central attention control processes and the representation of memory in the brain.

Recent studies of the role of attention have used two types of simulation. The first type, which is termed the connectionist model, is based on theoretical concepts regarding neurons and their connections and computer simulations are widely used to determine whether parallel processing of the properties can in principle be linked with sequential attention to increase the efficiency with which the system operates. Known data on the physiology of neurons and their connections are generally significantly simplified to allow the use of simple mathematical or computer simulation. This type of model is less a measure of the real function of attention than a device to allow the application of useful arithmetic algorithms to solve technical problems related to signal identification. A good example of this type of model is provided by models described in the review from the Pushchino group of mathematicians [1]. Connectionist models have recently been used to address explanation of the interaction between attention and memory (see, for example, [24]). The other type of model is the psychological or psychophysiological model. This is used to allow a minimum number of theoretical assumptions and concepts to be involved in explaining the detailed influences of attention on perception, behavior, and short-term and long-term memory. The greatest internal concordance of models of this type is seen in [19] and [20]. The present review covers the results obtained with simulations of this type, though the neuronal realization of these models rests on advances in Russian science and, in particular, those of Ukhtomskii’s school and Vinogradova, which lead to a new approach to

the whole set of problems in studies of attention. In particular, we present a neurobiological model for assessment of the following questions of attention.

I. The question of stimulus selectivity. What is the mechanism of stimulus selectivity: why do some simultaneously presented stimuli attract attention and thus receive access to higher sensory processing, while others do not?

II. The question of long-term memory. What is the mechanism of reliable long-term storage of memory traces? How is long-term memory linked to short-term memory, and what is the role of attention in this?

III. The question of integration. The fact that parallel processing of stimuli occurs by way of extracting the properties of these stimuli at the early stages of processing leads to the questions of how and where the integral image is reconstructed.

IV. The question of inertia. There is an enormous gulf between the duration of the effects of attention, such as its attraction, and the duration of neuronal events, such as the interspike interval. This gulf is sometimes of several orders of magnitude. What is the physiological mechanism which slows time during the transition from neurons to behavior? What is the neuronal basis for retention of long-term attention to some stimuli even when the stimulus is presented for a relatively short period of time?

V. The questions of the inhibition and suppression of interference. What type of processing is applied to stimuli to which attention is not paid: are they inhibited actively or are they simply not subjected to further processing? What is the mechanism of inhibition of interfering stimuli?

VI. The question of the “central controller.” Is there a neuronal modality-non-specific structure, separate from memory, which coordinates the operation of all components of the processes of attention and memory, or is attention a consequence of the self-organization of the structures of the neocortex?

2. THE PRINCIPLE OF THE DOMINANT AND THE MAJOR QUESTIONS OF ATTENTION

The dominant in physiology, according to the Great Soviet Encyclopedia, is “the focus of excitation which temporarily determines the nature of the body’s responses to internal and external stimulation.” Ukhtomskii wrote in 1923: “My colleagues understand the term ‘dominant’ to mean a more or less stable focus of centers with increased excitability compared to what would otherwise be evoked, such that excitation newly arriving in the center serves to increase (reinforce) excitation in the focus, while the rest of the nervous system is subject to widely diffused manifestations of inhibition” [10, p. 7]. Tens of years later he wrote: “The dominant is not a theory, nor even a hypothesis, but represents an experimentally determined and very widely used principle, an empirical law, like the law of gravity,

which may in itself be uninteresting but is troublesome enough not to be ignored. I regard it as a ‘principle’ of operation of centers not because seems in any way very rational, but because it is a very constant feature of the activity of centers. In fact, the dominant can be established not so much as a rational feature of the operation of centers as merely a very stable feature of their operation. In any case, the dominant is one of the cryptic factors of our neural activity, and not an innocent one as appeared initially to be the case” [10, p. 125].

Of the extensive classical data on which the dominant is based, we will cite only those most tightly associated with the corresponding five questions described in section 1 of this review.

1. Neuromuscular preparations have an *optimal frequency* for stimulation, which corresponds to the maximum *lability* of nerve tissue and the maximal intensity of muscular *tetanus*. Further increases in the stimulation frequency can pass through a critical value after which muscle activity starts to decrease until the state of a *pessimum* or a state of absolute unexcitability is reached. This manifestation of an optimum and a pessimum can also be seen in relation to stimulus intensity [3].

2. Along with propagating spike excitation of the “all or nothing” type, the nervous system also has gradual *local stationary excitations*. These can be created both with normal stimulation and by local cooling and by local stimulation by electrical currents or various chemical substances. The resulting state of altered loci may be determined by the functional parameter *physiological lability*. This corresponds to the maximum stimulation frequency above which a transformation of rhythm takes place in nerve tissue, such as *frequency division* [3].

3. Apart from the optimum frequency and strength of stimulation, there is an optimum lability and an *optimum polarization* at which stationary excitation is most easily transformed into local rhythmic activity which can propagate beyond the boundaries of the initial locus [3].

4. At the optimum levels of lability and polarization, the phenomenon of *rhythm assimilation* arises, whereby rhythmic activity arising synchronously with the rhythm of the incoming stimulation persists for some period of time after the incoming stimulation ceases [3].

5. The transition from local stationary excitation to propagating excitation is seen during the acquisition or extinction of conditioned reflexes. Transmission of excitation to the cortex from the center of the signal to the center of the effector is accompanied by synchronicity of biopotentials in these centers, such that propagation of excitation becomes easier as the surface involved in synchronous coherent activity increases and as the coefficient of mutual correlation of “closed” centers increases. For some animals, such as rabbits, this synchronicity and coherence occurs at the frequency of the theta rhythm. Extinctive inhibition (“fading” of centers) is accompanied by divergence of the

frequencies of the linked centers, with drops in the synchronicity and synphasicity of slow biopotentials and decreased lability in these centers [6].

Now, considering the above data, it is easier to understand the *systems nature* of the principle of the dominant, which is as follows. At each point in time, the nervous system has only one active, dominating constellation (focus) of coordinately excited neuronal groups or centers characterized by a single rhythm and a single action (behavior). The same individual centers or groups of neurons can be included in different dominant constellations, whereby entry into one constellation or departure from it determines the ability of these groups or centers to assimilate a single tempo and rhythm of activity. Traces of the previous dominant persist over long periods in the higher parts of the nervous system and, in conditions of complete or partial recovery of the initial conditions, can completely or partially reappear.

On the other hand, the dominant is characterized by the following five *local properties* [10, p. 52].

- increased excitability: for entry of a group of stimuli into the dominant, the threshold of excitability of the dominant must be lower than the strength of the incoming stimulation;
- stability of excitation: for excitation to produce any marked behavioral effect, it must not undergo rapid changes over time;
- ability to sum excitation: the ability to accumulate excitation not only from specific, but also from non-specific stimulation;
- inertia, i.e., the ability to retain the state of excitation once the initial stimulus has ceased;
- “linked” inhibition, i.e., the ability to exclude from the dominant those centers whose activity is functionally incompatible with the activity of the dominant constellation.

In some of the very earliest publications on the dominant, Ukhtomskii wrote: “At the higher stages and in the cerebral cortex, the principle of the dominant is the physiological basis of the act of attention and subject thought” [10, p. 11]. Now, knowing the main systems and local properties of the dominant, we will briefly address the fundamental possibility of solving the questions of attention as described in section 1 of this review.

I. The question of selectivity is associated with increased excitability of the dominant for particular external stimuli and with the explanation of a concrete mechanism for selecting these stimuli. Considering that the dominant is a constellation of neuronal groups and centers with the properties described above, the following simple but important suggestion, based on large quantities of current data, becomes almost obvious: “Although selective attention is well grounded by behavioral experiments, it remained unclear what, in biological terms, is selected and how. We now suggest that selective attention includes selection of one of a whole series of different global cellular ensem-

bles”* [46]. With the aim of evaluating the significance of this result and, thus, assessing the strength of the principle of the dominant, we introduce a difficult question, which is often repeated in different versions, and appears to be insoluble without the principle of the dominant. “It has been demonstrated that the effect of attention acts on neurons in the primary zones, such that signals controlling attention in most studies run downwards from above, from the higher stages to the lower. However, how these signals return to the earlier stage of processing in order to select certain synaptic inputs from a multiplicity of other inputs remains deeply obscure” [17]. “Attention is known to affect the early stage of visual perception. The question is that of how prefrontal cells in the attention control system identify which cells of the visual system they should activate” [64].

The answer in terms of the dominant is in both cases simple: the signal at the higher level does not need to perform special detection and selection of cells at the lower level of the primary analysis, as according to the principle of the dominant they pass together and simultaneously to one and the same dominant constellation if they assimilate a single tempo and rhythm of activity. We will discuss below how this is realized in the simulation.

II. The question of long-term memory. In terms of the dominant, this is the question of a long-lived state of excitation, i.e., the question of the retention of traces of past dominants. This definition of the principle of the dominant, as we will see, opens wide new possibilities in studies of the question of memory and, in particular, the question of the connection of long-term and short-term memory. That this question is difficult is evidenced by many specialists, for example: “Further studies are needed to address the question of the link between perception, memory, and response. Several models of memory use different synchronicities to hold several concepts simultaneously in short-term memory. However, long-term memory for objects and events must use longer-lasting traces...” [62, p. 175]. The only candidate for the mechanism of long-term memory – *long-term potentiation* (LTP) – after extensive discussion [59] induced profound doubts in many investigators, albeit on the principle that LTP settles over a number of hours or days while long-term memory operates throughout the lifetime [55]. Unexpected support was obtained that “LTP is neither a memory trace nor a mechanism of memory formation, but is the beginning of the end of the synaptic theory of memory” [59, p. 621]. How serious this support is will be discussed in section 6 of the present review.

III. The question of integration, or, as it is often termed, the binding problem, from the point of view of the dominant is also tightly associated with the question of memory, and not only with the synchronization of the elementary properties of the object and the possible errors of

the relationship between these properties and foreign objects, as is usually understood in most cases. The fact is that each of the dominants experienced are recovered not only by means of the signal components or object characteristics, but also by all the cortical components of previous experiences of the same dominant. The dominant “can be experienced as a transient recollection with negligible inertia. And when this is without changes, a constant and uniform integral image, it again vanishes into the folds of memory. However, it can be recovered and reexperienced with almost the previous completeness and revival of the operation of the whole somatic constellation. It then again produces long-lasting inertia in the operation of the centers, and again *selects* the biologically relevant stimuli from the new environment, enriching the brain with new information. After this revived experience of the dominant, the corresponding image is again reprocessed and vanishes into the folds of memory, more or less *reintegrated*” [10, p. 36].

This makes it understandable that the binding problem cannot be solved in isolation from all the major questions of attention and memory. The authors of [44] came to this important conclusion, laying out six criteria for solving the binding problem. Similar conclusions have been made by other investigators addressing this problem. There are two further authoritative comments. “Current models provide no possible solutions for the coordination of the memory of episodes, attention, and working memory... This is an example of the fact that the central puzzle of studies of consciousness are considered simultaneously: the binding problem” [48]. “The cryptic process controlled by multimodal binding continues to pose a major problem” [45].

IV. The question of inertia is the central problem of the dominant, the question of understanding the mechanisms of the “excitation of large intervals,” i.e., neuronal activity of long duration, significantly greater than the time constant of the membrane potential of a nerve cell: “The dominant is based on the principle of inertia: the tendency to respond to stimulation always on to one side, the side of the ongoing ‘reflex situation.’ The Flywheel. The Aperiodicity of CNS Reactions to the Environment” [10, p. 247].

We will present some new data supporting the relevance of this question. Cortical neurons performing the task of short-term memory retain activity for several seconds after the stimulus ceases [28, 43, 65, 72]. Neurons in the human and primate prefrontal cortex develop the unusual ability to process information “on-line” – an ability which, as thought by many, is the basis of understanding, attention, and thought. These neurons have the ability to support the transient maintenance of information in attention for up to 20 sec. This fact is usually explained by some clearly intercalated slow chemical process [29]. However, this explanation is not always appropriate, as the size of the time delay is often dependent on the functional load of the test and on external conditions, for example, subsequent experience [61]. Our magnetoencephalograms of humans during a task

*The author was apparently not familiar with the principle of the dominant.

involving short-term maintenance of information in memory shows an unusual dynamic oscillation of activity in the range 4–12 Hz. This activity starts simultaneously with onset of the stimulus (so-called reset) and can continue for up to 600 msec. The duration of this activity *increases linearly* with increases in loading on the human memory, i.e., with increases from one to five objects held in attention.

There is no explanation for these data; however, in terms of the dominant, it does not present any particular difficulties. The question of inertia is intrinsically associated with the key concept of the dominant – physiological lability. A significant point in the teachings on the dominant is that lability can change, and it changes during the course of a response, i.e., under the influence of stimuli and loading. The duration of a response can change accordingly. Lability can be regarded as a measure of the inertia of the active state. High inertia and low lability mean a poor ability to terminate developing excitation. However, explanation of lability in neuronal terms leads to the same problem: that of what the source of high inertia could be, given that the neural network is presumed to consist of an aggregate of low-inertia elements. In addition, dynamic systems theory shows that there is a loss of stability of movement as the amount of inertia associated with the system increases. These two difficulties are discussed in more detail and are resolved in sections 3 and 5 respectively.

V. The question of the “central controller” is one of the most acute unsolved problems not only in psychology, but also in relation to the dominant. On the one hand: “Neural coordination or linkage of physiological sets in the order of their neural correlation is a necessary process and does not require the intervention of any additional “coordinating” center. Coordination in the sense of linking neural acts arises unavoidably from the impossibility of simultaneously performing the various necessary mechanisms in one relatively small defined executive apparatus” [10, p. 197]. On the other hand: “Why should a center dominate? It is the strongest. The most biologically important. Pleonasms (verbosity) in place of responses. There is no answer to “why?”” [10, p. 250]. We also show that the answer to “why?” does exist and that the “coordinating” center, the “central controller,” arises, or, more precisely, starts to operate, as a new systems property. The latter appears as arousal increases to some critical level at which phase transitions become possible in the form of spatial synchronization.

VI. The question of inhibition and suppression of interference is directly related to an important property of the dominant – “connected” inhibition. This connection can be illustrated by the following example. Some investigators [57], simulating the integration of features into a single object via synchronization of their corresponding oscillators, introduced desynchronizing inhibitory connections between some of the oscillators to interrupt the erroneous inhibition of different objects which only transiently overlap in space. This raises the question: can this concept gen-

eralize to the case in which features other than position are common to the different objects needing to be separated? It seems obvious that desynchronizing inhibition must be dynamic and functional rather than fixed, in order to consider different versions of identical features. In other words, each dominant is preferentially linked to only one characteristic inhibition of other dominants interfering with it. This is “connected” inhibition. In principle, it is entirely possible for this to occur in a dominant, because “excitation and inhibition are not opposite processes, but are related to a single process with opposite end effects depending on the conditions in which it occurs. The main and defining conditions for the effect is the degree of lability of the effector acting at a given point in time” [10, p. 124]. But how is this possible? Can it be understood using a simple physical analogy? The following section answers these questions in the affirmative.

3. PHASE TRANSITIONS IN THE BRAIN AND THE PRINCIPLE OF THE DOMINANT

The concept that inhibition arises from excitation, which is analogous to phase transitions in contrast with inhibition due to structurally fixed inhibitory substance, permeates all aspects of the teaching of Vvedenskii and Ukhtomskii. This teaching has thus far encountered difficulties and lack of comprehension because of the quite complex, poorly understood, and often paradoxical physical manifestations of the *phase transitions* which underlie it. Ukhtomskii proposed that further development of the physics of magnetism and the mathematical theory of non-linear oscillations would clarify many of these paradoxes. Along these lines, in 1940 he and Gulyaev proposed the first model of non-linear oscillations explaining *rhythm assimilation*, the *optimum and pessimum of the frequency and strength of stimulation*, the *effects of forced synchronization*, *inertial delay*, and *frequency division*, and other effects of the linked non-linear oscillators considered at the start of section 2 of this review. In the report, he posed the important question: “Is it impossible to ascribe the general features of behavior of physiological systems to variable lability and non-linear oscillatory systems because we have to address, first, inertial systems, and, second, the different manifestations of overcoming inertia” [11, p. 194]; he also answered the question in another work: “If the role of traces over time is considered as a magnet, with colloidal hysteresis, in a segment of nerve, then it must be considered that in a cell, a nerve cell, a cortical cell, transmission of traces from moment to moment must be the primary role” [11, p. 62].

In other words, inertia as a central property of a dominant, as all the other properties of a dominant, is analogous to the collective effects seen in ferromagnets, in which there is known to be a so-called critical slowing of the kinetics close to the phase transition point. Thus,

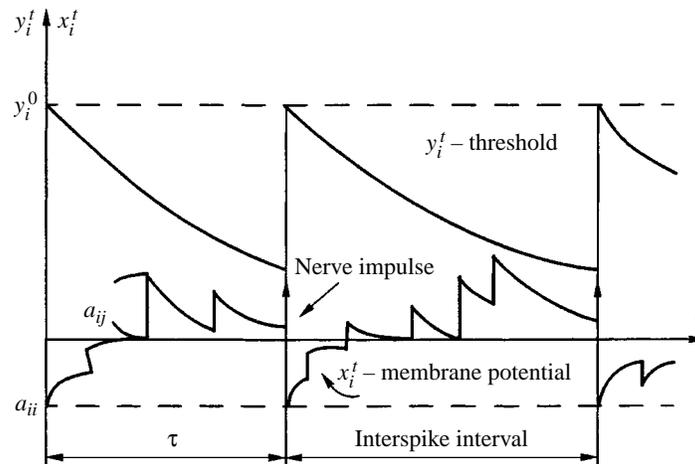


Fig. 1. Diagram showing the assumptions of the “basic neuronal model.” For explanation see text.

Ukhtomskii suggested the existence of different weighted phase transitions and metastability in the cerebral cortex: “The dominant can be regarded as a state of *unstable equilibrium* of molecules able to undergo catalytic (explosive) breakdown, while the usual reaction is an equilibrium of *reversible reactions*” [10, p. 241].

However, if the theory of phase transitions and metastability in the brain can provide consistent explanations for all the properties of the dominant, then, considering their tight connection with the questions of attention described above, it becomes possible to solve all these problems simultaneously, if phase transitions do in fact occur in the cerebral cortex. However, how would we demonstrate that phase transitions do actually occur in the brain?

The concept of non-equilibrium phase transitions in the brain has been developed by the well-known German scientist Haken in recent years [31, 32]. He believes that this concept is useful for brain theory. In particular, he repeated Ukhtomskii almost verbatim when writing that “the brain operates close to points of instability” [31, p. 2], and “we are not studying stable states, but transitions close to instability points, transitions which occur in the brain” [31, section VIII]. However, he was unable to demonstrate the existence of phase transitions in the brain, because his synergetic model of the brain was in essence phenomenological rather than microscopical, i.e., his model already contained an affirmative answer to the question posed, in the form of a special formulation of a differential equation proposed by him. The correct approach here was to describe the major elements of the model not with kinetic equations, but a simulated threshold neuron with quenching, accounting for the main experimental data on connections and the internal apparatus of a real neuron. It is the task of a strict mathematical theory to answer the question of the possibility or impossibility of phase transitions in a network con-

sisting of these “physiological” neurons. It is this approach that has been realized in our studies.

A relatively simple but quite physiological model of a network of non-formal neurons, termed a *basic neuronal model*, was laid out in [5] and [41]; the existence of phase transitions was strictly demonstrated in this model, within certain flexible formal limits. Generation of a nerve spike in this model (Fig. 1) occurs at the moment the threshold y_i^t and the membrane potential x_i^t (a stepwise process with exponential decay) are crossed, where the steps a_{ij} arises at the moment at which spikes arrive from neighboring neurons, where a_{ij} is the synaptic weighting of the connection from the i th to the j th neuron, satisfying the conditions of Ising’s potential (i.e., $a_{ij} = 0$, if neurons i and j are not neighbors); this potential is of the ferromagnetic type (i.e., $a_{ij} > 0$, if neurons i and j are neighbors). The model thus allows for the spike nature of neuron activity, exponential decay of the membrane potential, refractoriness, and interneuronal connections. Discussion of the existence of phase transitions in the nervous system is obviously virtually useless without consideration of these factors. Conversely, having demonstrated the existence of phase transitions in the “basic neuronal model,” it becomes possible to compare the consequences of this fact with real experiments and, ultimately, to approach the conclusion that phase transitions actually occur not only in the model, but also in the real brain. We will briefly outline the sequence of steps along this path.

The “basic neuronal model” is, mathematically, a system of interacting marker processes, obtained from a large number N of initially independent components connected by local connections of the physically paired potentials type to form a system. It follows from the details of the theory of such systems that the existence of a phase transition, when N tends to infinity, occurs when certain additional conditions for the existence of metastable states for the

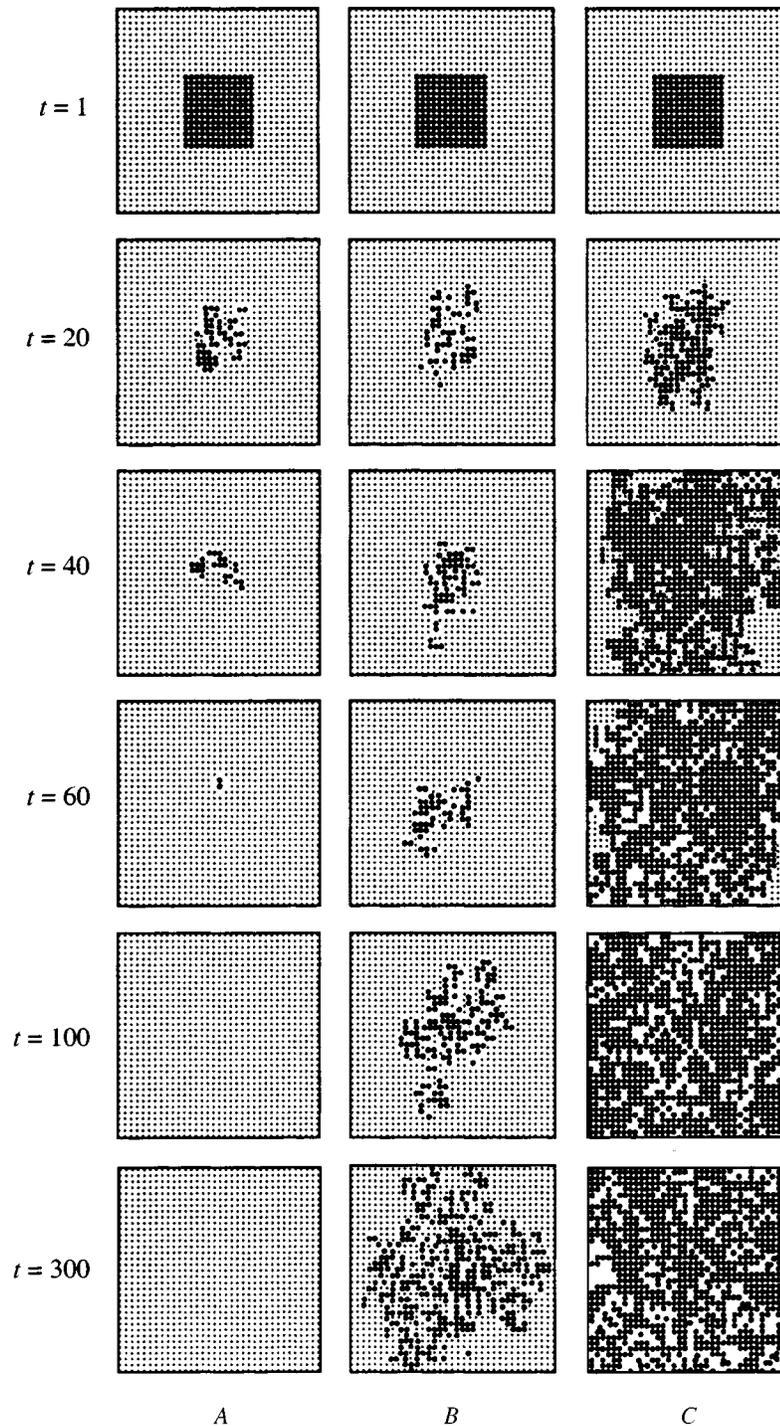


Fig. 2. The “spot persistence effect.” For explanation see text.

final N are fulfilled (metastable states – these are long-lived, but nonetheless unstable states of the system consisting of a large number of locally interacting components). Numerical experiments run on computers have supported this position. The novel effect of “spot retention” (Fig. 2)

was observed, in which the initial configuration of neuron activity in the form of a compact spot of defined size was rapidly compressed (case A – subcritical) or, conversely, expanded (case C – supracritical). However, in the case of the critical regime (case B), the lifetime was 200–300 units,

TABLE 1. Phase Transitions and Metastability in the Brain (adapted from [5])

Ferromagnetism	Neurophysiology
Distant ordering	Correlation of activity of distant neurons
Spontaneous magnetization	Spontaneous activity
Critical slowing	Inertia
Critical field	Critical depolarization level
Critical extension of autocorrelations	Extension of autocorrelations
Critical susceptibility	Increased susceptibility. Increased reactivity
Equilibrium phase transitions of type 1	Stepwise transitions between excitation and inhibition
Non-equilibrium phase transition. Hysteresis	Acquisition and extinction of conditioned reflexes. Hysteresis
Clusters, domains	Microfoci of excitation

which was two orders of magnitude greater than the characteristic relaxation time of the membrane potential, which in this case was 3 units. The fact that phase transitions occur in the “basic neuronal model” opened a path to developing an adequate dynamic theory of short-term neuronal memory as a critical manifestation occurring close to the phase transition point. This is because the “spot retention” effect is valid not only for the “basic neuronal model,” but is also general for a system of locally interacting oscillators. In the latter case, we can discuss the effect of “retention of a spot of synchronization.”

It was quite unexpected that a small modification of the physical theory of the intermediate field, presented in [5], would provide qualitative explanations for many experiments on different parts of the CNS which are very difficult to interpret, including all the basic properties of the dominant described by Ukhtomskii. We will list the *main results of studies* of the “basic neuronal model” [5]:

- the existence of a “distant order” with some limitations has been demonstrated mathematically. The existence of a metastable state has been suggested (“distant orders” in neuronal network indicates significant correlations between the activity of very distant neurons and, particularly, their synchronization with a zero phase shift despite the significant distance between them, as later observed experimentally [54]);
- the simulation model supported the predicted effect of metastability;
- a neuronal oscillator with unusual properties was proposed and investigated (for example, the unusually long period and the unusually small variability) on the basis of the metastability effect;
- the results obtained are useful for interpreting data from various brain structures which are difficult to understand, i.e., from the hippocampus, septum, cerebellum, and neocortex;
- it was concluded that the functional basis for the operation of the brain consists of metastable and unstable states, as the neuronal substrate of Ukhtomskii’s dominant.

Table 1, adapted from [5], shows the relationships between ferromagnetic data and neurophysiological data, and provides evidence for the appropriateness of the magnetic analogy for the brain. Attention is drawn to the similarity, sometimes identity, of the concepts and terms in such different fields (for example, spontaneous activity, susceptibility, hysteresis). We can then note significant differences in these effects which are difficult to explain, these seeming to be individual, independent of each other, and essentially different aspects of a single manifestation of phase transitions. Finally, the last row of Table 1 lays out a hypothesis presented to very many neurophysiologists and specialists working on simulation of the nervous system. In our view, this now appears to be an unavoidable consequence of the set of experimental and theoretical facts presented here.

The existence of phase transitions in the “basic neuronal model” provided for the first time [5]:

- an understanding of the inertia of the dominant as a critical slowing of the kinetics of microfoci;
- avoidance of the major defects in the reverberation hypothesis of short-term memory, such as the small period of recirculation and its low stability;
- presentation of a non-contradictory interpretation of the five local properties of the dominant focus;
- an explanation of several paradoxes associated with the dynamics of brain electrical activity during the acquisition and extinction of conditioned reflexes;
- identification of microfoci of excitation in neocortical columns as the basis of long-term memory;
- a solution for the main question regarding the cerebellum (the functional significance of its two main outputs) and, on this basis, proposal of a new model of the organization of movement as a system of phase autotuning of frequency, this implying the unification of perception and action, which needs the principle of the dominant, realized by the interaction of two identical phase-frequency regulatory systems.

The existence of long-term microfoci of excitation allows us to make the following new suggestion of the *hier-*

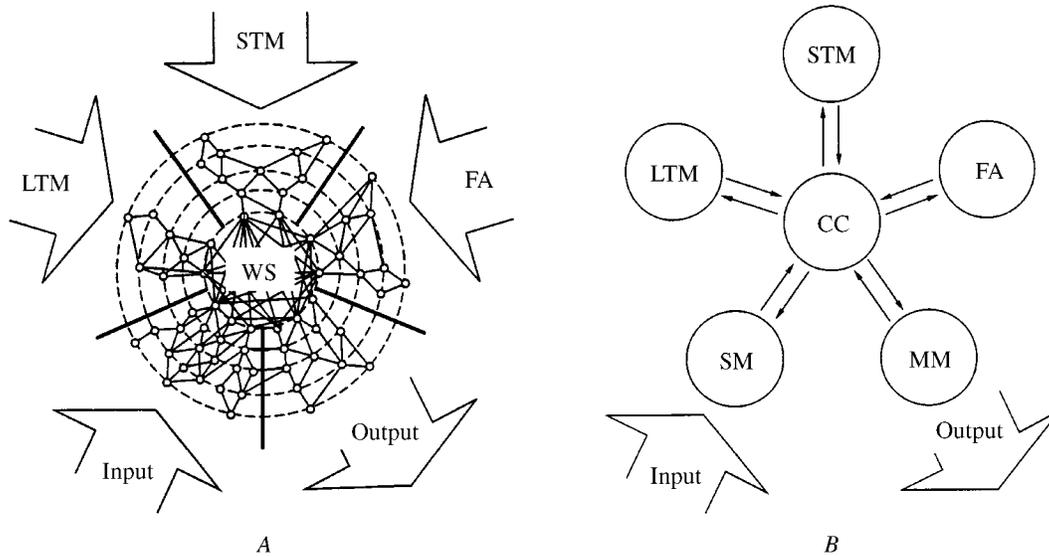


Fig. 3. Comparison of two types of information processing. A) Connectionist (adapted from [24]); B) the dominant. WS = working space; LTM = long-term memory; STM = short-term memory; FA = focus of attention; CC = "central controller"; SM = sensory memory; MM = motor memory.

archical type of relationship between long- and short-term memory. While short-term memory is identified with the activity of reverberators, which do not have to be synchronous and indeed are more often asynchronous, while long-term memory is associated with stationary local excitation in microfoci in the cortical layers, the transition from one type of memory to the other must be identified with the well studied phenomenon of transition of local stationary excitation to excitation propagating beyond the limits of the microfocus (see section 2, part 2). The Ukhtomskii dominant, as a constellation of microcolumns coordinately excited at the same frequency, is a means of "reviving" inactive long-term traces, providing for the easy and rapid selection of only those which were at some time "recorded" in an analogous situation (i.e., at the same frequency and phase of oscillations), and combining different modalities in terms of a single integral image, again transferring the result into the inactive form of a stationary local excitation protected from external and internal interferences. This suggestion relates to both the sensory system and the motor system.

However, this does not answer the question of the stability of the dominant, which was raised by Ukhtomskii [10, p. 264]; it merely clarifies. Now, taking long-lived microfoci of excitation into consideration, the question can be formulated as follows: how do we combine high reliability and stability of neuronal memory (traces of past dominants) with the fundamentally unstable regime of metastable states?

It is suggested below (section 5) that the answer to the question of the stability of the dominant is provided by the fact that the brain contains, along with long-lived microfoci of excitation, a *phase-frequency trace system* which syn-

chronizes these microfoci and induces within them the critical regime only for the period of time required for writing or reading of information from long-term memory. This answer to the question of stability can be likened to the dynamic stability of a moving bicycle: the initially unstable metastable state of synchronization can, in certain conditions, be turned by a source of flexible and economical CNS control, the *central controller*. This new concept in psychology was proposed in [15] and radically alters the whole essential architecture of information processing in the brain, as we have demonstrated in section 4. Its neuronal realization is given in section 5, and it is largely based on the "Vinogradova bicycle" idea (see Fig. 5, C).

4. THE DOMINANT ARCHITECTURE OF INFORMATION PROCESSING IN THE BRAIN

The human brain is a multifunctional organ mainly supporting the integrity of perception of the world, storing experience, and planning and performing actions. Processing of sensory information is not its main function. Nonetheless, we can regard the architecture of such processing in isolation from higher mental functions to clarify the close relationship between attention and memory in humans. (The terms to be encountered, such as "volitional control," "conscious action," etc., should be understood as a special way of performing a verbal task by the subject and not an attempt to provide a model explanation.)

With the aim of understanding the difference between the dominant architecture and the more traditional connec-

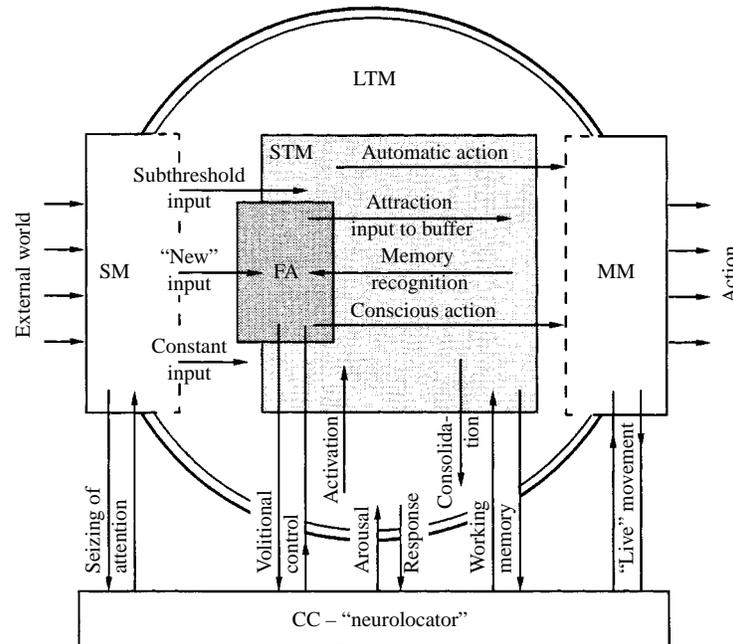


Fig. 4. The dominant architecture for sensory information processing in the brain. For further details see caption to Fig. 3. For explanation see text.

tionist architecture, we note the characteristic features of the latter.

1. The functional unit of information processing here is the neuron-like element, the formal neuron. Each element performs a simple binary operation on signals at its inputs.

2. All information is accumulated in the weightings of connections between elements, which are modified during the process of learning on the basis of local information in accordance with the so-called *Hebb rule*.

3. The typical connectionist architecture is a *parallel* three-layer network of these elements in which, apart from the input and output layers, there is an internal, "cryptic" layer. More developed versions of the architecture, which take greater cognizance of neurobiology, include up to four or more layers (Fig. 3, A).

4. Along with local layers, a so-called global working space is sometimes introduced. The latter has the role of a commutator of local spaces or local processes: long-term memory, short-term memory, and the focus of attention.

5. Attention corresponds to activation of one of the local processors, whose function is to enhance, like a *projector*, the illumination of some part of the global working space, strengthening the corresponding connections and weakening the connections from interfering layers. Information processing here is of the *parallel type*.

The dominant architecture differs from the connectionist in each of these five features.

1. The functional unit is not an individual neuron, but an elementary oscillator, i.e., a relatively small network of

"physiological" neurons, described above as the "basic neuronal model," in which an inhibitory interneuron is added to allow oscillations to occur by means of inhibitory feedback.

2. Information is "described" in a system of elementary oscillators, whose lability is modified during the process of learning in accord with the phenomenon of assimilation of the rhythm from a central oscillator. (In fact, this is coding in terms of isolability plots; the ribs of the plot connect oscillatory nodes with the corresponding frequencies, which initially differ by, say, 10–15%, but which are able to synchronize at a common frequency by assimilating the rhythm at this frequency.)

3. A simplified dominant architecture, as shown in Fig. 3, B, consists of a system of groups of independent peripheral oscillators and a single central oscillator with bilateral connections with each of the peripheral oscillators.

4. The central oscillator functions as a "central controller," a "conductor," which can alter its intrinsic frequency and produce different constellations of peripheral oscillators for long- and short-term memory and the focus of attention, as well as oscillators for sensory memory and motor memory.

5. The manifestation of attention corresponds to the appearance of synchronized and to some extent synphasic oscillations in the central and one or several peripheral oscillators working in *parallel*. Switching of attention from one group of oscillators to another *occurs sequentially*, by means of changes in the frequency of the central oscillator. Information processing is thus of the *parallel-sequential type*.

Overall, comparing this architecture with the connectionist architecture leads us to note some formal similarity, expressed as the identical composition of components and the ability to perform parallel processing. However, the difference between them is fundamental and qualitative, and it is as big as the difference between the wire telegraph and a radio connection. In the dominant architecture, all connections between components are functional rather than fixed. Therefore, all the known problems of connectionism are absent, such as the combinatorial break, the overload catastrophe, the connection catastrophe, and others [44]. Another important characteristic is that information is stored not in synaptic connections, but in the temporospatial configurations, such that learning is not localized but is centrally controlled [40].

Figure 4 helps to clarify the advantages of the dominant architecture; this illustrates both the properties of the dominant as described above and psychological, neuropsychological, and neurophysiological data. Firstly, the dominant has a *hierarchical* structure: the focus of attention enters short-term memory, and this in turn enters long-term memory: short-term memory is the activated part of long-term memory, and the focus of attention is the synchronous part of short-term memory. Second, the possible participation of the “central controller” in all the major transfers is clearly demonstrated, along with the fact that it is separate from memory. Third, all possible transfers from one memory type to the other under the influence of the “central controller” or external conditions are clearly indicated. Fourth, the compatibility and, to some extent, interchangeability, of the sensory and motor systems is evident – motor and sensory memory enter long-term memory, as discussed in section 3; more detailed consideration of this point will be presented at the end of the current section.

We will consider the interaction of the major components in more detail. The description presented below is largely in agreement with the accepted cognitive-psychological model [19, 20], but differs significantly in terms of its concrete neuronal realization. First we note that the arrows between components designate not only logical transfers of information, but also physical phase transitions as described in section 3. These are transitions between different states of a single neuronal environment performing *parallel-sequential* processing of the information arriving from the formation of a focus of attention before recording in long-term memory or before the movement reaction in motor memory. The intermediate, *buffer*, state in this process is the formation of short-term memory which is sometimes identified as so-called *working memory* which, like short-term memory, has limited information capacity and storage time. (Working memory means only a subcomponent of short-term memory, associated with regulation from the “central controller.”) the focus of attention includes that group of cortical oscillators which at a given moment is synchronized *in terms of frequency and phase* with the “central controller.” Cortical oscillators activated by exter-

nal influences not in the focus of attention or already lost from the focus of attention, for example as a result of *habituation*, form short-term memory. Switching of the focus of attention from one group of oscillators to another does not lead to rapid desynchronization of groups of synchronous oscillators which have already formed. These groups retain synchronicity for some period of time without regulation from the “central controller,” fixing the resulting associations and realizing *consolidation* of memory traces and thus mediating the transfer from short-term memory to long-term memory. This involves utilization of the phenomenon of *rhythm assimilation* as described in section 2 and in this case meaning that the oscillators which had operated for a quite long period of time in the synchronous regime gradually alter their intrinsic frequency to the frequency to which they were synchronized. This frequency or, more precisely, its corresponding *lability*, persists and, after extinction of oscillations in long-term memory, is reproduced on reading, when these oscillators again become active in short-term memory in response to the external stimulation. Thus, the configuration of activation of the “central controller” stored in long-term memory can be regenerated in short-term memory, from which it can pass to the focus of attention, for example during the processes of *remembering* and *recognition* of external signals. In addition, the presence of oscillators in the focus of attention may be induced by three external principles: the appearance of a “new”* *input*, replacing the existing input with a “new,” *volitional control* by the “central controller.” The latter ensures both single readings of information about individual facts from long-term memory – *semantic memory* – and time-interrupted reading by synchronization of different groups of oscillators in some temporal sequence – *episodic memory*. In the latter case, it is sufficient merely to reproduce the context and initial conditions present at learning for sequential displacement of associated representations to occur automatically, without the involvement of volitional control. Both of these require, albeit to different extents, the involvement of the focus of attention and the “central controller” and are therefore combined under the term *declarative memory*, i.e., memory whose content can be communicated to another.

We will now present new data providing significant support for this architecture.

1. Short-term memory and long-term memory have one and the same neuronal substrate in the neocortex. This substrate consists of a system of cortically disseminated, dispersed, overlapping, *hierarchically* organized networks of cortical neurons [26, 27].

2. A single given cell can be involved in different memory configurations, including different inputs and dif-

*“New” means not only a stimulus absent from both short-term and long-term memory, but generally any other stimulus adequately strong or biologically significant to be able to capture the focus of attention.

ferent modalities. In particular, such combinations occur in the perceptive and motor cortex [27].

3. Cognitive psychology data [13, 14] indicate that the “central controller” can perform the following operations:

- creation and manipulation of long-term memory;
- performance of intermodal binding of information from long-term memory and sensory systems, placing it in the short-term memory buffer;
- reading this information from the short-term memory buffer to the focus of attention (when consciousness is present);
- generation of a sequence of representations, extracting these from the buffer and long-term memory;
- coordination of the simultaneous performance of two separate tasks, including movement;
- inhibition of other interfering influences.

4. The hippocampus plays a critical role in the formation of long-term memory [30]. It is needed not only for recording information in long-term memory, but also for reading information from long-term memory [47].

5. Support is provided for the hypothesis that consolidation of memory traces is mediated via formations of the cortical cellular ensemble, which corresponds to the coherence of multiple neocortically disseminated memory traces, in which interruption of this consolidation leads to degradation of memory as seen in amnesiacs [50].

6. Support is provided for the hypothesis that the two different neuronal systems of motor and sensory memory have similar characteristics as regards coding and recovery of new information [16].

The latter point is entirely in agreement with the principle of the dominant, which requires not only a concordant rhythm, but also a concordant action. However, it will also be in agreement with the proposed architecture illustrated in Fig. 4, if we additionally propose that the “central controller” contains two phase-frequency systems operating in parallel and, evidently, tightly interacting: one for cognitive systems, the other for motor systems. This suggestion leads to a more general hypothesis for the possible interaction of *declarative* and *procedural* memory, which are usually placed in counterpoint to each other: declarative (consisting of episodic and semantic memory) requires attention, and its content can be communicated to another, while the latter is automatic and therefore does not require attention, but is not communicable to another. The hypothesis is that declarative memory, realized via hippocampal systems [25], cannot from the point of view of the theory of automatic control be fundamentally different from *procedural* (unconscious, automatic) memory, considering that the latter, like the former, is realized by a system of phase-frequency synchronization based on the olivo-cerebellar system which is functionally analogous to the septo-hippocampal system [5]. According to this hypothesis, which has already received partial support [42], both systems can operate at one and the same theta-rhythm frequency and can to some

extent substitute for each other. This may be the reason for surprising data showing that learning can occur without consciousness, the successful execution of unconscious movement activity, and data on blind vision and other puzzling phenomena which have yet to receive satisfactory explanations [23, 37].

5. THE “NEUROLOCATOR” – THE “CENTRAL CONTROLLER”

In many respects, the brain resembles a radar system, at both the behavioral and neuron levels. “Every time previous conditions required to produce a dominant are met, dominant responses will also arise” [10, p. 82]. Every time a dominant symptom complex is present, there is a corresponding behavioral vector. Simonov termed this relationship the “radar principle,” meaning the “selective readiness of the brain to make a response to a particular stimulus when it appears in the environment, the active search for this stimulus” [9, p. 37]. This analogy is followed in more detail in the classical psychological theory of the human operator [21], where the theory of automatic tracking systems is widely used with success. Similar concepts have also been advanced at the neuronal level on a number of previous occasions. Thus, in his theory of neuronal memory, John organized his radar analogy around the central node of the radar system – a *comparator* responsible for comparing the arrival times of the *interrogator* signal and the *response* signal from long-term memory (the coincidence detector hypothesis) [34]. John, *working from Ukhtomskii’s dominant* (!) put forward one of the most important concepts of the theory of neuronal memory, that an almost instantaneous search for a concrete memory trace can be made without scanning the whole memory. This will be the most probable temporospatial pattern, with the best correspondence to the animal’s ongoing external and internal environmental conditions, and will be the pattern most easily generating a response to the interrogation signal for a given symptom complex.

We will clarify that in this context memory, according to teachings on the dominant, means fixation of the topology of neuron coactivity as frequency and phase synchronization in different cortical and hippocampal areas at the theta rhythm [6]. Reading and recognition of memory traces correspond to the regeneration of former configurations of coactive neurons on recurrence of previous and similar conditions.

Studies described in [2, 67, 68] addressed in detail the application of the comparator hypothesis to the hippocampus and neocortex, and these results provided the basis for our model. In accordance with the comparator hypothesis of Vinogradova, the brain, in our interpretation, like a radar, sends an interrogation signal from its stem structures to all cortical structures of the brain and, after some delay,

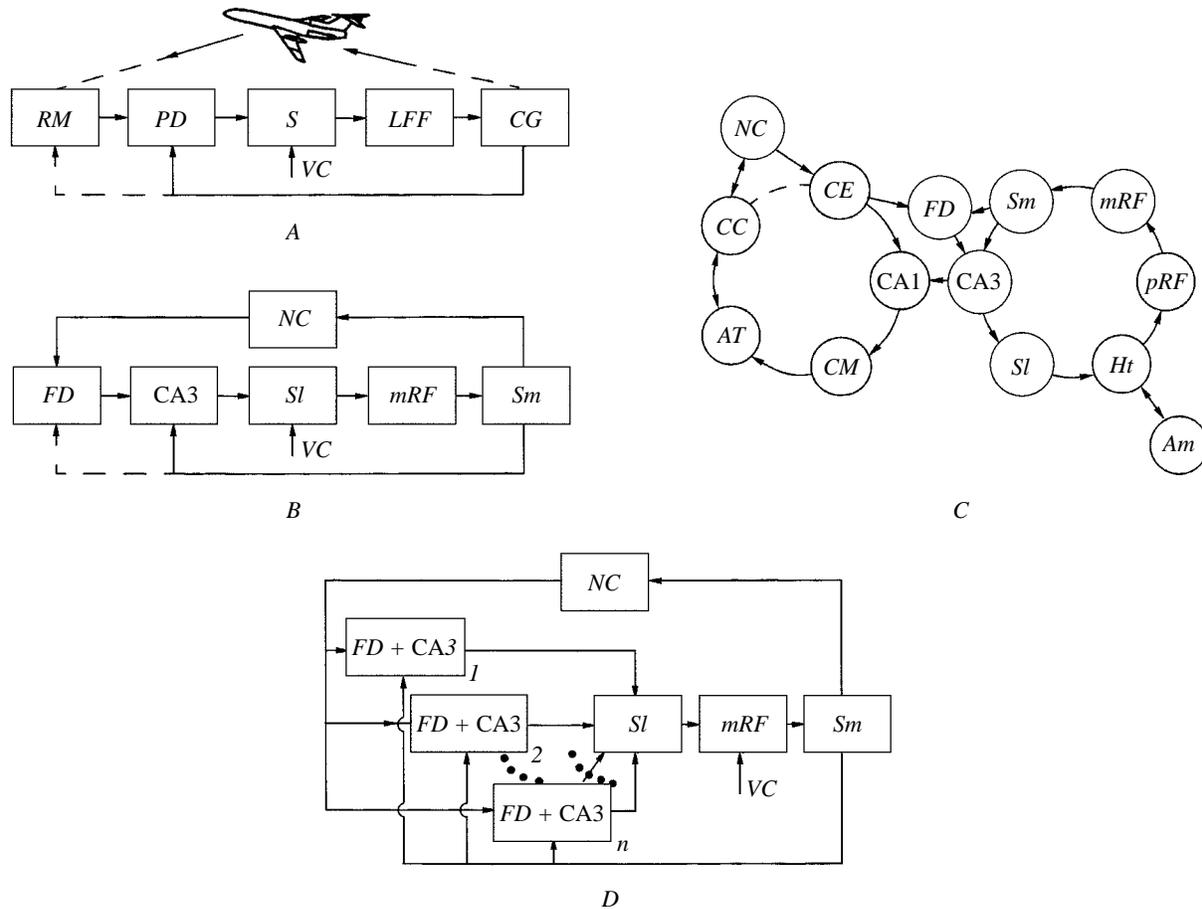


Fig. 5. Block diagram of a radar tracking system (A), the “Neurolocator” (B), the limbic system of the brain (C) and the “Neurolocator” allowing for the lamellar nature of the hippocampus (D). *CG* = controllable generator, *PD* = phase detector; *LFF* = low-frequency filter; *RM* = receiver-mixer; Σ = summator; *VC* = volitional control; *CE* = entorhinal cortex; *CM* = nuclei of the mammillary bodies; *AT* = anterior thalamus; *Ht* = hypothalamus; *Am* = amygdala; *CC* = cingulate cortex. For explanation see text.

receives a corresponding response signal at the input of the hippocampus, resembling a radar impulse reflected from the target. The comparator compares the response signal with endogenously generated reference signal consisting of the septal theta rhythm; when these coincide in time, i.e., when the two signals have the same frequency and phase, an error signal is generated which decreases the possible phase discordance between the reference and response signals.

A simplified block diagram of an actual radar tracking system is shown in Fig. 5, A. This consists of five standard nodes: a controllable generator, a phase detector, a low-frequency filter, a receiver-mixer, and a summator. Its neuronal “realization” is shown in Fig. 5, B. We rely on the well established functions of the various parts of the limbic system [2, 68]: the hippocampal dentate fascia (*FD*), which is the mixer and integrator of specific inputs; the medial septum (*Sm*), which is the central oscillator or synchronizer not only for hippocampal field CA3, but also for many other brain structures; and the lateral septum (*SI*), which is the output

mixer of field CA3 for individual lamellae, i.e., parallel structurally-functionally independent sections of the hippocampal formation [12], corresponding to various independent groups of cortical oscillators. All these structures are concordantly [2] connected to each other, forming two closed outlines, as shown in Fig. 5, C. This is the so-called Vinogradova bicycle. The first outline is informational; it includes hippocampal field CA1, the anterior thalamus (*AT*), the neocortex (*NC*), and other structures retaining, at least partially, signal specificity. This outline is active during “on-line” information processing, for example to ensure a prolonged delay or recirculation of signals for *working memory*. The second outline is the non-specific regulatory loop, including field CA3, the medial septum (*Sm*), and the lateral septum (*SI*), the mesencephalic reticular formation (*mRF*), and several other structures including the amygdala (*Am*), which is involved in emotional and volitional control. This loop provides non-specific activation of the brain and controls arousal. Given that the hippocampus is known to inhibit the activatory reticular for-

mation, the second outline is actually the *negative feedback loop* both for control of arousal in the brain and for control of the frequency of the septal oscillator. (This explains the “linked” inhibition of all signals not in the focus of attention – see below, property 6.) Both loops are important in our model, though in Fig. 5, *B* both are shown in very reduced form, especially the informational loop. Detailed consideration of the functions of all limbic system structures is not required here, so we present only a minimal version of the systems model shown shaded in Fig. 5, *C*, to provide a better explanation of the key role of the septohippocampal system in the operation of the whole brain. However, this now requires us to formulate the assumptions for a minimal model in neuronal terms and extract the main properties of this model as an automatic control system.

Assumptions of the “Neurolocator” Model

1. The “Neurolocator” systems model includes significant main subsystems (see Fig. 5, *B*): the septal oscillator (*Sm*); n independent groups of cortical oscillators (*NC*); the phase detector or comparator (*CA3*); the low-frequency filter (*mRF*); the input mixer (*FD*); and the output mixer (*SI*). Each subsystem is a small modification of the “basic neuronal model” described in section 3.

2. All cortical oscillators apart from the non-specific inputs from *mRF* have sensory inputs of defined modality, thus being analyzers of stimulus properties (shape, color, brightness, etc.). Their intrinsic oscillation frequencies are distributed uniformly in the alpha/theta range, and the non-specific inputs from *mRF* produce no oscillations in the absence of specific inputs, but only facilitate the appearance of these oscillations if the corresponding stimuli produce enough arousal.

3. The phase detector of field CA3 is a detector of coincidences between the two main inputs, the septal and the neocortical. Their maximum joint actions on field CA3 will be identified as the coefficient A , expressing the strength of the correlational relationship between these inputs, whose dependence on the time shift of these inputs is identified as $g(\tau)$.

4. The low-frequency filter *mRF* is characterized by a first-order transfer function $F(p) = K/(1 + Tp)$, where T is the time constant amounting to several hundred milliseconds – in accord with the explanation of inertia given in section 3.

5. The main assumption of the model is that attention can be described as a global property of the system whose block scheme is shown in Fig. 5, *B*, and this property corresponds to the phase synchronization of many cortical oscillators at the frequency of the septal oscillator.

6. When the number of synchronized cortical oscillators in an ensemble is greater than a critical value, the optimal labilities of these oscillators become equal, i.e., those oscillators which assimilate the general rhythm gradually, over some period of time, change their intrinsic frequencies. Memory of a stimulus which had evoked this type of syn-

chronization fades if the oscillators are immediately involved in other ensembles after the relevant one, without managing to extinguish these oscillations for transfer into long-term memory.

Studies in [41] demonstrated that the “Neurolocator” model is described by the following system of integral-differential equations:

$$\frac{d\varphi_i}{dt} = \Lambda_{0i} - \left[\sum_{j=1}^n A_{0j} g(\varphi_j) + N_j(t) \right] F(p), \quad (1)$$

$$(i = 1, \dots, n),$$

where φ_i is the difference in the phases of oscillations of the septal and i th cortical oscillators; Λ_{0i} shows their frequency tunings; A_{0i} shows the strengths of their correlational relationships; $g(\tau)$ is the non-linear phase discriminator function; $N_j(t)$ is the “white” noise of the i th lamella; n is the total number of lamellae; $F(p)$ is the transfer function of the low-frequency filter *mRF*. (The operator multiplication by $F(p)$ in the right side of Equation (1) identifies the convolution with the function $f(t)$, for which $F(p)$ is a Laplace representation, and the coefficients A_{0i} and A are linked by the

$$\text{relationship } A = \sum_{j=1}^n A_{0j}.)$$

This system is a generalization of the main equation of the standard phase-frequency tracking system based on several synchronized oscillators [7, p. 131]. This is used below as the basis for presenting the main properties of the “Neurolocator” as the known properties of a standard first-order phase tracking system, i.e., for the case in which $F(p) = K$. This latter suggestion was made only for simplicity and will be excluded at the end of the present section.

Properties of the “Neurolocator” Model

1. Attention is impossible both in conditions of low arousal ($AK < \Lambda_0$) and when arousal is too high ($AK \gg \Lambda_0$).

2. Attention arises stepwise as arousal increases ($AK > \Lambda_0$) and appears as a “choice” of one of several external or internal signals which at the ongoing time point “seizes” control of the “Neurolocator.”

3. Attention is controlled by changes in the tuning (Λ_0) both automatically and by volitional effort. There is a regime in which the i th lamella is functionally automatically disconnected from the system (for example, automatically when $A_{0i}K$ is high).

4. Attention is unitary, though with sufficiently high arousal AK , it can be divided in accord with Miller’s rule (7 ± 2) or Couen’s rule (4 ± 1). In other words, the system can simultaneously follow the frequency and phase of 4–7 initially independent cortical oscillators simultaneously being controlled by the septal oscillator.

5. Disruption of attention can be sudden (due to noises or synchronization from another group of cortical oscilla-

tors) or gradual. In the latter case, there is an interesting non-linear “slipping cycle” effect.

6. “Linked” inhibition of all stimuli not falling within the focus of attention occurs via two mechanisms: firstly, because of desynchronization, and secondly, because of decreases in arousal, which, according to assumption 2, is only effective for signals falling in the focus of attention.

7. Attention is possible not only when the septal and cortical oscillators have identical frequencies (1:1 synchronization), but also when their frequencies are different ($p:q$ synchronization). When $p = 1$, the most stable synchronization for the case $g(\varphi) = \sin(\varphi)$ will be at $q = 4$ (for example, synchronization of oscillations with frequencies of 10 and 40 Hz is not only possible, but may be predominant).

Studies reported in [40] showed that the “Neurolocator” can explain the following set of classical psychophysiological effects:

- focused, shared, and controlled attention;
 - the Yerkes–Dodson rule for optimum arousal;
 - binding or multimodal synthesis of an integral image;
 - the stroop effect or the effects of intra- and inter-modal interference;
 - instantaneous learning;
 - the multifunctionality of the hippocampus;
 - acclimation and operation of long-term potentiation in the hippocampus;
 - generation of temporal sequences of configurations.
- Recent studies have yielded significant quantities of new data which can be considered as supporting the “Neurolocator” model. Only a selection will be referenced here:
- the hippocampus is a critical component for the integration (binding) of multiple input signals [18, 66];
 - the existence of “linked” inhibition of the dominant has been supported: attention to one object leads to suppression of the activity of all other interfering objects [33, 49, 60];
 - the theta rhythm, which is difficult to observe in humans in normal conditions, appears on performance of cognitive tests [35, 56];
 - the EEG alpha rhythm and the theta rhythm reflect memory and attention functions in humans [38], the transition from theta to alpha synchronization arising in conditions of maximal loading of episodic memory [39];
 - memory is located not in the hippocampus, but in the neocortex; the hippocampus plays a key role in organizing memory in humans [25, 45, 51, 63, 64];
 - the hippocampus is always required both for the recording and consolidation, and for the reproduction of different types of memory, including spatial, declarative, and episodic memory [47, 53];
 - the gradient of retrograde amnesia [51] provides evidence for the involvement of the hippocampus not only in the recording, but also in reading of sensory information;
 - humans can simultaneously track several independently moving targets; the imprecision of tracking and the

reaction time increase sharply when the number of targets is more than five [52] (cf. property 4).

A more complete listing of supportive evidence, along with new data inexplicable in terms of existing models of attention but fitting with the “Neurolocator” model, will be presented in a future article.

All the data discussed above provide evidence for the correctness of the *basic predictions of the “Neurolocator” model*, as formulated in [40].

The CNS contains a phase-frequency tracking system of at least the second order, controlling attention and memory in humans and functioning in the range of frequencies of the alpha and theta rhythms. The septohippocampal formation plays a leading role in this system as the driving oscillator and phase discriminator. This system, interacting with other cortical and subcortical structures (for example, *mRF*), performs the functions of a “central controller.”

It remains to be clarified that a first-order tracking system whose properties are given above becomes, with inclusion of the inertial filter $F(p) = K/(1 + Tp)$, a second-order system, acquiring a number of interesting new properties. For example, theory and experiment show that the operation of such a system is not in any way disturbed when the signal from the target being tracked disappears behind an obstacle at a time of the order T . However, this very effect was recently observed experimentally in studies on the visual tracking of a moving target with obstructions [70]. The mean duration of the “safety” of a target behind an obstacle was 300 msec [58], which is in good agreement with the predictions of the model described above. This fact, particularly as observed during the simultaneous tracking of several targets, provides convincing evidence of the correctness of the main predictions of the “Neurolocator” model and, thus, of the key role of the septohippocampal system in information processing in the brain.

6. CONCLUSIONS

The simulation and experimental data presented here lead us to the view that there is a simultaneous solution to all six problems involved in simulating attention described in section 1 of the present review. This solution is implicitly present in Ukhtomskii’s principle of the dominant and Vinogradova’s comparator hypothesis. All that remains to be done is to describe the physical mechanism underlying the dominant, to link it to data on the hippocampus, and to apply the theory of automatic control tracking systems. Having done this, we will summarise our results as follows.

I. The Question of Stimulus Selectivity

The “Neurolocator” selects groups of cortical oscillators encoding the individual properties of an object or the objects themselves, on the basis of the ability of these oscillators to assimilate the common rhythm using phase-frequency synchronization at the frequencies of the theta and alpha rhythms

and their harmonics. The selection process itself is reminiscent of a competition between several groups of oscillators for seizing control of the central oscillator. Once control has been seized, cooperation occurs in the victorious group, reinforcing this group by recruitment (by frequency coupling) of many other oscillators, enriching the memory of new associations. The victor may be not one, but several in number (see section 5, properties 4 and 7). Our model can explain the selection and automatic simultaneous linking of up to 4–7 independently moving objects. Those stimuli or objects which cannot evoke sufficient arousal and the septal theta rhythm, even if they are completely novel and unfamiliar, will not attract attention and will not receive access to higher sensory processing. It is of note that excessively attractive objects, inducing significant synchronization, are also likely automatically to lack attention to protect the brain from overload and epileptiform pathology.

II. The Question of Long-Term Memory

The high reliability of long-term memory results from the fact that the functional element of our model of attention is not a single neuron but a small ensemble, a local group of neurons, which can reproduce the cooperative effect of local stationary excitation (the “spot persistence” effect), and can increase the duration of the period and stability of oscillations of local oscillators by an order of magnitude. The actions of external stimulation and non-specific activation from the reticular formation have the result that this local activity becomes oscillatory activity propagating across the network. This article is based on the view that the transition from synchronization to localization is the mechanism of the phase transition from short-term memory to long-term memory and vice versa. In this regard, it is relevant that long-term memory is not a property of synaptic contacts, but is a property of nervous tissue as a whole, which has virtually unlimited capacity for the long-term constant maintenance of its optimum lability as established by the previous dominant. Attention is required both for writing to long-term memory and for reading from long-term memory, as integration of contrasting traces and object properties is needed in both cases.

III. The Question of Integration

The question of the integration of the properties of an object in a single entire image is resolved in our model by phase-frequency synchronization of a large number of spatially separated cortical oscillators, using attention and the “central controller.” The suggestion that this integration could occur at the pre-attention level is refuted by psychological experiments [71]. Thus, the binding problem, even in its simplified formulation relating to the correct relationship between the properties and the object, cannot be resolved without attention and memory. Integration itself of the properties and reconstruction of objects occurs at the focus of attention both at the late and the early stages of information processing.

IV. The Question of Inertia

This is solved in our model by the automatic appearance within the components of the model of a critical regime during the active perception of external stimuli. The individual subsystems of oscillators and the whole system of attention in general operate close to the point of unstable equilibrium, which is a close analog of the critical point of the physical manifestations of phase transitions and metastability. This latter is accompanied by the so-called critical slowing of dynamics, which fills the temporal gap between the psychological and neurological time scales. Long-lived corticohippocampal reverberations support the state needed for resolving many vital problems “on-line” by using working memory.

V. The Question of Inhibition and Suppression of Interference

Attempts to understand Ukhtomskii’s concept that “linked” inhibition is a dynamic process rather than the effect of fixed inhibitory connections, led us via phase transitions to the theory of critical inertial synchronization-desynchronization as the main property of the dominant. It is not surprising that Ukhtomskii termed “linked” inhibition “completely half” of the principle of the dominant [10, p. 65]. It is so important for understanding attention that discussion of its mechanisms continue to be of interest (see review [36]). Initial solutions to the problems identified in [40] are now more widely understood, and they provide for desynchronization and simultaneous control of non-specific excitation. Recent studies using new scanning techniques completely support this solution: inhibition of interfering objects occurs not only as a local process of desynchronization, but also as a global process involving a change in the null level of the background activity (i.e., decreases in arousal) of all items not present in the focus of attention [33, 49, 60]. This double mechanism protects attention on the selected object even when an interfering object and the target are superimposed one on the other. We believe that this closes the discussion on the mechanism of “linked” inhibition.

VI. The Question of the “Central Controller”

The CNS contains a neuronal modal-non-specific structure, separate from memory, which coordinates the operation of all components of the processes of attention and memory. Functions which can be performed by the “central controller” are very wide-ranging. They include: selection of individual objects, selection of groups of moving objects, extraction of “novelty,” binding, spatial frequency-phase encoding, reading, consolidation, generation of sequences and different types of memory (declarative, episodic, semantic, working, short-term, long-term, instantaneous). Many of these functions have recently been reviewed [1]. The “Neurolocator” model solves the problem of simulating these functions, *all of them simultaneously in time and place*. The key concept in this, apart from the dominant, is the application of the theory

of phasic frequency autotuning. The “heart” of this system is the septohippocampal structure, which has been studied in detail by Vinogradova. In our view, this approach can pull the theory of long-term neuronal memory out of the blind alley described in section 1, while the theory of working memory provides a flexible and universal mechanism for maintaining multimodal information “on-line” for resolving ongoing cognitive and motor tasks.

In conclusion, we note the current state of theories of the brain. In 1979, Crick [22, p. 188] predicted: “...if there is actually a hiatus in brain studies, it will probably be at the level of the overall control of the system. If the system were as chaotic as it is sometimes presented, it would be unable to perform even simple tasks satisfactorily. A possible but unlikely example would be the discovery that brain processing depends on the phasic control of some kind of *periodic system*, like a computer, and this would probably be the main hiatus.”

In 1995, von der Malsburg predicted [44, pp. 523]: “We are in the middle of a scientific revolution, which will establish binding as a fundamental aspect of the neuronal code, alongside the classical frequency code. With luck, this revolution will lead to a satisfactory paradigm, improving on the current focus on the single cell, advancing towards more internally consistent concepts of brain function and perception and to a solution of many remaining problems in terms of clear models.” This prediction already identified the characteristics of the future hiatus as a general features of science, as per the need for unavoidable scientific revolutions to overcome stagnant situations and obsolete paradigms as previously noted by Kuhn (1962).

Another major brain researcher, Rose, wrote recently [8, p. 326]: “According to the concepts of Imre Lakatos, the gradual advancement of science is based on ‘study programs.’ Programs may be progressive, and in this case progress occurs via the posing of tasks which can be resolved within its framework, remaining fertile in relation to understanding the world, or it can be degenerative, when it will produce more significant anomalies. In this sense the associative program of Hebb is progressive, as it can include Kandel’s ‘cellular alphabet of memory’ and *long-term potentiation* as a model mechanism, while the experiments performed by Harsy and myself on brain damage are anomalies of the type which may demonstrate the degenerative nature of the Hebbian program; in other words, we need a revolution, if we are to use Kuhn’s terms.” We note that long-term potentiation as discussed here is, according to a review of extensive experimental studies [55, 59] performed by many experimenters, unsuitable as a model mechanism of learning; the “Hebb model is inadequate even for the relatively simple memory of a chicken and does not even address the much more complex memories of humans” [8, p. 354].

Which of these is right? Are we in the middle of a scientific revolution or merely at its beginning? Given that brain theoreticians, unlike experimenters, hardly ever write about the degenerative nature of the Hebbian program, it

would seem that the revolution has still not started. Almost all brain theoreticians seek support for the Hebbian program, taking the hypothesis as irrefutable fact. The truth, in our view, is very different – and lies in Ukhtomskii’s studies of the dominant and the teachings of Vinogradova on the functions of the septohippocampal system.

REFERENCES

1. G. N. Borisyyuk, R. M. Borisyyuk, Ya. B. Kazanovich, et al., “Model of the dynamics of neuron activity during information processing by the brain,” *Usp. Fizich. Nauk.*, **172**, No. 10, 1189–1214 (2002).
2. O. S. Vinogradova, *The Hippocampus and Memory* [in Russian], Nauka, Moscow (1975).
3. N. V. Golikov, “The question of local and disseminated excitation in contemporary neurophysiology,” in: *Mechanisms of Local Reactions and Propagating Excitation* [in Russian], Nauka, Leningrad (1970), pp. 5–12.
4. V. F. Kichigina, “Mechanisms of the regulation and functional importance of the theta rhythm: the role of the serotonergic and noradrenergic systems,” *Zh. Vyssh. Nerv. Deyat.*, **52**, No. 2, 195–204 (2002).
5. V. I. Kryukov, G. N. Borisyyuk, R. M. Borisyyuk, et al., *Metastable and Unstable States in the Brain* [in Russian], ONTI NtsBI, Academy of Sciences of the USSR, Pushchino (1986).
6. M. N. Livanov, “The neuronal mechanisms of memory,” *Usp. Fiziol. Nauk.*, **6**, No. 3, 66–89 (1975).
7. V. Lindsey, *Synchronization Systems in Communication and Control* [Russian translation], Sovetskoe Radio (1978).
8. S. Rose, *The Making of Memory: From Molecules to Mind* [Russian translation], Mir, Moscow (1995).
9. P. V. Simonov, *The Emotional Brain* [in Russian], Nauka, Moscow (1981).
10. A. A. Ukhtomskii, *The Dominant* [in Russian], Nauka, Moscow, Leningrad (1966).
11. A. A. Ukhtomskii, *Selected Works* [in Russian], Nauka, Moscow (1978).
12. D. G. Amaral and M. P. Witter, “The three-dimensional organization of the hippocampal formation: a review of anatomical data,” *Neurosci.*, **31**, 571–591 (1989).
13. A. Baddeley, “The episode buffer: a new component of working memory?” *Trends Cogn. Sci.*, **4**, 417–423 (2000).
14. A. Baddeley, “Exploring the central executive,” *Quart. J. Exptl. Psychol.*, **49A**, 5–28 (1996).
15. A. Baddeley and G. J. Hitch, “Working memory,” in: *Recent Advances in Learning and Motivation*, G. Bower (ed.), Academic Press, New York (1974), Vol. 8, pp. 47–90.
16. T. Brashwars-Krug, R. Shadmehr, and E. Bizzi, “Consolidation in human motor memory,” *Nature*, **382**, 252–255 (1996).
17. K. H. Britten, “Cortical neurophysiology. Attention is everywhere,” *Nature*, **382**, 497–498 (1996).
18. N. J. Cohen, J. Ryan, C. Hunt, et al., “Hippocampal system and declarative (relational) memory: summarizing the data from functional neuroimaging studies,” *Hippocampus*, **9**, 83–98 (1999).
19. N. Cowan, “Attention and memory: An integrated framework,” in: *Oxford Psychology*, Oxford University Press, New York (1995; paperback edition 1997), Series 26, Vol. xv.
20. N. Cowan, “Evolving concepts of memory, storage, selective attention and their mutual constraints within human information processing systems,” *Psychol. Rev.*, **104**, 163–191 (1988).
21. K. J. W. Craik, “Theory of the human operator in control. I. The operator as an engineering system. II. Man as an element in control system,” *Brit. J. Gen. Sect.*, **38**, 142–148 (1948).
22. F. Crick, “Thinking about the brain,” *Sci. Am.*, **241**, 181–189 (1979).

23. J. Danckert and M. A. Goodale, "Blindsight: A conscious route to unconscious vision," *Curr. Biol.*, **10**, R64–R67 (2000).
24. S. Dehaene, M. Kerszberg, and J.-P. Changeux, "A neuronal model of a global workspace in effortful cognitive tasks," *Proc. Natl. Acad. Sci. USA*, **95**, 14529–14534 (1998).
25. H. Eichenbaum, G. Schoenbaum, B. Young, and M. Bunsey, "Functional organization of the hippocampal memory system," *Proc. Natl. Acad. Sci. USA*, **93**, 13500–13507 (1996).
26. J. M. Fuster, "Cellular dynamics of network memory," *Z. Naturforsch. (C)*, **53**, 670–676 (1998).
27. J. M. Fuster, "Distributed memory for both short and long term," *Neurobiol. Learn. Mem.*, **70**, 268–274 (1998).
28. J. M. Fuster, "The prefrontal cortex – an update: time is of the essence," *Neuron*, **30**, 319–333 (2001).
29. P. S. Goldman-Rakic, "Regional and cellular fractionation of working memory," *Proc. Natl. Acad. Sci. USA*, **93**, 13473–13480 (1996).
30. P. M. Grasby, C. D. Frith, K. Friston, et al., "Activation of the human hippocampal formation during auditory-verbal long-term memory function," *Neurosci. Lett.*, **163**, 185–188 (1993).
31. H. Haken, *Principles of Brain Functioning: A Synergetic Approach to Brain Activity, Behavior and Cognition*, Springer, Berlin (1996).
32. H. Haken, "Synopsis and introduction," in: *Synergetics of the Brain*, E. Basar, et al. (eds.), Springer, New York (1983), pp. 3–25.
33. S. A. Hillyard, E. K. Vogel, and S. J. Luck, "Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence," *Phil Trans Roy. Soc. Lond.*, **B353**, 1257–1270 (1998).
34. E. R. John, *Mechanisms of Memory*, Academic Press, NY (1967).
35. M. J. Kahana, D. Seelig, and J. R. Madsen, "Theta returns," *Curr. Opin. Neurobiol.*, **11**, 739–744 (2001).
36. N. Kanwisher and E. Wojciulik, "Visual attention: insights from brain imaging," *Nat. Rev. Neurosci.*, **1**, 91–100 (2000).
37. R. W. Kentridge, C. A. Heywood, and L. Weiskrantz, "Attention without awareness in blindsight," *Proc. Roy. Soc. Lond. Biol. Sci.*, **B266**, 1805–1811 (1999).
38. W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain Res. Brain Res. Rev.*, **29**, 169–195 (1999).
39. W. Klimesch, M. Doppelmayr, J. Schwaiger, et al., "'Paradoxical' alpha synchronization in a memory task," *Brain Res. Cogn. Brain Res.*, **7**, 493–501 (1999).
40. V. I. Kryukov, "An attention model based on the principle of dominance," in: *Neurocomputers and Attention. I. Neurobiology, Synchronization and Chaos*, A. V. Holden and V. I. Kryukov (eds.), Manchester University Press, Manchester (1991), pp. 319–351.
41. V. I. Kryukov, G. N. Borisyuk, R. M. Borisyuk, et al., "Metastable and unstable states in the brain," in: *Stochastic Systems: Ergodicity, Memory, Morphogenesis*, R. L. Dobrushin, V. I. Kryukov, and A. L. Toom (eds.), Manchester University Press, Manchester, UK, (1990), pp. 226–357.
42. R. Llinas and J. P. Welsh, "On the cerebellum and motor learning," *Curr. Opin. Neurobiol.*, **3**, 958–965 (1993).
43. Z.-L. Lu, S. J. Williamson, and L. Kaufman, "Physiological measures predict behavioral lifetime of human auditory sensory memory," *Soc. Neurobiol. Abstr.*, **18**, 1212 (1992).
44. C. von der Malsburg, "Binding in models of perception and brain function," *Curr. Opin. Neurobiol.*, **5**, 520–526 (1995).
45. M.-M. Mesulam, "From sensation to cognition," *Brain*, **121**, 1013–1052 (1998).
46. R. Miller, "Cortico-hippocampal interplay: Self-organizing phase-locked loops for indexing memory," *Psychobiol.*, **17**, 115–128 (1989).
47. L. Nadel and M. Moscovitch, "Hippocampal contributions to cortical plasticity," *Neuropharmacology*, **37**, 431–439 (1998).
48. J. Newman and A. A. Grace, "Binding across time: The selective gating of frontal and hippocampal systems modulating working memory and attentional states," *Consc. Cognit.*, **8**, 196–212 (1999).
49. L. Nyberg, A. R. McIntosh, R. Cabeza, et al., "Network analysis of positron emission tomography regional cerebral blood flow: ensemble inhibition during episode memory retrieval," *J. Neurosci.*, **16**, 3753–3759 (1996).
50. K. A. Paller, "Consolidating dispersed neocortical memories: the missing link in amnesia," *Memory*, **5**, 73–88 (1997).
51. A. J. Parkin, "Human memory: the hippocampus is the key," *Curr. Biol.*, **6**, 1583–1585 (1996).
52. Z. W. Pylyshyn and R. W. Storm, "Tracking multiple independent targets: evidence for a parallel tracking mechanism," *Spat. Vis.*, **3**, 179–197 (1988).
53. G. Riedel, J. Micheau, A. G. Lam, et al., "Reversible neural inactivation reveals hippocampal participation in several memory processes," *Nat. Neurosci.*, **2**, 898–905 (1999).
54. E. Rodriguez, N. George, J. P. Lachaux, et al., "Perception's shadow: long-distance synchronization of human brain activity," *Nature*, **397**, 430–433 (1999).
55. E. Russo, "Controversy surrounds memory mechanism," *The Scientist*, **13**, 1–5 (1999).
56. J. Sarnthein, H. Petsche, P. Rappelsberger, et al., "Synchronization between prefrontal and posterior association cortex during human working memory," *Proc. Natl. Acad. Sci. USA*, **95**, 7092–7096 (1998).
57. T. B. Schillen and P. Konig, "Binding by temporal structure in multiple feature domains of an oscillatory neuronal network," *Biol. Cybern.*, **70**, 397–405 (1994).
58. B. J. Scholl and Z. W. Pylyshyn, "Tracking multiple items through occlusion: clues to visual objecthood," *Cogn. Psychol.*, **38**, 259–290 (1999).
59. T. J. Shors and L. D. Matzel, "Long-term potentiation: What's learning got to do with it?" *Behav. Brain Sci.*, **20**, 597–655 (1997).
60. A. T. Smith, K. D. Singh, and M. W. Greenlee, "Attentional suppression of activity in the human visual cortex," *Neuroreport*, **11**, 271–277 (2000).
61. C. D. Tesche and J. Karhu, "Theta oscillations index human hippocampal activation during a working memory task," *Proc. Natl. Acad. Sci. USA*, **97**, 919–924 (2000).
62. A. Treisman, "The binding problem," *Curr. Opin. Neurobiol.*, **6**, 171–178 (1996).
63. E. Tulving and H. J. Markowitsch, "Memory beyond the hippocampus," *Curr. Opin. Neurobiol.*, **7**, 209–216 (1997).
64. L. G. Ungerleider, "Functional brain imaging studies of cortical mechanisms for memory," *Science*, **270**, 769–775 (1995).
65. M. A. Uusitalo, S. J. Williamson, and M. T. Seppa, "Dynamical organisation of the human visual system revealed by lifetimes of activation traces," *Neurosci. Lett.*, **213**, 149–152 (1996).
66. M. Verfaellie and M. M. Keane, "The neural basis of aware and unaware forms of memory," *Semin. Neurol.*, **17**, 153–161 (1997).
67. O. S. Vinogradova, "Discussion," in: *Functions of Septo-Hippocampal System*, Elsevier, Amsterdam (1978), pp. 171–177.
68. O. S. Vinogradova, "Hippocampus as comparator: role of the two input and two output systems of the hippocampus in selection and registration of information," *Hippocampus*, **11**, 578–598 (2001).
69. O. S. Vinogradova, E. S. Brazhnik, and V. S. Stefekina, "Septo-hippocampal system, rhythmic oscillations and formation selection," in: *Neurocomputers and Attention. I. Neurobiology, Synchronization and Chaos*, A. V. Holden and V. I. Kryukov (eds.), Manchester University Press, Manchester (1991), pp. 129–148.
70. S. N. Watamaniuk and S. P. McKee, "Seeing motion behind occluders," *Nature*, **377**, 729–730 (1995).
71. J. M. Wolfe and S. C. Bennett, "Preattentive object files: shapeless bundles of basic features," *Vision. Res.*, **37**, 25–43 (1997).
72. D. Zipser, B. Kehoe, G. Littlewort, and J. Fuster, "A spiking network model of short-term active memory," *J. Neurosci.*, **13**, 3406–3420 (1993).